Supplement of The Cryosphere, 14, 3017–3032, 2020
https://doi.org/10.5194/tc-14-3017-2020-supplement

The Cryosphere Open Access EGU

Supplement of

Bayesian calibration of firn densification models

Vincent Verjans et al.

Correspondence to: Vincent Verjans (v.verjans@lancaster.ac.uk)

S1 Separation between the calibration and the evaluation data

The 91 sites of the dataset span a broad range of temperature and accumulation rate conditions (Table S1 and Fig. S1). As explained in the main text, our objective is to select the evaluation data (22 cores) randomly but still making it representative of (i) all climatic conditions and (ii) the ratio of GrIS to AIS sites of the dataset. We separate the 91 observed cores in three tiers of lowest, middle and highest T_{av} and we select randomly 7 cores in each tier for the evaluation data. We repeat this random selection until 5 to 10 out of the 21 cores are from GrIS, with the remainder from AIS. Finally, our dataset includes two sites that are climatic outliers with respect to the others (DML and spencer4 in Table S1) with high T_{av} and $\dot{b}$ values (Figure S1). We select randomly one of these for the evaluation data. Proceeding to the selection based on $\dot{b}$ rather than T_{av} would be similar given the strong correlation between both variables.

S2 Application of random noise in the boundary conditions

In order to let uncertainty in RACMO2 output affect the calibration process, we perturb the temperature and accumulation time series that serve as climatic forcing for the firn models. At each iteration (a round of simulations with a given parameter set at all the calibration sites) and for each individual calibration site, we randomly draw an individual climatic perturbation value c_p from a standard Normal distribution (Eq. (S1)). As such, every calibration site has a specific c_p value, which changes at each iteration. We use observed statistics of RACMO2 errors in temperature and Surface Mass Balance to determine the perturbation. For GrIS, Noël et al. (2019) report RMSE values with respect to field observations for temperature and surface mass balance flux of 2.1 K and 69 mm w.e. yr^{-1} respectively (in their Supplementary Material). Each monthly value of the RACMO2 time series is therefore perturbed by the corresponding RMSE value scaled by c_p (Eq. (S2), (S3), (S4)).

We favour this approach rather than drawing a different random perturbation at each time step. The latter method would cause perturbations of opposite signs to occur on a very short timescale, which would result in unrealistic short term climatic variability (e.g. a very warm perturbation could be immediately followed by a very cold perturbation in the next month). Also, using the same c_p value to quantify the magnitude of the perturbation for temperature and accumulation preserves the strong correlation between these variables. Warm (cold) temperature perturbations coincide with high (low) accumulation perturbations, which keeps our random perturbations physically plausible.

The part of the total accumulation perturbation attributed to each monthly time step is weighted by the actual accumulation at that time step. This attributes larger absolute noise in accumulation to high-accumulation months and lower absolute noise to low-accumulation months (Eq. (S3), (S4)).

Our approach is summarized in Eq. (S1), (S2), (S3) and (S4). These equations are applied at all iterations of the calibration process.

$$c_p \sim N(0,1) \quad \text{at all calibration sites} \tag{S1}$$

$$T_t^* = T_t + c_p \sigma_T \quad \text{at all } t \tag{S2}$$

$$\dot{b}_{tot}^* = n_{yr} c_p \sigma_{SMB} \tag{S3}$$

$$\dot{b}_t^* = \dot{b}_t + \dot{b}_{tot}^* \frac{\dot{b}_t}{\sum_t \dot{b}_t} \tag{S4}$$

where T_t and $\dot{b}_t$ are temperature and accumulation rate as computed by RACMO2 at time step t and the $*$ superscript denotes the perturbed quantity. n_{yr} is the total number of years in a given simulation, $\dot{b}_{tot}^*$ is the total accumulation perturbation applied for that simulation and σ_T and σ_{SMB} are the temperature and surface mass balance flux RMSE values (as mentioned above, σ_T = 2.1 K and σ_{SMB}= 69 mm w.e. yr^{-1} for GrIS). Note that by using a RMSE value on the surface mass balance flux, we overestimate uncertainty because the observed RMSE is mostly driven by errors in melt amounts which do not apply at the sites of our dataset, all from the dry snow zone area. For AIS, we apply the exact same process for perturbing the temperature variables. We use the RMSE value reported by van Wessem et al. (2018) and set σ_T = 1.3 K. The accumulation conditions of AIS forces the use of a slightly different method for perturbing the accumulation rate. In terms of magnitude, RACMO2 errors are much larger in coastal areas, where accumulation rates are high. In contrast, in the dry interior of the ice sheet where most of the cores of our dataset come from, the magnitude of RACMO2 errors is small due to low accumulation rates. As such, applying noise based on the ice sheet wide RMSE value would result in noise signals larger than actual accumulation values at most of our dry

sites. We thus use the average RACMO surface mass balance bias of 5% (van Wessem et al., 2018) as a proxy for one standard deviation. For AIS, Eq. (S3) and (S4) are replaced by Eq. (S5).

$$\dot{b}_t^* = \dot{b}_t + 0.05\, c_p\, \dot{b}_t \quad \text{(S5)}$$

As explained in Sect. 2.2, we also let uncertainty in fresh snow density, ρ_0, affect the calibration process by applying random perturbations to each ρ_0^t. In contrast to the climatic perturbation, the perturbation in ρ_0 can be specific to each single time step t, and the perturbation thus varies throughout the duration of a firn model simulation. Indeed, it is not unrealistic that a month with anomalously low fresh snow density is immediately followed by a month of anomalously high fresh snow density for example. We determine surface density values at each site from the firn cores of our dataset, ρ_0^{core}, and we perturb these values based on a standard deviation of 25 kg m^{-3}. This value goes in line with a typical window of local variability of 50 kg m^{-3} for ρ_0 (Reeh et al., 2005). As such, adding noise to ρ_0 simplifies to Eq. (S6).

$$\rho_{0,t}^* \sim N(\rho_0^{core}, 25) \quad \text{(S6)}$$

We emphasize that the aim of this study is not to conduct a complete sensitivity analysis of firn densification to climatic forcing and to fresh snow density. The objective of the perturbations is to let reasonable estimates of errors in those fields to be accounted for in the calibration process.

S3 Prior correlations in HL and Ar

The Arrhenius form of HL and Ar (Eq. (4) and (5)) allows us to include some correlation in the prior distributions over the parameters of these models. The values of the Arrhenius pre-exponential factors (k_0^*, k_1^*, k_0^{Ar} and k_1^{Ar}) are correlated with their corresponding activation energies (E_0, E_1 and E_g). For any given constant temperature, modelled densification rates, $\frac{d\rho}{dt}$, can be kept constant despite a change in the pre-exponential factor if the corresponding activation energy is changed accordingly and vice versa. As such, changes in these parameters can potentially compensate in the calculation of DIP values and of the likelihood function (Eq. (8)).

By enforcing constant $\frac{d\rho}{dt}$, exact compensation is ensured by the following equalities:

$$k_{0,mv}^* = k_{0,HL}^* \exp\left(\frac{E_{0,mv} - E_{0,HL}}{R\,T}\right) \quad \text{(S7)}$$

$$k_{1,mv}^* = k_{1,HL}^* \exp\left(\frac{E_{1,mv} - E_{1,HL}}{R\,T}\right) \quad \text{(S8)}$$

$$k_{0,mv}^{Ar} = k_{0,Ar}^{Ar} \exp\left(\frac{E_{g,Ar} - E_{g,mv}}{R\,T}\right) \quad \text{(S9)}$$

$$k_{1,mv}^{Ar} = k_{1,Ar}^{Ar} \exp\left(\frac{E_{g,Ar} - E_{g,mv}}{R\,T}\right) \quad \text{(S10)}$$

where HL and Ar subscripts denote the values from the original publications of the HL and Ar models, and the mv subscript denotes a modified value of the parameter. Firstly, we generate 10000 random values of temperature T in the range of annual mean temperatures covered by our dataset. Secondly, for each random temperature, we generate random values of $E_{0,mv}$, $E_{1,mv}$ and $E_{g,mv}$ in an interval of ±500 J mol^{-1} around the original values. Thirdly, we calculate the corresponding values in the pre-exponential factors from Eq. (S7), (S8), (S9) and (S10). This results in 10000 pairs of $(k_{0,mv}^*, E_{0,mv})$, $(k_{1,mv}^*, E_{1,mv})$, $(k_{0,mv}^{Ar}, E_{g,mv})$ and $(k_{1,mv}^{Ar}, E_{g,mv})$, from which we calculate correlation coefficients. The absolute values of all four correlation coefficients lie in the interval [0.75; 0.78]. We decide to fix all prior correlation coefficients to -0.75 (HL parameters, negatively correlated) and 0.75 (Ar parameters, positively correlated). The process described necessarily results in perfectly correlated $k_{0,mv}^{Ar}$ and $k_{1,mv}^{Ar}$. We also set the prior correlation between these parameters to 0.75.

We emphasize here that any other pair of *a priori* uncorrelated parameters can certainly be correlated *a posteriori* if the calibration process identifies such quantitative behaviour when the observed data is considered. This is highlighted and further discussed in Sect. S7.

S4 The likelihood function, Eq. (8)

The covariance matrices Σ_{15} and Σ_{pc} that appear in Eq. (8) are diagonal matrices with the site-specific variances on the diagonal. At each site, we take 10% of the observed $DIP15$ and 20% of the observed $DIPpc$ as the standard deviation, and the variance value is the square of the standard deviation. We take the higher value of 20% for $DIPpc$ because density errors propagate in firn models. Equation (3) shows that densification rates depend on the density value itself, resulting in error propagation through time. As such, if a model shows an offset compared to observations at 15 m depth, it is likely to show an even stronger offset at z_{pc}. Taking a higher variance alleviates the strength of this effect on the likelihood calculations by allowing a larger spread of model results compared to observed $DIPpc$ values.

The form of Eq. (8) corresponds to a normal likelihood function. This assumes that model DIP results are normally distributed around the observed values. To support this assumption, we conducted a preliminary verification of errors in $DIP15$ $(X_{15} - Y_{15})$ and $DIPpc$ $\left(X_{pc} - Y_{pc}\right)$ computed with the three original models (HL, Ar, LZ) on the entire dataset. We compute a basic Kolmogorov-Smirnov test for both sets of errors: residuals in $DIP15$ and in $DIPpc$. The resulting p-values are very large: 0.94 and 0.86 respectively. The distributions of these errors are thus in line with a normal distribution. We show the Quantiles-Quantiles plots for both sets of residuals in Figure S2. As explained in the main text, the form of Eq. (8) also assumes independence between errors in $DIP15$ and $DIPpc$, which is the reason why $DIPpc$ is calculated only from depths below 15 m. As such, observations-model discrepancies are essentially governed by parameter values of stage-1 densification for $DIP15$ and by parameter values of stage-2 densification for $DIPpc$, with little interaction between both. The same preliminary verification as mentioned above allows us to evaluate the correlation between $DIP15$ and $DIPpc$ errors for all three original models on the entire dataset. This yields correlation coefficients of 0.13, 0.60 and -0.01 for the original models HL, Ar and LZ respectively. **S5**

Convergence diagnostics

For convergence of the RWM algorithm, the chain must traverse between the peaks of the target posterior distribution multiple times. Simply examining the trace of the RWM algorithm for each parameter provides an effective way to verify this criterion. The trace is the history of accepted parameter values over the entire chain. We show this sampling history in Fig. S3. The fuzzy appearance for each parameter of each model indicates an efficient exploration of the parameter space as the samples from RWM algorithm oscillate around the posterior mode.

In addition to this, we compute the Gelman-Rubin statistic, which provides a numerical test for convergence (Gelman et al., 2013). The motivation behind this test is that if each chain (run independently) converges to the same posterior distribution, then the variances within each chain should be approximately the same. For each model (HL, Ar, LZ), we launch three different chains from different initial parameter values. For each parameter of each model, we calculate the mean within sample variance W:

$$W = \frac{s_1^2 + s_2^2 + s_3^2}{3} \quad \text{(S1)}$$

where s^2 denotes the variance of an individual chain. We then calculate the between sample variance:

$$B = \frac{n}{3-1}\sum_{i=1}^{3}\left(\bar{\theta}_i - \bar{\bar{\theta}}\right)^2 \quad \text{(S2)}$$

where n denotes the number of iterations within each chain, $\bar{\theta}_i$ the mean parameter value within each chain and $\bar{\bar{\theta}}$ is the mean of $(\bar{\theta}_1, \bar{\theta}_2, \bar{\theta}_3)$. From there, the estimate of the variance of the posterior distribution is given by:

$$\hat{\sigma}^2 = \frac{n-1}{n}W + \frac{1}{n}B \quad \text{(S3)}$$

And the Gelman-Rubin statistic is defined as:

$$R = \sqrt{\frac{\hat{\sigma}^2}{W}} \quad \text{(S4)}$$

Large values of R indicate that estimates of θ values between the different chains are significantly different. With more iterations, the chains progressively converge to the same stationary distributions and the estimates of θ become similar, resulting in values of R close to 1. We reach $R < 1.1$ for all parameters, which proves adequate convergence (Gelman et al., 2013). Two parameters of the LZ model needed a larger number of iterations to reach $R < 1.1$.

S6 Normal approximation to the posterior

The ensembles of parameter combinations obtained for each model provide large samples, representative of the posterior probability distributions over their respective parameter space. The most efficient way to assess parameter-related uncertainty is to run a model with a high number of random parameter combinations from these ensembles, which is demonstrated in Sect. 3. However, this means that for any firn modelling study, access must be easy to such posterior ensembles or an MCMC algorithm must be re-executed. To circumvent these practical difficulties, it is approximately correct to sample random parameter combinations from a multivariate normal distribution centred about the mean of the posterior ensemble and with covariance matrix set to the posterior ensemble covariance matrix. This is commonly referred to as a normal approximation to the posterior (Gelman et al., 2013). Table S2 provides both the posterior mean and posterior covariance for the HL, Ar and LZ models.

We assess how random samples from the normal approximations compare to samples from the posterior ensembles in Fig. S4. Posterior samples and the normal approximations are very similar, with correlations only slightly less well captured in the tails of the distributions. As a consequence, the normal approximation results in a slight overestimation of uncertainty and thus conservative estimates of uncertainty. This has been confirmed by additional model simulations with values sampled from the normal approximations (not shown).

S7 Posterior correlation between parameters

The joint posterior distributions for the parameters of each model also allow us to analyse the models' internal structure, i.e. the correlation between their different parameters. The full correlation matrices are given in Fig. S5. In HL, the strongest correlation coefficients r are unsurprisingly found for the pairs of pre-exponential factor and activation energy governing densification in stage-1 (k_0^* and E_0) and in stage-2 (k_1^* and E_1) with r of 0.91 and 0.92 respectively. Higher activation energies (E_0 and E_1) imply stronger thermal barriers to the densification process and thus slower densification, and the pre-exponential factors (k_0^* and k_1^*) counter-balance the effect to still match observed DIP values. In the same way, the activation energies are negatively correlated with their respective accumulation rate exponent (a and b), but more moderately (r values of approximately -0.5). The negative correlation of -0.28 between a and b themselves might be linked to the density at 15 m being the lower boundary and the upper boundary condition for the calculation of $DIP15$ and $DIPpc$ respectively. Higher values of a tend to cause lighter firn at 15 m depth. Lower E_0 values compensate for this effect on $DIP15$ because the shallow firn densifies faster due to its greater sensitivity to temperature variations. The lighter 15 m depth density also affects $DIPpc$, and lower values of b compensate for this by enhancing the densification rate, which explains the negative correlation between a and b. In Ar, the interpretation is more challenging due to the use of a same activation energy in both stages. There is a strong correlation between the activation energy E_g and both pre-exponential factors k_0^{Ar} (r = -0.89) and k_1^{Ar} (r = -0.90), for the same reason as in HL. As such, this induces a strong positive correlation between the latter parameters (r = 0.76). The negative correlation between α and k_1^{Ar} (r = -0.41) is more surprising because these parameters apply to different stages, but it reveals an interesting pattern. Higher temperatures raise densification rates at warmer sites, where accumulation rates are also higher thus further amplifying this effect. Higher accumulation rates nevertheless cause light recently deposited firn to be buried rapidly, which may cause lower density firn governed by the fast stage-1 densification to extend below 15 m. To avoid underestimation of $DIPpc$ at such sites, stage-1 densification rates must remain low enough but there is no possibility for adjusting a stage-1 specific activation energy. Lower α values generate this effect while only marginally affecting densification at colder low-accumulation sites. Thus, high k_1^{Ar} without a complementary lower α would cause $DIPpc$ underestimation at warm and high accumulation sites. We note here that, through the calibration process, the data enhanced the prior correlations we assigned in the HL and Ar models. Analysis of correlation coefficients in LZ is less straightforward because its governing equations, Eq. (6), are less interpretable than the plain Arrhenius relationship of HL and Ar. Still, we highlight some correlated pairs of parameters. As could be expected from Eq. (6), lz_a and lz_b are negatively correlated (r = -0.80). Also, the independent term of stage-1 densification lz_{11} is strongly correlated with the corresponding temperature-related parameter (lz_{13}, r = 0.94). The same is valid for stage-2 densification between lz_{21} and lz_{23} (r = 0.90). The positive correlation between lz_{12} and lz_{22} (r = 0.74) is discussed in the main text.

Tables

Site	Lat	Lon	Core depth [m]	Year	Mean $\dot{b}$ [m w.e. yr^{-1}]	Mean T [°C]	ρ_0 [kg/m^3]	$DIP15$ [m]	Var $DIP15$ [m^2]	$DIPpc$ [m]	Var $DIPpc$ [m^2]
EGRIP	75.63	-35.98	20.1	2017	0.113	-29.0	285	7.816	0.611	/	/
Summit *	72.58	-38.47	22.1	2017	0.205	-28.4	330	7.500	0.562	/	/
id359	73.94	-37.63	102.4	1993	0.124	-28.8	240	6.708	0.450	11.456	5.250
id369	75.00	-30.00	19.9	1997	0.135	-27.6	335	7.454	0.556	/	/
id373	75.25	-37.62	100.8	1993	0.106	-29.5	275	7.826	0.612	12.372	6.123
id385	76.00	-43.49	109.8	1995	0.124	-29.3	315	7.857	0.617	13.186	6.955
id423 *	76.62	-36.40	143.2	1993	0.093	-29.1	310	7.716	0.595	10.666	4.550
id514	77.25	-49.22	119.6	1995	0.163	-28.3	300	7.575	0.574	13.217	6.987
id531 *	77.45	-51.06	75.0	2009	0.198	-27.4	320	7.434	0.553	/	/
id534	80.00	-41.14	96.0	1994	0.105	-28.4	335	7.811	0.610	11.345	5.148
Basin8	69.80	-36.49	29.8	2003	0.350	-25.6	300	7.396	0.547	/	/
D2	71.80	-46.34	101.3	2003	0.421	-23.4	370	7.051	0.497	14.097	7.949
D4	71.39	-43.94	143.9	2003	0.390	-24.6	300	7.394	0.547	12.770	6.523
HumboldtM *	78.47	-56.98	141.9	1995	0.384	-24.8	280	8.062	0.650	10.947	4.794
NASAE1 *	74.98	-29.97	19.9	1997	0.135	-27.6	340	7.394	0.547	/	/
spencer6 *	72.57	-37.62	82.3	1994	0.176	-29.0	360	4.889	0.239	/	/
spencer16	71.75	-40.75	15.0	1954	0.289	-27.0	340	7.216	0.521	/	/
spencer17	77.95	-39.18	60.0	1973	0.080	-29.3	300	5.002	0.250	7.781	2.421
spencer66 *	70.75	-35.96	109.0	1987	0.247	-27.3	300	7.340	0.539	14.852	8.823
spencer67	70.63	-35.83	128.6	1988	0.262	-27.0	325	7.098	0.504	14.114	7.968
spencer68 *	70.65	-37.48	105.6	1988	0.263	-26.9	325	7.172	0.514	14.505	8.416
spencer69	70.67	-38.79	24.8	1988	0.252	-27.1	305	7.184	0.516	/	/
spencer70	70.64	-39.62	100.1	1988	0.262	-27.0	290	6.772	0.459	14.026	7.869
spencer71	71.76	-35.87	77.8	1988	0.203	-28.2	275	7.043	0.496	13.094	6.858
spencer72	71.48	-35.88	25.7	1988	0.207	-28.0	330	7.223	0.522	/	/
spencer73	71.15	-35.85	70.8	1988	0.214	-27.7	340	7.230	0.523	/	/
spencer74	70.85	-35.85	26.2	1988	0.264	-26.9	330	7.087	0.502	/	/
SouthPole	-90.00	0.00	122.9	2001	0.055	-47.8	325	7.613	0.580	22.312	19.913
Newall	-77.58	162.50	111.1	1989	0.043	-31.2	305	7.160	0.513	4.132	0.683
Berkner *	-79.61	-45.72	178.2	1995	0.124	-28.3	345	6.255	0.391	9.658	3.731
DML *	-71.41	-9.92	78.2	2007	0.902	-20.6	410	6.037	0.364	10.228	4.185
id9	-76.77	-101.74	111.6	2012	0.313	-24.7	470	6.194	0.384	12.119	5.875

id10	-76.95	-121.22	62.0	2011	0.213	-28.4	355	6.947	0.483	/	/
id11	-77.06	-89.14	114.5	2001	0.346	-26.5	415	5.879	0.346	11.201	5.019
id12	-77.61	-92.25	67.8	2001	0.301	-27.8	350	6.019	0.362	/	/
id13	-77.68	-124.00	59.3	2000	0.155	-28.2	350	6.411	0.411	/	/
id14	-77.76	153.38	97.1	2006	0.048	-44.6	360	6.833	0.467	17.516	12.272
id15 *	-77.84	-102.91	70.7	2001	0.486	-25.1	415	5.853	0.343	/	/
id17	-77.88	158.46	98.5	2006	0.058	-41.1	350	6.419	0.412	11.687	5.464
id18	-77.96	-95.96	57.4	2010	0.354	-28.0	335	6.752	0.456	/	/
id19	-78.08	-120.08	57.8	2000	0.171	-27.7	315	6.253	0.391	/	/
id20	-78.12	-95.65	70.5	2001	0.324	-27.7	385	6.265	0.393	/	/
id22 *	-78.33	-124.48	59.9	2000	0.152	-27.7	285	6.509	0.424	8.989	3.232
id24	-78.43	-115.92	59.8	2000	0.318	-27.8	390	6.295	0.396	/	/
id26	-78.73	-111.50	60.7	2000	0.329	-27.8	350	6.427	0.413	/	/
id28	-79.04	149.68	100.1	2006	0.040	-44.6	405	6.703	0.449	15.584	9.714
id29 *	-79.13	-122.27	63.1	2000	0.127	-27.8	300	6.507	0.423	9.926	3.941
id30	-79.16	-104.97	72.7	2001	0.306	-28.7	400	5.921	0.351	/	/
id33	-79.38	-111.24	104.8	2000	0.239	-28.2	370	6.159	0.379	12.943	6.701
id35 *	-79.48	-112.09	160.0	2011	0.162	-28.0	460	6.181	0.382	11.824	5.592
id39	-80.62	-122.63	57.5	1999	0.094	-25.9	370	6.253	0.391	/	/
id43	-81.20	-126.17	48.3	1999	0.070	-24.5	325	6.268	0.393	4.975	0.990
id46	-82.00	-110.01	62.2	2002	0.180	-27.8	340	6.161	0.380	/	/
id48	-83.50	-104.99	61.7	2002	0.220	-31.0	360	6.098	0.372	/	/
id49 *	-84.40	140.63	50.1	2007	0.023	-45.4	340	6.886	0.474	/	/
id50	-85.00	-105.00	44.9	2002	0.157	-36.3	360	6.422	0.412	/	/
id51	-85.78	145.72	41.7	2007	0.033	-46.1	310	6.767	0.458	/	/
id52 *	-86.50	-107.99	71.6	2002	0.147	-38.8	340	6.882	0.474	/	/
id53	-86.84	95.31	20.8	2003	0.042	-53.3	355	6.535	0.427	/	/
id54 *	-88.00	-107.98	54.1	2002	0.133	-41.4	355	7.009	0.491	/	/
id55	-88.51	178.53	99.3	2007	0.081	-48.2	320	6.880	0.473	/	/
id56	-89.93	144.39	139.5	2002	0.080	-48.6	345	6.319	0.399	25.046	25.092
spencer1	-80.00	-120.00	307.0	1968	0.120	-27.2	350	6.987	0.488	10.314	4.255
spencer4	-66.72	113.18	200.9	1989	1.060	-22.0	380	7.848	0.616	12.847	6.602
spencer5	-74.50	123.17	49.5	1980	0.037	-51.8	345	8.262	0.683	/	/
spencer7	-85.25	166.50	79.9	19997	0.028	-39.7	305	7.003	0.490	8.202	2.691
spencer8	-66.77	112.80	180.0	1997	0.488	-22.7	385	7.385	0.545	10.640	4.528
spencer22	-73.60	-12.43	25.5	1996	0.220	-22.5	380	3.920	0.154	/	/
spencer25	-74.02	-12.02	26.5	1996	0.171	-30.7	390	5.412	0.293	/	/

spencer29 *	-75.00	2.00	20.6	1996	0.072	-42.9	320	7.602	0.578	/	/
spencer33	-70.68	44.32	123.5	1978	0.114	-33.1	385	6.385	0.408	7.022	1.972
spencer34 *	-70.68	44.32	109.0	1978	0.114	-33.1	375	6.161	0.380	6.909	1.909
spencer61	-73.10	39.75	99.7	1978	0.069	-42.3	360	7.005	0.491	16.245	10.556
spencer62 *	-71.18	45.97	100.2	1997	0.091	-38.2	395	7.049	0.497	16.344	10.686
spencer76	-90.00	0.00	122.1	1997	0.055	-47.8	360	4.906	0.241	25.586	26.185
spencer77	-75.00	147.00	15.8	1961	0.042	-46.1	385	7.184	0.516	/	/
spencer78 *	-74.00	143.00	16.0	1961	0.043	-45.5	375	7.205	0.519	/	/
spencer79	-73.00	142.00	15.7	1961	0.057	-44.0	325	7.148	0.511	/	/
spencer80	-73.00	141.00	16.0	1961	0.057	-44.0	355	6.876	0.473	/	/
spencer81	-72.00	140.00	16.9	1961	0.080	-42.7	335	6.936	0.481	/	/
spencer82 *	-71.00	139.00	15.6	1961	0.120	-41.6	375	6.848	0.469	/	/
spencer83	-72.00	143.00	15.7	1961	0.087	-41.3	405	6.796	0.462	/	/
spencer84	-72.00	146.00	16.2	1961	0.086	-40.9	410	6.876	0.473	/	/
spencer85	-72.00	148.00	15.9	1961	0.096	-40.2	360	6.745	0.455	/	/
spencer86	-72.00	151.00	15.8	1961	0.103	-39.7	400	6.963	0.485	/	/
spencer87	-72.00	154.00	15.9	1961	0.130	-38.0	355	6.430	0.414	/	/
spencer88	-72.00	156.00	15.7	1961	0.130	-37.6	395	7.050	0.497	/	/
spencer89	-72.00	159.00	15.7	1961	0.115	-35.7	370	6.665	0.444	/	/
spencer90	-83.47	138.80	340.5	1994	0.020	-45.2	420	/	/	10.046	4.037
spencer91	-83.47	-138.80	47.0	1987	0.058	-27.1	295	7.037	0.495	3.530	0.499
spencer92	-78.47	106.80	179.3	1996	0.022	-54.6	360	8.790	0.773	20.368	16.594

Table S1. The 91 firn core dataset used in this study. * symbols indicate the core is part of the evaluation data. Lat and Lon designate latitude and longitude respectively. Year indicates the year of drilling of the core. $\dot{b}$ is the accumulation rate. T is the temperature. Values for both $\dot{b}$ and T are computed from the RACMO2 model. ρ_0 is the surface density boundary condition that was derived individually for each core by extrapolating density measurements until the surface (random noise is added to ρ_0 as discussed in Sect. S2). Var designates the site-specific variance used for the terms of Σ_{15} and Σ_{pc} (see Text S4 for their calculation). The core spencer90 has only a single density measurement above 15 m depth and its $DIP15$ is discarded.

	Parameters	Posterior mean	Posterior covariance matrix
HL	$k_0^*, k_1^*, E_0, E_1, a, b$	$[16.7, 649, 10760, 21000, 0.88, 0.66]$	$\begin{bmatrix} 34.4 & 40.2 & 4500 & 324 & -0.0685 & -0.0195 \\ 40.2 & 44000 & 618 & 161000 & 1.087 & -3.670 \\ 4502 & 618 & 710000 & 7080 & -29.95 & 1.94 \\ 324 & 1610000 & 7080 & 694000 & 7.86 & -27.51 \\ -0.0685 & 1.087 & -29.95 & 7.86 & 0.0051 & -0.0012 \\ -0.0195 & -3.670 & 1.94 & -27.51 & -0.0012 & 0.0036 \end{bmatrix}$
Ar	$k_0^{Ar}, k_1^{Ar}, E_g, \alpha, \beta$	$[0.080, 0.028, 40900, 0.78, 0.69]$	$\begin{bmatrix} 5.62\times10^{-4} & 1.55\times10^{-4} & -12.66 & 9.65\times10^{-5} & -3.23\times10^{-4} \\ 1.55\times10^{-4} & 7.41\times10^{-5} & -4.64 & -2.04\times10^{-4} & 1.05\times10^{-4} \\ -12.66 & -4.64 & 360000 & 11.0 & 4.67 \\ 9.65\times10^{-5} & -2.04\times10^{-4} & 11.0 & 0.00330 & -0.00101 \\ -3.23\times10^{-4} & 1.05\times10^{-4} & 4.67 & -0.00101 & 0.00312 \end{bmatrix}$
LZ	$lz_a, lz_b, lz_{11}, lz_{12}, lz_{13}, lz_{21}, lz_{22}, lz_{23}$	$[7.56, -2.091, -14.71, 7.269, -1.019, -1.513, 6.0203, -0.09127]$	$\begin{bmatrix} 5.27 & -0.198 & -1.20 & -1.68 & -0.0239 & 0.00553 & -0.0606 & 0.00413 \\ -0.198 & 0.0116 & 0.218 & -0.0612 & 0.0134 & -0.0158 & -0.00229 & -7.37\times10^{-4} \\ -1.20 & 0.218 & 14.6 & -3.96 & 0.801 & 0.368 & 0.354 & 0.0129 \\ -1.68 & -0.0612 & -3.96 & 13.3 & -0.309 & -0.0850 & 5.40 & 0.0166 \\ -0.0239 & 0.0134 & 0.801 & -0.309 & -0.0502 & -0.0173 & 0.0252 & -4.42\times10^{-4} \\ 0.00553 & -0.0158 & 0.368 & -0.0850 & -0.0173 & 0.446 & -0.429 & 0.0131 \\ -0.0606 & -0.00229 & 0.354 & 5.40 & 0.0252 & -0.429 & 3.94 & -2.59\times10^{-4} \\ 0.00413 & -7.37\times10^{-4} & 0.0129 & 0.0166 & -4.42\times10^{-4} & 0.0131 & -2.59\times10^{-4} & 4.80\times10^{-4} \end{bmatrix}$

Table S2. The posterior means and covariance matrices for the free parameters of HL, Ar and LZ. These statistics can be used to generate random parameter combinations following a normal approximation.

Figures

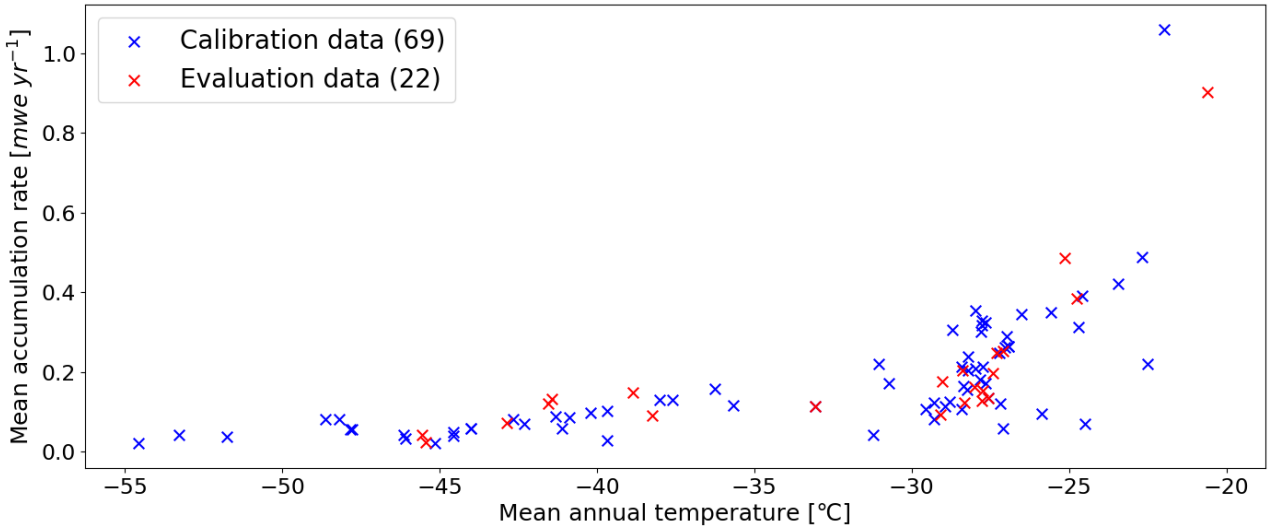

Figure S1. Climatic conditions at the 91 sites of the dataset

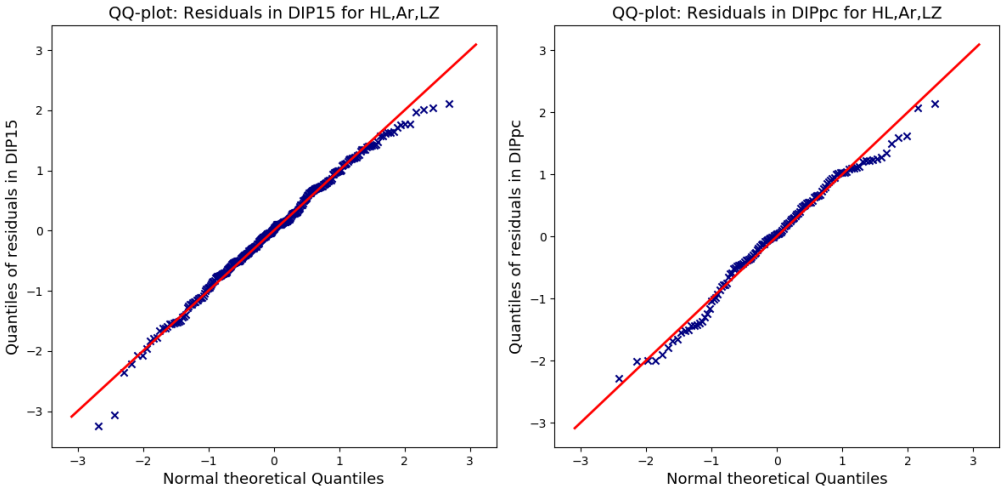

Figure S2. Quantiles-Quantiles plots for the errors of the three original models (HL, Ar, LZ) computed on the entire dataset. The alignment of the points along the red line informs about the fit to a normal distribution.

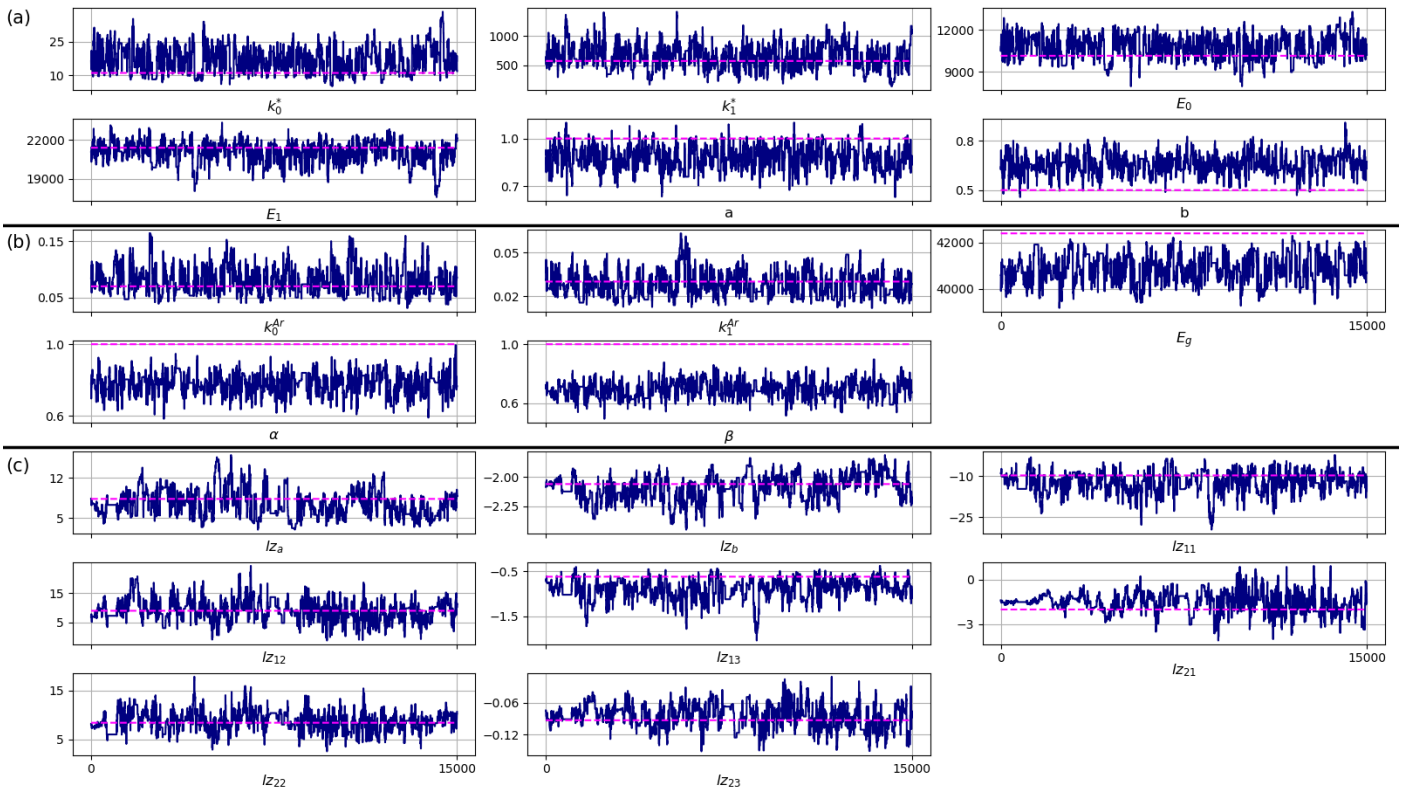

Figure S3. Sampling chains of each parameter for (a) HL, (b) Ar, (c) LZ. The x-axis displays the iteration number, the y-axis displays the parameter value. The dashed pink line shows the value of the original model, which is also the starting point of each chain.

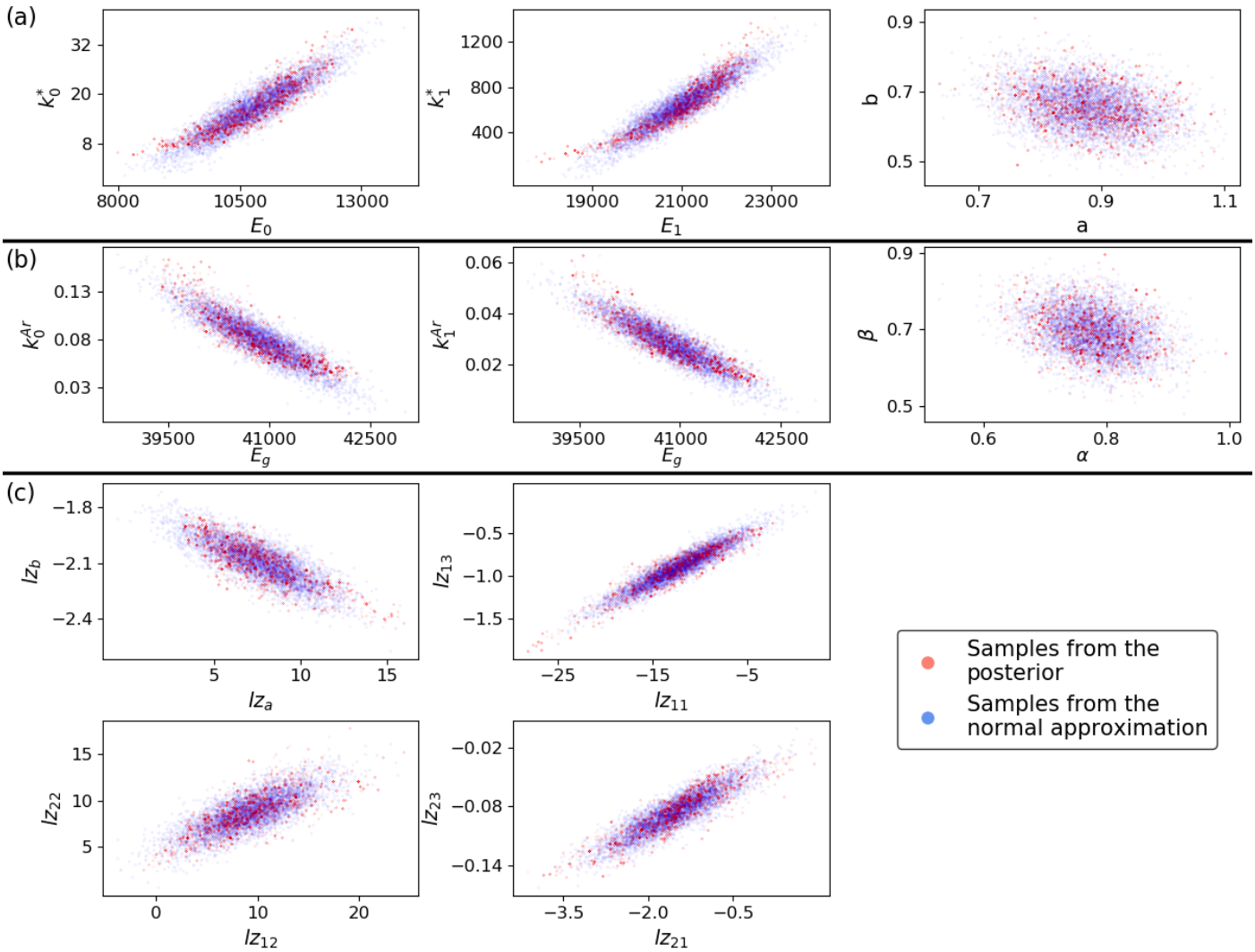

Figure S4. Evaluation of the normal approximations to the posterior distributions for (a) HL, (b) Ar, (c) LZ. Where possible, correlated parameters share a same graph.

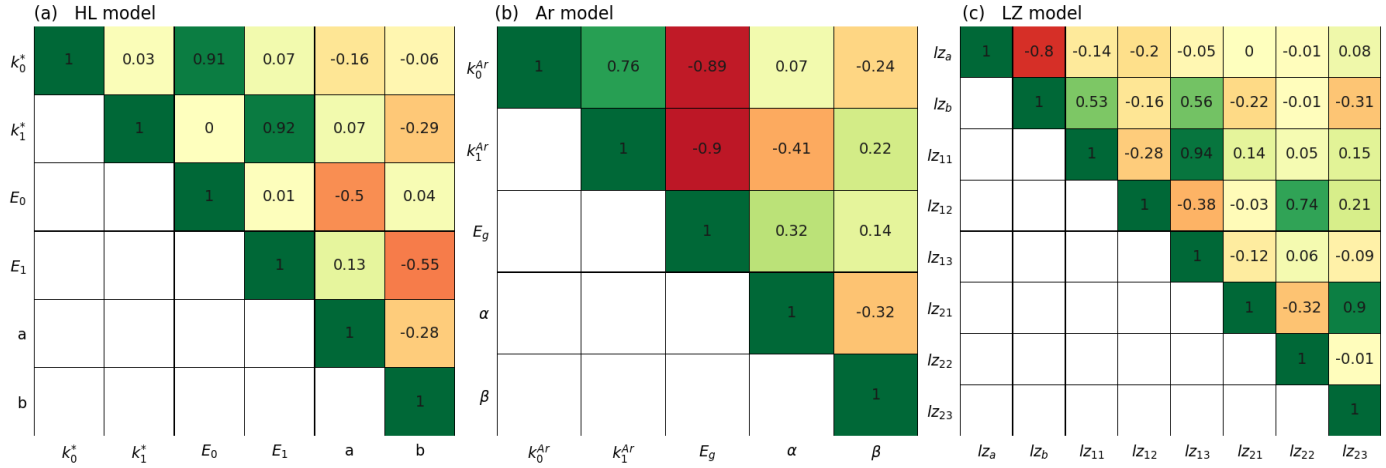

Figure S5. Posterior correlation matrices.

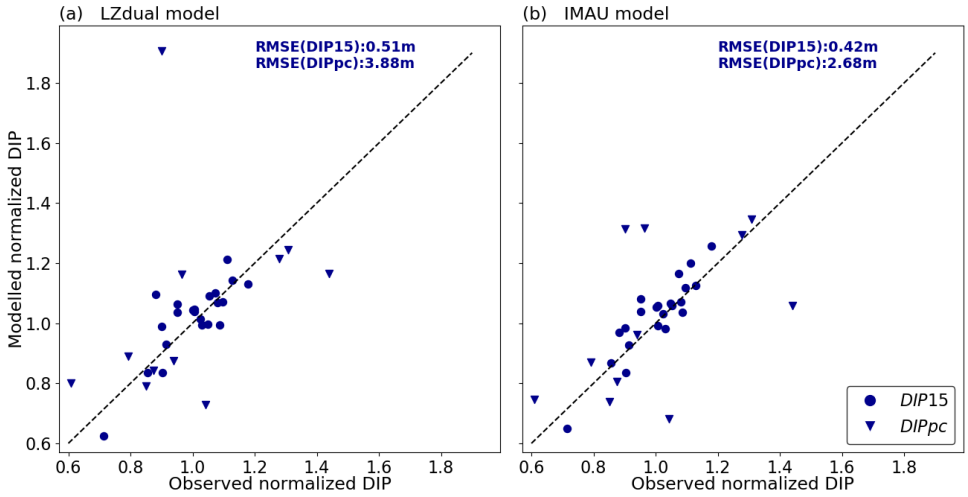

Figure S6. Comparison of evaluation data *DIP* with model results for the LZ dual and IMAU models.

References

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A. and Rubin, D.: Bayesian Data Analysis, Third Edit., CRC Press Taylor & Francis Group, Boca Raton., 2013.

Noël, B., van de Berg, W. J., Lhermitte, S. and van den Broeke, M. R.: Rapid ablation zone expansion amplifies north Greenland mass loss, Sci. Adv., 5(9), eaaw0123, doi:10.1126/sciadv.aaw0123, 2019.

Reeh, N., Fisher, D. A., Koerner, R. M. and Clausen, H. B.: An empirical firndensification model comprising ice lenses, in Annals of Glaciology, vol. 42, pp. 101–106., 2005.

van Wessem, J. M., Jan Van De Berg, W., Noël, B. P. Y., Van Meijgaard, E., Amory, C., Birnbaum, G., Jakobs, C. L., Krüger, K., Lenaerts, J. T. M., Lhermitte, S., Ligtenberg, S. R. M., Medley, B., Reijmer, C. H., Van Tricht, K., Trusel, L. D., Van Ulft, L. H., Wouters, B., Wuite, J. and Van Den Broeke, M. R.: Modelling the climate and surface mass balance of polar ice sheets using RACMO2 - Part 2: Antarctica (1979-2016), The Cryosphere, 12(4), 1479–1498, doi:10.5194/tc-12-1479-2018, 2018.