Using a process-based dendroclimatic proxy system model in a data assimilation framework: a test case in the Southern Hemisphere over the past centuries

Jeanne Rezsöhazy[1,2], Quentin Dalaiden[1], François Klein[1], Hugues Goosse[1], and Joël Guiot[2]

[1]Université catholique de Louvain (UCLouvain), Earth and Life Institute (ELI), Georges Lemaître Centre for Earth and Climate Research (TECLIM), Place Louis Pasteur, B-1348 Louvain-la-Neuve, Belgium
[2]Aix Marseille University, CNRS, IRD, INRA, College de France, CEREGE, Aix-en-Provence, France

Correspondence: J. Rezsöhazy (jeanne.rezsohazy@uclouvain.be)

Abstract. Currently available data assimilation-based reconstructions of past climate variations have only used statistical proxy system models to make the link between the climate model outputs and the indirect observations from tree rings. However, the linearity and stationarity assumptions of the statistical approach may have limitations. In this study, we incorporate the process-based dendroclimatic model MAIDEN into a data assimilation procedure, using as a test case the reconstruction of near-surface air temperature, precipitation and winds in the mid-latitudes of the Southern Hemisphere over the past 400 years. We compare our results with a data assimilation approach including a linear regression as a proxy system model for tree-ring width proxies. Overall, when compared to instrumental data, the reconstructions using MAIDEN as a proxy system model offer a skill equivalent to the experiment using the regression model. However, knowing the advantages that a process-based model can bring and the improvements that can still be made with MAIDEN, those results are promising.

1 Introduction

Indirect observations of climate from proxies such as tree-ring width (hereafter TRW) or isotopic content in ice cores and coral inform on past climate variability beyond the instrumental era (Jones et al., 2009). Those proxy records have been used in numerous studies devoted to climate variability over the last millennium. This has been classically achieved by using empirical (i.e., statistical) relationship between climate variables of interest (e.g., surface air temperature) and proxy observations (e.g. Mann et al., 1999; Luterbacher et al., 2004; Mann et al., 2008; Büntgen et al., 2011; Neukom et al., 2011; Tingley and Huybers, 2013; Wilson et al., 2016). Tree rings represent one of the main source of information for those studies as it is the most available and spatially distributed proxy to reconstruct past climates at high temporal resolution (e.g. Fritts, 1976; Jones et al., 2009; Cook et al., 1999; Esper et al., 2002; Wilson et al., 2016; Anchukaitis et al., 2017; Esper et al., 2018).

20   Although the network of TRW records is well developed, mostly in the extratropical latitudes of the Northern Hemisphere (PAGES 2k Consortium, 2017), these records only provide insight about past climates at a specific location. In order to spread the local information to the large spatial scale, statistical methods have been developed to provide spatially gridded climate reconstructions of the past centuries, generally on the basis of linear regressions between the network of proxy records and the climate field of interest, calibrated over the instrumental period (e.g. Fritts et al., 1971; Mann et al., 1998, 2007; Cook et al.,

25   2010; Wang et al., 2015; Anchukaitis et al., 2017). ~~The~~ In this so-called transfer function approach (Fritts et al., 1971), the climate field of interest is expressed as a function of the tree-ring proxy, i.e. in the inverse mode (as opposed to the natural direction where tree growth responds to climate). However, while this assumes an invertibility of the relationship between climate and tree growth, this may not always be the case. Additionally, the use of linear regression is based on the assumption that the relationship between climate and tree growth is linear and stationary over time. However, many studies have highlighted

30   the shortcomings associated with these assumptions for TRW records (Wilmking et al., 2020; Briffa et al., 1998; Wilson and Elling, 2004; Wilson et al., 2007; D'Arrigo et al., 2008; Guiot et al., 2014; Babst et al., 2018).

On the other hand, global climate model simulations offer a complete spatial picture of the past climate variability. Nevertheless, these simulations, if used alone, cannot follow the observed climate evolution because of the large ~~impact of natural~~ contribution of internal climate variability that is unpredictable over those timescales and thus cannot be reproduced by models,

35   and because of potential biases. Over the last decades, data assimilation (DA) has been increasingly used in paleoclimatology to bring together the information from both indirect climate records from proxies and climate models to provide a more complete picture of the climate changes (Goosse et al., 2010; Widmann et al., 2010; Goosse et al., 2012; Steiger et al., 2014; Franke et al., 2017; Steiger et al., 2018; Tardif et al., 2019).

Initially, the DA-based reconstructions were developed using temperature reconstructions derived from the statistical re-

40   lationship between temperature and proxy observations, and not the proxy time series themselves (e.g. Crespin et al., 2009; Goosse et al., 2010). More recently, the use of forward proxy system models (hereafter, PSMs) has emerged in order to directly assimilate the proxy time series. This gives the opportunity to consider the relationship between the proxy and the climate in the natural direction, as opposed to inverse models, and to take into account the combined influence of multiple climate variables on the proxy data. Specifically, the PSMs – one specific form of the observation operator in the DA framework

45   – make the link between the outputs of the climate model included in the DA procedure and the assimilated proxy observations, for example TRW (Evans et al., 2013; Dee et al., 2016). To date, reconstructions based on data assimilation using observed tree-ring data have only included statistical PSMs (uni- or multivariate regressions; e.g. Franke et al., 2017; Steiger et al., 2018; Tardif et al., 2019). Contrary to statistical models, process-based dendroclimatic PSMs are able to account for the complexity of the relationship between climate and tree-ring proxy data by explicitly simulating the biologi-

50   cal processes governing the climate dependency of tree-ring growth. In this regard, they may overcome some of the limitations of statistical models. However, so far, they have never been used in a DA procedure with actual tree-ring proxy data, but only with pseudoproxies ~~(Dee et al., 2016; Acevedo et al., 2017; Steiger and Smerdon, 2017)~~(i.e., synthetic proxy data; Dee et al., 2016; Acevedo et al., 2017; Steiger and Smerdon, 2017). In these studies, the VS-Lite model (Tolwinski-Ward et al., 2011) ~~has shown to significantly~~ is forced by climate model outputs in order to provide synthetic tree-ring proxy data. These

synthetic proxy data are then incorporated in a DA framework, providing a controlled environment to evaluate the DA procedure and resulting reconstruction skill. In those experiments, using the non-linear VS-Lite model as a PSM for tree-ring data has the potential to improve the quality of the reconstruction compared to the traditional statistical linear approach of tree-ring models. However, VS-Lite cannot be considered as entirely process-based as it does not include any biological processes to drive tree growth, but is rather based on the implementation of the principle of limiting factors with threshold growth response functions (Tolwinski-Ward et al., 2011).

In this study, we introduce for the first time a process-based dendroclimatic model in ~~an offline~~ a data assimilation procedure with actual tree-ring width records. Among all the available mechanistic tree-growth models (e.g Misson, 2004; Dufrêne et al., 2005; Vaganov et al., 2006; Drew et al., 2010; Tolwinski-Ward et al., 2011), we chose to work with the ecophysiological model MAIDEN (Modelling and Analysis In DENdroecology; Misson, 2004). MAIDEN includes the atmospheric CO_2 concentration among other inputs, which represents a major advantage for taking into account the potential effect of the recent exponential increase of CO_2 on tree growth (Myhre et al., 2013; Boucher et al., 2014). In Rezsöhazy et al. (2021), the MAIDEN model has been successfully applied over the last century to the PAGES2k TRW database ~~using the calibration protocol developed~~ by comparing the simulated annual quantity of carbon allocated to the stem with the observed tree-ring widths (both normalized), using as inputs of the model high-resolution climate data. The model has been calibrated over the second half of the 20th century using the Bayesian procedure with Markov Chain Monte Carlo sampling elaborated in Rezsöhazy et al. (2020), ~~highlighting~~ and successfully verified over the first half of the century at 20% of the PAGES2k tree-ring sites. The studies have highlighted the potential of MAIDEN to be used as a PSM for DA-based reconstruction of past climate variability, compared to a simple model like VS-Lite (Tolwinski-Ward et al., 2011).

Here, we perform a comparative analysis on the impact of using a simple regression model or a process-based dendroclimatic model as PSMs for TRW records in the DA framework. The goal is to evaluate if and how a complex process-based model of tree growth like MAIDEN can contribute to the improvement of large-scale DA-based reconstructions of past climate variability compared to the regression model classically used so far. Because MAIDEN has shown to be skillful in different regions of the Southern Hemisphere (SH; Rezsöhazy et al., 2021), we focus on that region and provide climate reconstruction of near-surface air temperature, precipitation and winds in the mid-to-high latitudes of the SH over the past 400 years. The Southern Hemisphere also constitutes an interesting study case, as few reconstructions have been performed in this region so far (Neukom et al., 2011, 2014). The DA experiments are based on both TRW and ice core ($\delta^{18}O$ and snow accumulation) proxy data to ensure the consistency of the reconstructed large-scale circulation pattern by assimilating proxy records at mid and high (i.e., Antarctica) latitudes of the Southern Hemisphere and thus avoiding to only reconstruct small-scale features. They both represent the best available continental annually resolved proxies in the SH (PAGES 2k Consortium, 2017). While the TRW records require the use of a PSM, the ice core records can be directly compared with the outputs of an isotope-enabled climate model.

First, DA experiments are performed and evaluated against state-of-the-art gridded climate datasets over the last century to evaluate the impact of the assimilation of the tree-ring data and the associated PSMs on the skill of the reconstruction. In Sect. 4.1, we compare the performance of the DA-based reconstructions using a regression- or a process-based dendroclimatic PSM

in our framework over the last century. Sensitivity experiments are then performed in Sect. 4.2 to identify the contribution of tree-ring width proxies in our data assimilation framework using a regression- or a process-based dendroclimatic PSM as well as to quantify the impact of the uncertainty of the records. Secondly, the DA-based reconstructions performed with a statistical or process-based dendroclimatic PSM are compared to each other over the past four centuries in Sect. 4.3. To match the seasonality of tree growth in the Southern Hemisphere, all the analysis in this study are performed on a July-June year (e.g the year 1992 corresponds to July 1992-June 1993), instead of the usual January-December year, suited for trees growing in the Northern Hemisphere.

2   Data assimilation

In paleoclimatology, the objective of data assimilation is to optimally combine the information from the physics of the climate system as included in climate models and indirect observations of climate from proxies (Kalnay, 2003; Goosse et al., 2010; Widmann et al., 2010; Steiger et al., 2014; Franke et al., 2017; Hakim et al., 2016). More specifically, the DA procedure is based on the Bayes theorem (van Leeuwen, 2009): starting from the prior estimate of the state of the climate system provided by the model, DA produces a reconstruction of the climate system as accurate as possible (i.e, the posterior) given the information provided by the climate observations. The updated estimate of the climate system is often called a reanalysis. One of the advantages of DA is its ability to propagate the information brought by the observations both in space and between the climate variables, by relying on the spatial covariance of a specific variable and on the covariance between different variables in the model. In other words, DA gives the opportunity to reconstruct the climate at places where there is no proxy sites, and for variables for which we do not have any observations to assimilate.

In this study, we use an offline DA approach, by opposition with the online DA approach. The offline approach samples a pre-existing ensemble of climate model simulations to build the prior state of the climate system. This has the great advantage of being considerably computational time saving. In this configuration, the different years of the sampled ensemble are assumed to be independent of each other. This assumption is reasonable if the predictability of the variables of interest is smaller than the temporal resolution of DA (i.e, one year in our study), as is the case for atmospheric variables (Okazaki et al., 2021). With the online approach, the prior state of the climate system is sequentially updated. In this case, the years are not assumed to be independent of each other. Instead, the update made at the previous time step is transferred to the next one. When focusing on the atmosphere, as in our study, the online approach is not expected to outperform the offline approach (Matsikaris et al., 2015; Okazaki et al., 2021).

Here, the offline approach is based on a particle filter method described in Sect. 2.1. In addition, the framework of DA includes three main components: (1) prior climate model simulations (Sect. 2.2); (2) indirect observations of climate provided by proxy data (Sect. 2.3); and (3) proxy system models, to make the link between climate models outputs and indirect observations of climate (Sect. 2.4). The error associated with the data used in the DA procedure is also considered and defined in Sect. 2.5. Finally, in Sect. 3, we elaborate the DA experimental design and diagnostics of this study.

2.1 Particle filter

In this study, the offline DA procedure employs a particle filter as in Dubinkina et al. (2011). The particle filter (van Leeuwen, 2009) describes the prior of each climate variable by a probability density function, which is represented discretely on the basis of independent states of the model (Sect. 2.2), i.e., the particles. These model states are sampled at a chosen frequency in the model simulations, that determines the number of sampled particles. For instance, for a chosen frequency of one year, we sample a particle or model state in the model simulation every one year. It will be characterized later in Sect. 3. The prior stays the same throughout the DA procedure. As the observations are annually resolved (see Sect. 2.3), the time step of the DA is also annual. As a consequence, the particles are defined as the annual mean of the model state variables.

For each year of the reconstruction period (i.e, the period covered by observations), all the particles are compared to the available observations through the use of a PSM (Sect. 2.4). This allows estimating the likelihood of each particle knowing the observations at that time. This computation also takes into account all known sources of errors (see Sect. 2.5). The likelihood then determines the weight given to each particle. The weight given to a particle will be higher if the particle is closer to the observations, and lower if further to the observations. In this way, starting from a prior distribution where all the particles have the same weight, the particle filter produces a posterior distribution where the weights of the particles are redistributed according to the observations, for each year of the reconstruction, and independent of the location. In this framework, the ~~dynamical consistency of the model is preserved, since only the weights of the particles are updated~~ individual particles for a given timestep are dynamically consistent climate states. However, it is important to keep in mind that even though atmospheric states are almost independent on an annual timescale (as in this study), the dynamics of the ocean and cryosphere components of the climate system involves very long timescales, and their states also influence the atmospheric variables. At the end of the DA procedure, the weighted mean of the particles for each year over the assimilation period provides an estimate of the climate variables (i.e., the ensemble mean). ~~The~~ Since dynamical consistency in space and between variables involves non-linear equations, the weighted ensemble mean is therefore not fully dynamically consistent. Finally, the weighted standard deviation of the particles is used to estimate the range of the reconstruction (i.e., the ensemble spread).

More details on the implementation of the particle filter in our experiments can be found in Rezsöhazy (2021).

2.2 Climate model simulations

We use the ensemble of three simulations performed with the isotope-enabled Community Earth System Model version 1 (iCESM1; Brady et al., 2019; Stevenson et al., 2019; publicly available at https://www.earthsystemgrid.org/dataset/ucar.cgd. ccsm4.CESM_CAM5_LME.html; last access: 12 May 2021) as a basis for building the prior of our data assimilation experiments. The iCESM1 is a coupled atmosphere–ocean–sea-ice model at 1.9° latitude and 2.5° longitude resolution for the atmosphere component and around 1° for the ocean–sea-ice component. The model simulations include the anthropogenic (aerosol emissions and land use changes) and volcanic aerosol forcings, the greenhouse gas emissions, solar irradiance variations and Earth's orbital changes (Brady et al., 2019; Stevenson et al., 2019). The annually resolved iCESM1 simulations used in this study span the July 850 – June 2005 CE time period. This corresponds to the time span of the simulations used to sample

155   the particles during the data assimilation procedure, by contrast with the time span of the reconstructions (i.e., the past four centuries). For our DA experiments (Sect. 3), all the particles are sampled in the three iCESM1 members for a total of 3465 particles.

2.3 Proxy data

In this study, we make use of three types of proxy data from two types of archives: first, tree-ring width (Sect. 2.3.1) and,
160   second, δ^{18}O (i.e, the ratio of stable isotopes oxygen in the ice core) and snow accumulation from ice cores (Sect. 2.3.2). They represent the best available continental proxies in the SH continents (South America, Australia, Tasmania, New Zealand and Antarctica) for reconstructing the past climate at high-resolution (PAGES 2k Consortium, 2017).

2.3.1 Tree-ring width data

The PAGES2k database (hereafter PAGES2k2017) includes 692 proxy records (PAGES 2k Consortium, 2017) selected for
165   reconstructing the temperature of the last two millennia. The records are from a variety of archives (tree rings, ice cores, lake sediments, coral, speleothems, documentary evidence, etc). TRW records represent a bit more than the half of the database with 354 records, among which 12 are located below 30°S, in New Zealand (four records), Tasmania (four records) and South America (four records). The previous version of the PAGES2k TRW database (PAGES 2k Consortium, 2013; hereafter PAGES2k2013) contains 28 TRW records below 30°S. Nine records have been excluded from the recent version of the
170   PAGES2k database because they were too short (less than 300 years). Those records are also included in this study and are all located in South America. The location of the 21 TRW records from the two databases can be seen on Fig. S1 (blue stars). The location of the 15 TRW records that are actually used in this study can be observed on Fig. 1 (see Sect. 2.4.1 and 2.4.2 for more details on the included sites). The TRW records are normalized (i.e, with a null mean and a unity standard deviation) relative to the 1900–2000 CE time period for a use with the MAIDEN model (Sect. 2.4.1) while for the regression-based model (Sect.
175   2.4.2), the anomalies are computed relative to the 1900–2000 CE time period~~are subtracted from the TRW time series~~.

2.3.2 Ice core data

In addition to tree-ring proxy records which are the focus of this paper through the use of a process-based PSM (Sect. 2.4.1), we also assimilate other proxy data, i.e. ice core records, to cover a vast area of the Southern Hemisphere (i.e., Antarctica) where no tree-ring proxy data can be found, in order to provide a comprehensive reconstruction of the Southern Hemisphere
180   climate.

     The ice core data selected here are the same as in Dalaiden et al. (2021), a study focusing on the mechanisms behind the West Antarctic climate changes. This includes the 48 annually resolved ice core snow accumulation records in Antarctica from the database of Thomas et al. (2017) spanning the last millennium (Medley and Thomas, 2019). The B40 ice core snow accumulation record of Medley et al. (2018) has also been added as this record has been published after the publication of
185   the Thomas et al. (2017) ~~'s~~database. Finally, as in Dalaiden et al. (2021), we use the 29 annually resolved $\delta^{18}O$ records in
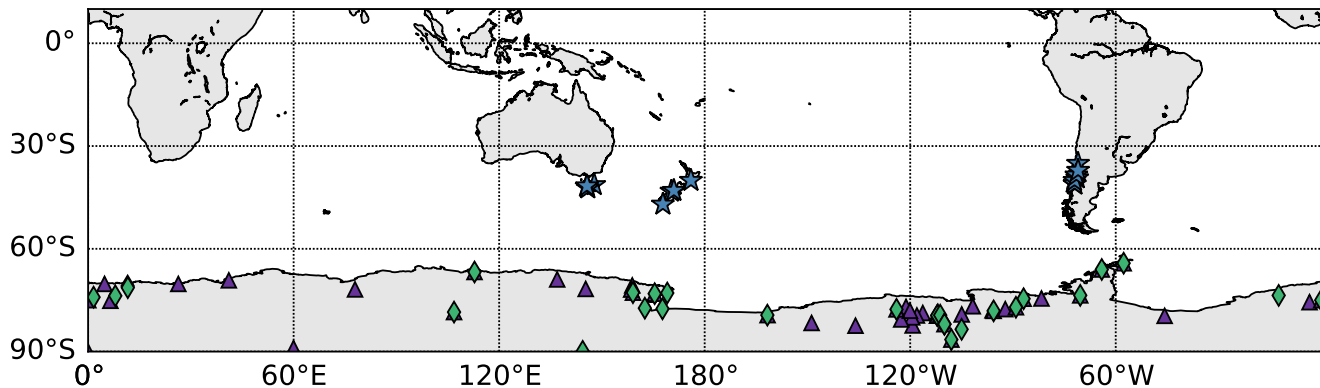
Figure 1. Location of the 15 tree-ring width (TRW) records (blue stars) from the PAGES2k databases (PAGES 2k Consortium, 2013, 2017) used in this study; the 49 ice core snow accumulation records (purple triangles) from Thomas et al. (2017) and Medley et al. (2018); and the 29 $\delta^{18}O$ records (green diamonds) from Stenni et al. (2017) (Sect. 2.3). Background map from Hunter (2007).

.

Antarctica from Stenni et al. (2017). Fig. 1 represents the location of the snow accumulation and $\delta^{18}O$ records (purple triangles and green diamonds, respectively). They are mostly situated in West Antarctica. The ice core records are averaged over a 500 km grid in order to reduce the non-climatic noise in the observations. This represents a maximum of 20 snow accumulation and 13 $\delta^{18}O$ time series over the last four centuries. The anomalies of the ice core data are computed relative to 1900–2000 CE.

2.4 Proxy System Models

Proxy System Models, also referred to as observation operators, are a crucial component of the DA framework that makes the link between the climate state variables of the prior such as near-surface air temperature and precipitation and the assimilated indirect observations from proxies, here tree-ring width. In this study, we make use of two types of tree-ring PSMs, at opposite ends of the PSM spectrum: first, a process-based ecophysiological model, MAIDEN (Sect. 2.4.1); second, a linear regression model (Sect. 2.4.2). If we neglect here potentially important elements such as post-depositional effects, the ice core records can be directly compared with the climate model. In particular, the difference between precipitation and sublimation/evaporation ($P – E$) from iCESM1 is compared to the snow accumulation observations. It has already been shown in different studies that the main contributor to the snow accumulation variability is indeed $P – E$, especially when working at large spatial scale (Agosta et al., 2019; Souverijns et al., 2018; van Wessem et al., 2018). Finally, the ratio of stable isotope oxygen $\delta^{18}O$ is simulated in iCESM1 and can be directly compared to $\delta^{18}O$ records.

2.4.1 MAIDEN

MAIDEN (Modelling and Analysis In DENdroecology; Misson, 2004; Gea-Izquierdo et al., 2015; Gennaretti et al., 2017) is a process-based dendroclimatic model which simulates the course of photosynthesis and the allocation of carbon to different pools (stem, roots, leaves, storage) on a daily time step as a response to climatic drivers (daily maximum and minimum temperature, precipitation and atmospheric CO_2 concentration). The annual quantity of carbon which is allocated to the stem (hereafter Dstem) is assumed to be proportional to tree-ring growth and comparable to TRW data, after normalization. Therefore, this variable is used for the comparison with the observed TRW. ~~We~~ In practice, as previously discussed in Rezsöhazy et al. (2020, 2021), the comparison of the tree-growth observations with the forest carbon accumulation simulated by MAIDEN implies some limitations. For instance, Gennaretti et al. (2018) computed a wood biomass index directly comparable to what MAIDEN simulates. This involves using both tree-ring width and density (earlywood and latewood fractions) data, while the latter is less broadly available. Finally, we use a combined version of the model from Gea-Izquierdo et al. (2015) and Gennaretti et al. (2017), developed by Fabio Gennaretti (unpublished).

We must first determine where the MAIDEN model can be skillfully applied among the 21 available PAGES2k sites (Sect. 2.3.1). MAIDEN was calibrated at the 12 SH sites from PAGES2k2017 in Rezsöhazy et al. (2021) driven by the daily maximum and minimum temperature and precipitation from the *Global Meteorological Forcing Dataset for land surface modelling* (v2) (http://hydrology.princeton.edu/data.php, last access: 5 May 2021; Sheffield et al., 2006) at 0.5° resolution (hereafter GMF), over 1950–2000 CE. The model was validated using the same dataset over the 1901–1949 CE time period. The calibration method is based on a Bayesian procedure with Markov Chain Monte Carlo sampling using the DREAMzs algorithm (Hartig et al., 2019). This Bayesian approach is used to calibrate the most sensitive parameters of the model from the photosynthesis and carbon allocation modules (18 parameters). MAIDEN was found to be robustly calibrated with significant (95% confidence level) calibration and verification correlations greater than 0.3 at five sites south of 30°S. The same methodology was used here for calibrating and validating MAIDEN at the nine records of PAGES2k2013. One site was found to be well calibrated and validated. The six sites for which MAIDEN calibration has successfully passed those tests will thus be incorporated into the DA procedure with MAIDEN. They can be seen on Fig. S2 (orange circles). In particular, there is no site in Tasmania.

The iCESM1 data used in the DA framework are available at a monthly resolution. Yet, MAIDEN runs on a daily basis. To convert monthly data into daily data, we apply the same methodology to all the particles. We use a reference year of the daily GMF dataset (1950) which we modified to match the mean seasonal cycle over 1950–2000 CE, i.e., a climatological seasonal cycle. Then, for a given particle, the monthly anomalies of the iCESM1 data relative to 1950–2000 CE are added (near-surface air temperature) or multiplied (precipitation) to the reference year of the daily GMF data. Note that the iCESM1 does not include maximum and minimum temperature data. We use the anomalies of iCESM1 near-surface air temperature on both maximum and minimum daily GMF temperature. As for atmospheric CO_2 data, the annual data from Rubino et al. (2019) (154–1996 CE; publicly available at https://doi.org/10.25919/5bfe29ff807fb, last access: 10 January 2021) followed from 1996 by the annual data from Sato and Schmidt (1880–2011 CE; https://data.giss.nasa.gov/modelforce/ghgases/, last access: 10 January 2021) were linearly interpolated at a daily time step.

The temporal autocorrelation in the simulated Dstem can be important, through the storage pool (Gea-Izquierdo et al., 2015; Gennaretti et al., 2017; Guiot et al., 2014), so that the performance of the model in reproducing the observed tree growth of one year highly depends on the performance of the previous years. As a consequence, we cannot run the MAIDEN model independently for each year driven by the iCESM1 data as is the case with the linear regression (Sect. 2.4.2). Instead, MAIDEN is sequentially run year by year using the reconstructed climate for the previous years and thus accounting for the effect of past climate history on tree-ring growth. For any particle, we use the CO_2 data corresponding to the assimilated year.

Dstem is not directly comparable with TRW data. The observations and simulation results are thus normalized (Sect. 2.3.1) to produce unitless indexes. The mean and standard deviation of a reference run of MAIDEN with iCESM1 inputs over 1900–2000 CE are used to normalize the annual Dstem output throughout the DA procedure.

2.4.2 Linear regression

A regression-based model was developed for each of the 21 TRW records (Sect. 2.3.1), similarly to Dalaiden et al. (2021) and to the Last Millennium Reanalysis (Tardif et al., 2019). Different climate variables were considered to model the relationship between climate and tree-ring growth. They are either related to temperature, precipitation, or both (univariate or bivariate model). For temperature, we considered the current year July to next June temperature, the December-January-February (i.e., austral summer) temperature and the September to March (i.e., the growth period temperature). For precipitation, we considered the July to June precipitation, the July-August precipitation and the July-August-September precipitation. We built the different climate variables by averaging the monthly near-surface air temperature and precipitation from the GMF interpolated on the iCESM1 grid (Sect. 2.2) over the period of months of interest. The climate data were extracted from the continental grid cell closest to each individual TRW record.

Similarly to the MAIDEN calibration, the regression parameters of each statistical PSM were first calibrated over the 1950–2000 CE time period using the ordinary least squares method. For each record, we tested all the possible combinations of climatic variables. As in Dalaiden et al. (2021), the best model was selected for each TRW record using the *Bayesian Information Criterion* (BIC) (Schwarz, 1978), which has the advantage to penalize the complexity (i.e, the number of explanatory variables) of a model.

The regression models were then run independently over the 1901–1949 CE verification time period. Among the 21 sites, only two ~~sites were validated following the same~~ locations satisfied the same selection criterion as MAIDEN (Sect. 2.4.1), i.e, a significant calibration and verification correlations greater than 0.3 (Table S1). As a consequence, the validity of the regression models out of the calibration interval is not ensured for most of them. Alternatively, we decided to use a less strict selection criterion similar to what is generally done in paleoclimate DA-based reconstruction using TRW proxies with a statistical model (e.g., the Last Millennium Reanalysis; Tardif et al., 2019). The regression parameters are thus calibrated using the same method over the entire period covered by climate data, i.e., 1901–2000 CE, in order to have the largest possible sample and thus the most robust estimate of the relationship between climate and tree growth. We then excluded nine sites (one in New Zealand and eight in South America) for which the explanatory climate variables were not significantly and positively correlated with observed TRW series. Following Cook et al. (2013), and as stated in PAGES 2k Consortium (2017)

270 , trees growing better under increasing temperature should be more reliable recorders of past temperature variability, because they respond to temperature-limiting conditions. The same assumption can hold for precipitation. The 12 sites included in the DA procedure with the regression model (hereafter, selected sites) can be seen on Fig. S3 (blue circles). A similar criterion based on the calibration only could not be used to select the TRW records for MAIDEN. The complexity of the model and the number of calibrated parameters make it more prone to overfitting problems and biological inconsistency during the calibration

275 (Rezsöhazy et al., 2020, 2021). As a consequence, the selection of the TRW records has to be more restrictive but may also lead to a better stability of the model. For comparison with the observed TRW series, anomalies of the simulated tree-ring indexes are computed relative to 1900–2000 CE.

2.5 Observation error

In the DA framework, the observation error determines the confidence that we can have in the climate information retrieved

280 from the available proxy records. It directly influences the magnitude of the weight given to a particle when compared to an assimilated proxy observation, which has in turn direct consequences on the resulting reanalysis. We generally identify three types of observation error (e.g. Badgeley et al., 2020).

The first type of error is related to the inaccuracy of the PSM in reproducing the proxy dependency on climate as the PSMs cannot fully account for the complexity of the processes by which tree growth responds to climate and non-climatic drivers.

285 The error of MAIDEN is computed using the standard deviation of the residuals over the verification period (1901–1949 CE; Sect. 2.4.1). The error of the TRW regression-based PSM (Sect. 2.4.2) is computed as in Tardif et al. (2019), using the standard deviation of the residuals over the calibration period (1901-2000 CE). However, computing the error over the calibration period can have serious shortcomings. It does not account for the actual skill of the model to accurately estimate tree growth outside the calibration interval and is biased by possible overfitting problems associated with any statistical model. This could result

290 in an underestimation of the error associated with the statistical PSM and, as a consequence, to overconfidence in the PSM performance and finally in the climate reconstruction. The snow accumulation and δ^{18}O records can be directly compared to the model outputs (Sect. 2.4) so that a PSM is not necessary and therefore an error related to the performance of the PSM is not needed for these records.

Second, the representativeness error (e.g. Oke and Sakov, 2008) corresponds to the uncertainty related to the unresolved

295 processes at the resolution of the climate model relative to the smaller scale potentially represented by proxy observations (tree-ring width, snow accumulation and δ^{18}O in our study). For the regression-based TRW PSM, we consider that the representativeness error is implicitly taken into account by computing the error related to the PSM using observed data at the same resolution as the iCESM1 data (Sect. 2.4.2). For MAIDEN, the bias correction applied to the iCESM1 data compared with the high-resolution GMF data (Sect. 2.4.1) used for the calibration of the model is associated with an error but we assume that it is

300 weaker compared to the error related to the PSM.

Regarding the ice core records (snow accumulation and δ^{18}O), the representativeness error was computed as in Dalaiden et al. (2021), on the basis of an Antarctic simulation performed with the latest version of the polar-oriented Regional Atmospheric Climate MOdel (RACMO2.3p2, hereafter RACMO) at 27 km horizontal resolution (van Wessem et al., 2018). They

computed the standard deviation of the differences between the annual RACMO time series (1979–2016 CE) of snow accu-
mulation averaged on all the ice core sites and the time series averaged over the 500 km grid, following the same approach as
in Valler et al. (2020). As for the δ^{18}O, the same methodology was applied but with the iCESM1 linearly interpolated on the
RACMO grid over 1950–2005 CE, as RACMO does not simulate the water isotopes. The error was then corrected to account
for the underestimation of spatial variability of snow accumulation and δ^{18}O in models compared to point measurements as
ice cores, by multiplying the resulting standard deviation by a factor three (see Dalaiden et al. (2021) for more detail). The
last type of error is the instrumental error related to the measurement of the proxy data, often negligible compared to the other
sources of error, particularly in paleoclimatology (Steiger et al., 2018; Tardif et al., 2019).

3 Experiment design and diagnostics

Different DA experiments are run in this study (Table 1), varying between them by (1) the records assimilated, (2) the tree-ring
PSM included and its associated period of calibration, (3) the PSM error and (4) the period of the reconstruction.

A first set of experiments is conducted over the 1900–2000 CE time period and compared with gridded climate datasets in
order to evaluate the performance of our DA-based reconstructions regarding the use of statistical (TIC-reg) and process-based
(TIC-MDN) dendroclimatic models separately, assimilating both ice core and tree-ring proxies.

A second set of experiments is performed over the 1900-2000 CE time period to evaluate the contribution of tree-ring
proxies in the data assimilation framework over the last century. Specifically, we perform an experiment assimilating tree-ring
proxies only with the regression (TREES-reg) and MAIDEN (TREES-MDN) models and an experiment assimilating ice core
proxies only (IC). These experiments are compared with the TIC-reg and TIC-MDN experiments incorporating both types of
proxy data. The impact of the observation error associated with the tree-growth PSMs (Sect. 2.5) on the performance of the
reconstruction is also evaluated by reducing the error by a factor two in the experiments constrained by tree-ring proxies only
(TREES-reg-05 and TREES-MDN-05).

Lastly, we perform a DA-based reconstruction over the period covered by a sufficient number of TRW proxy records, i.e.,
1600–2000 CE, with the regression model (TIC-reg) and the MAIDEN model (TIC-MDN), with both tree-ring and ice core
proxies assimilated.

Our analysis concentrates on four regions or sub-regions of the mid-latitudes of the SH: South America, defined between
25° and 56°S and between 55° and 85°W (SA); the South-Western Andes, where the TRW proxy sites are located, defined
between 35° and 42.5°S and between 70° and 85°W (SAmSW); New Zealand, defined between 29° and 51°S and between
160° and 179°E (NZ); and Tasmania, defined between 39.5° and 47°S and between 140° and 154°E (Tas). We did not include
Antarctica in our study as tree-ring records have not shown to significantly affect the reconstruction in this region (see Dalaiden
et al. (2021) for more details on the reconstructions in Antarctica).

We evaluate our reconstructions against the most recent atmospheric reanalysis from the ECMWF, ERA5 (Hersbach et al.,
2020) over the 1979–2000 CE time period, and against the GMF dataset (see Sect. 2.4.1 for more details on the dataset) over
1901–2000 CE, after linearly interpolating them on the iCESM1 grid (Sect. 2.2). We focus on the near-surface air temperature,

Table 1. Data assimilation experiments in this study and experimental design: experiment name; types of proxies assimilated; Proxy System Model (PSM) used to assimilate the proxy data; error associated with the PSM; calibration period of the PSM; the proxy records assimilated. *N.A* stands for Not Applicable. See Sect. 2.4.1 and 2.4.2 for details on the *Validated* and *Selected* sites, respectively. The number of tree-ring width records assimilated is in brackets.

Experiment name	Proxies	PSM	PSM error	Calibration period of PSM	Proxy records
TIC-reg	Ice cores	N.A	N.A	N.A	All
	Tree-ring width	Regression-based	Calibration	1901–2000 CE	Selected (12)
TIC-MDN	Ice cores	N.A	N.A	N.A	All
	Tree-ring width	MAIDEN	Validation	1950–2000 CE	Validated (6)
TREES-reg	Tree-ring width	Regression-based	Calibration	1901–2000 CE	Selected (12)
TREES-MDN	Tree-ring width	MAIDEN	Validation	1950–2000 CE	Validated (6)
IC	Ice cores	N.A	N.A	N.A	All
TREES-reg-05	Tree-ring width	Regression-based	Calibration error divided by two	1901–2000 CE	Selected (12)
TREES-MDN-05	Tree-ring width	MAIDEN	Validation error divided by two	1950–2000 CE	Validated (6)

the cumulative precipitation and the 500-hPa geopotential height (ERA5 only). In addition to these variables, we also evaluate the Southern Annular Mode (SAM) over 1957–2000 CE. The SAM is the principal mode of atmospheric variability in the extratropics of the SH (Marshall, 2003). It is associated with the north-south movement of the Westerlies and often defined

340 as the difference between the normalized monthly zonal mean sea-level pressure at 40° and 65°S (Gong and Wang, 1999; Marshall, 2003), as chosen here. The SAM index from Marshall (2003) is used as a reference for our evaluation.

For each of the experiments, the performance of our reconstruction is evaluated with the computation of two statistical indicators: the Pearson correlation coefficient (r; considered as significant at the 95% confidence level) and its associated p-value; and the coefficient of efficiency (hereafter CE). The coefficient of efficiency is computed as in Nash and Sutcliffe

345 (1970):

$$CE = 1 - \sum_{i=1}^{n} \frac{(x_i - y_i)^2}{(x_i - \bar{x})^2} \tag{1}$$

with n, the number of samples; x, the observation vector; y, the reconstruction vector. The overbar characterizes a time-averaged vector of values. The average is made over 1901-2000 CE in all experiments, except for the SAM index when compared to the observations from Marshall (2003), as they only cover the 1957-2000 CE period. While the Pearson correlation coefficient

350 gives an estimation of the linear fit between the reconstructed and observed values, the CE ~~also informs on the amplitude of the observed and reconstructed signal~~compares the variance of the residuals with the variance of the predictand. A CE equal to 1 is associated with perfect predictions, a CE equal to 0, with a residual variance identical to predicting the observed mean,

and a negative CE, with larger residual variance than when predicting the mean. For regional near-surface air temperature and precipitation, the metrics are computed between the mean of the reconstruction and the mean of the ERA5 or GMF data over the continental mass that lies in the region delimited by the latitudes and longitudes above.

It is important to note at this stage that the number of proxy records, but also the tree-ring proxy records themselves, that are assimilated with the statistical or MAIDEN PSMs are different. The regression-based PSM experiment includes twice the number of tree-ring proxy records compared to the one using MAIDEN. This results from different performance of the PSMs when applied to the same subset of SH tree-ring sites (Sect. 2.4.1 and 2.4.2). When working with actual proxy data, such a situation is difficult to avoid but we must be aware that this may have consequences on the skill of the resulting DA-based reconstructions.

4 Results and discussion

4.1 Comparison of the reconstructions using regression- and process-based dendroclimatic proxy system models over the last century

The spatial distribution of the correlation coefficients between different ERA5 and reconstructed climate fields over the recent period (1979–2000 CE) indicates a general good agreement of the reconstruction using MAIDEN as a PSM south of 30°S (TIC-MDN; Fig. 2a). When compared with the DA experiment including the regression model to assimilate 12 TRW records (TIC-reg; Fig. 2b), the hemispheric 500 hPa geopotential height, near-surface air temperature and precipitation fields are reconstructed with a similar performance in both cases.

At the regional scale and over the recent period (1979–2000 CE), the correlation coefficients of the ERA5 and GMF near-surface air temperature with the reconstructed near-surface air temperature in the four regions (Table 2) are also generally similar for both TIC-reg and TIC-MDN. Except in Tasmania (Tas), where the TIC-MDN near-surface air temperature is highly and significantly correlated with both ERA5 and GMF, the correlation coefficients are low and not statistically significant for both experiments. This differs from the good general pattern previously noted (Fig. 2). Locally, the correlations can indeed be lower (see Figs. S4–S9 for the regional pattern of the correlations). It is also important to highlight that these correlations have been computed between the integrated fields and are not the average of the correlations represented on the map, which may result in slight differences. Regarding precipitation (Table 2), the correlation coefficients are particularly high and significant in South America (SA) and in the South-Western Andes (SAmSW) for TIC-reg and TIC-MDN.

Over the whole century (1901–2000 CE), both DA experiments have similar skills (ranging from -0.12 to 0.60 for the correlation coefficients, and from -0.12 to 0.36 for the CE, for the four regions; Fig. 3; Table 3) in reconstructing the near-surface air temperature when compared to the GMF climate dataset. Except in the South-Western Andes region (SAmSW in Table 3), the correlation coefficients and CE between reconstructed and GMF temperature are close for both experiments and particularly high in New Zealand (NZ). In Tasmania, no TRW records have been assimilated with MAIDEN, which may explain the lower CE (Table 3) for the TIC-MDN experiment, while the correlation coefficients are similar between both experiments.
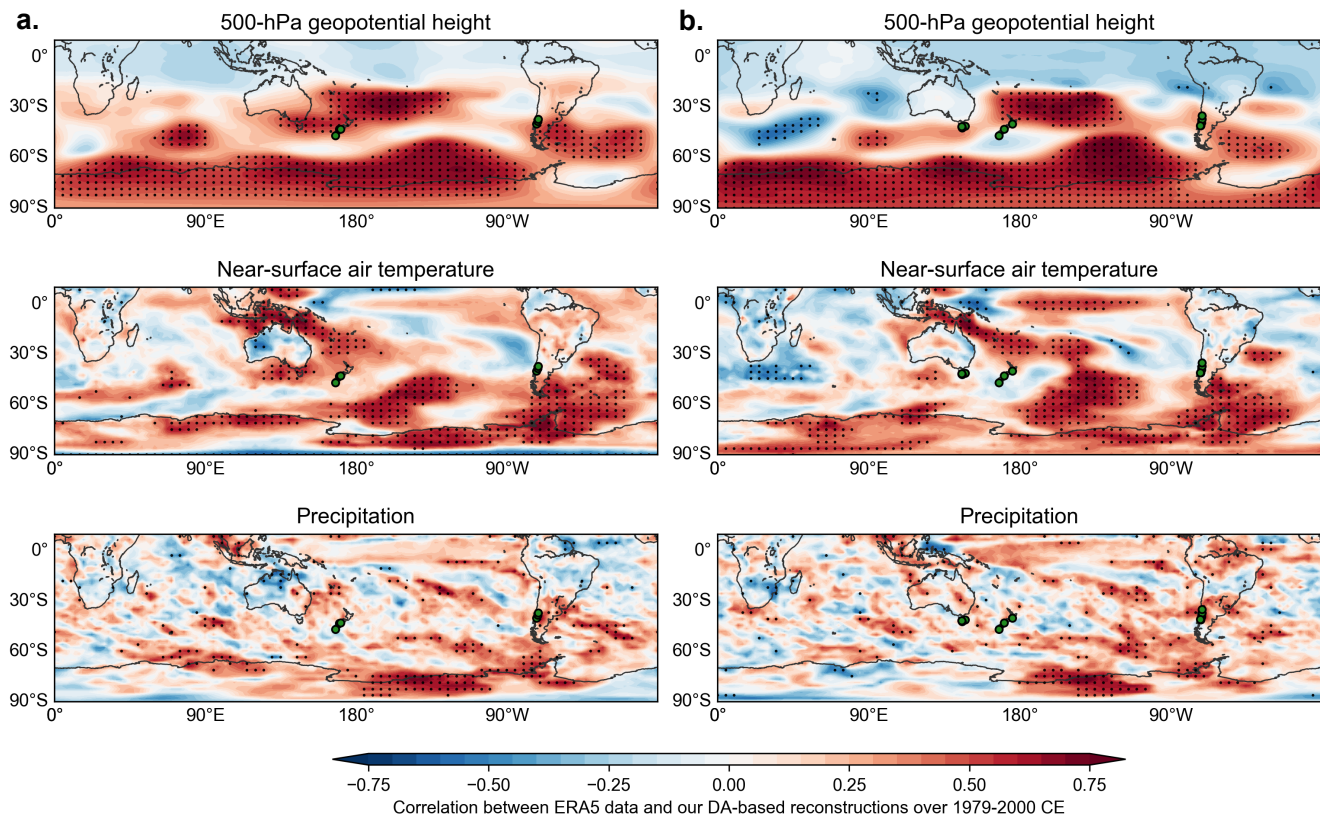
Figure 2. Pearson correlation coefficient between the reconstructed and ERA5 geopotential height at 500 hPa, near-surface air temperature or cumulative precipitation over 1979–2000 CE for the TIC-MDN (a) and TIC-reg (b) experiments (see Sect. 3 and Table 1 for details on the experiments). The green dots indicate the localization of the assimilated tree-ring width records. The black dots indicate a significant correlation coefficient at the 95% confidence level.

385 As for precipitation (Table 3 and Fig. 4), both TIC-MDN and TIC-reg are more skillful in the vicinity of the South American TRW proxies (SAmSW), with significant correlation coefficients of 0.35 and 0.37, respectively, and positive CE.

Overall, the variance of the ensemble mean reconstruction is far smaller than the observed variance, for both near-surface temperature and precipitation (Figs. 3 and 4), but observations are generally well in the range of the reconstruction (except for extreme years). Since this pattern is apparent in both experiments, i.e., when using a regression and when using a process-based 390 model as a PSM, this is likely not predominantly related to the intrinsic characteristics of each PSM. Our DA experiments only include 6 and 12 TRW proxy sites in the mid-latitudes of the SH for TIC-MDN and TIC-reg, respectively. Comparatively, in the high latitudes, the ice core records are far more numerous (Sect. 2.3.2). However, the ice core records can only have an influence on the reconstructed variance over a limited area. As a consequence, the number of assimilated TRW proxies in the mid-latitudes is probably not sufficient to effectively constrain the climate model simulations during the DA procedure. 395 In the DA framework, when no proxy is assimilated, the posterior as defined in Sect. 2.1 is equivalent to the prior, in which

Table 2. Pearson correlation coefficient between the reconstructed and GMF and ERA5 near-surface air temperature or precipitation over 1979–2000 CE for different DA experiments with both ice core and tree-ring proxies (Table 1) and for different regions (Sect. 3). *SA* stands for South America; *SAmSW* for the South-Western border of South America in the Andes where the TRW proxy sites are located; *NZ* for New Zealand; *Tas* for Tasmania. Asterisks stand for significant correlation coefficients at the 95% confidence level.

	ERA5		GMF	
	TIC-reg	TIC-MDN	TIC-reg	TIC-MDN
Near-surface air temperature				
SA	-0.07	0.04	0.12	0.20
SAmSW	-0.15	0.13	-0.21	0.14
NZ	0.11	0.33	0.14	0.32
Tas	0.05	0.56*	0.17	0.52*
Precipitation				
SA	0.72*	0.71*	0.65*	0.64*
SAmSW	0.54*	0.44*	0.53*	0.44*
NZ	0.25	0.26	0.28	0.44*
Tas	0.10	0.00	0.16	0.14

Table 3. Pearson correlation coefficient and coefficient of efficiency (CE) between the reconstructed and GMF near-surface air temperature or precipitation over 1901–2000 CE for different DA experiments with both ice core and tree-ring proxies (Table 1) and for different regions (Sect. 3). *SA* stands for South America; *SAmSW* for the South-Western border of South America in the Andes where the TRW proxy sites are located; *NZ* for New Zealand; *Tas* for Tasmania. Asterisks stand for significant correlation coefficients at the 95% confidence level.

	Correlation		CE	
	TIC-reg	TIC-MDN	TIC-reg	TIC-MDN
Near-surface air temperature				
SA	0.07	-0.12	-0.06	-0.12
SAmSW	0.34*	0.00	0.12	-0.07
NZ	0.55*	0.54*	0.30	0.27
Tas	0.60*	0.52*	0.36	0.18
Precipitation				
SA	0.12	-0.05	0.00	-0.06
SAmSW	0.37*	0.35*	0.12	0.12
NZ	-0.01	0.10	-0.12	-0.02
Tas	0.08	0.04	-0.02	-0.02

all the particles have the same weight. The posterior and the resulting ensemble mean are thus the same for each year of the reconstruction. As a consequence, the reconstruction has, by construction, a variance equal to zero. For a small number of proxies providing a weak ~~constrain~~constraint, many particles receive a similar weight, resulting in a small variance too. By contrast, the uncertainty of the reconstruction (as measured by the standard deviation of the posterior or the ensemble spread; Sect. 2.1) remains large. Additionally, both the MAIDEN and regression models work on the basis of normalized observed and simulated tree-ring growth indexes. Consequently, in the DA procedure, the error in the simulated variance of the climate signal is likely not properly assessed ~~with the TRW PSMs~~ when using such TRW PSMs to compare the climate model outputs with the proxy observations. It is in practice possible to transform tree-ring data in order to directly compare them with the output of the MAIDEN model, without normalization (Gennaretti et al., 2018). This implies to combine both tree-ring width and density measurements, that are less widely available (PAGES 2k Consortium, 2017). This could improve the model-data comparison in the DA procedure, and, consequently, the resulting reconstruction. This is not the case for the regression model.
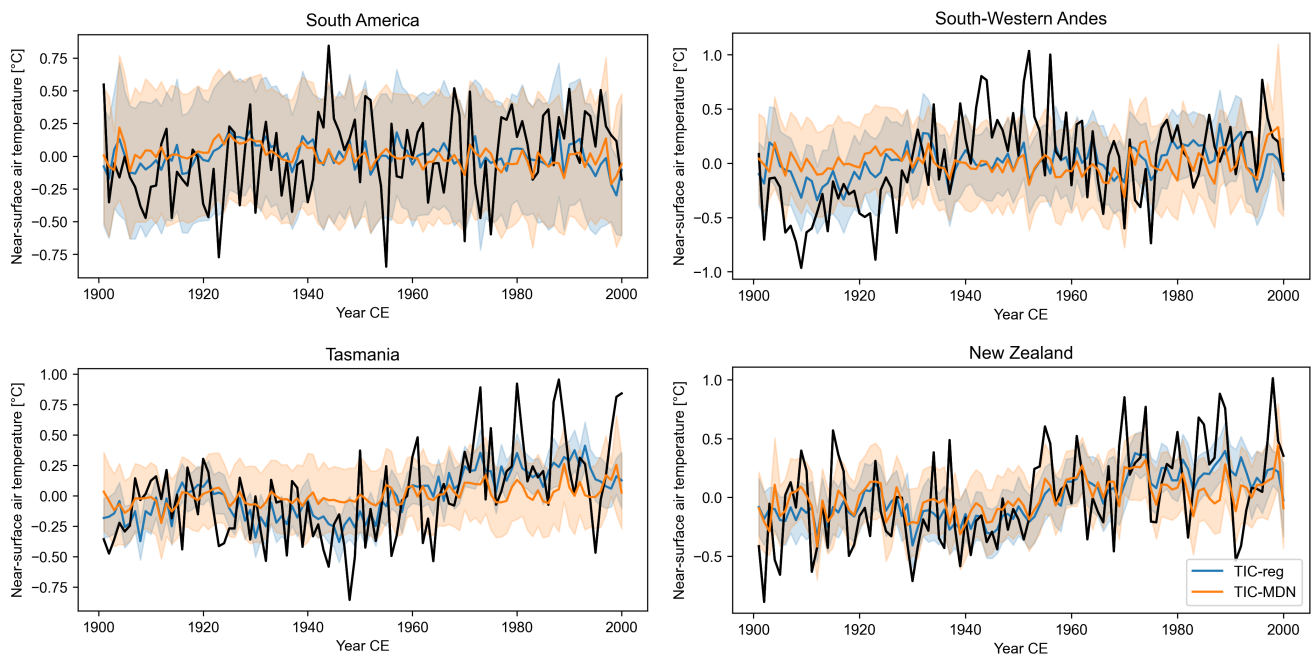
Figure 3. Reconstructed (in color) and GMF (in black) near-surface air temperature anomalies (in °C; relative to 1901–2000 CE) over 1901–2000 CE for different DA experiments with both ice core and tree-ring proxies (Table 1) in different regions (Sect. 3). The uncertainty (i.e., the ensemble spread; Sect. 2.1) associated with each reconstruction is represented by the shaded areas.

The DA procedure seeks to identify the climate conditions (i.e., the inputs of the TRW PSM; near-surface air temperature and precipitation) that gave rise to the observed TRW, through the use of a PSM. As a consequence, ~~to fully verify our DA procedure,~~ we must also establish ~~if~~ whether the reconstruction of near-surface air temperature and precipitation translates into a local ability of the MAIDEN model to reproduce tree growth at the six assimilated TRW sites over the 20th century. First, we
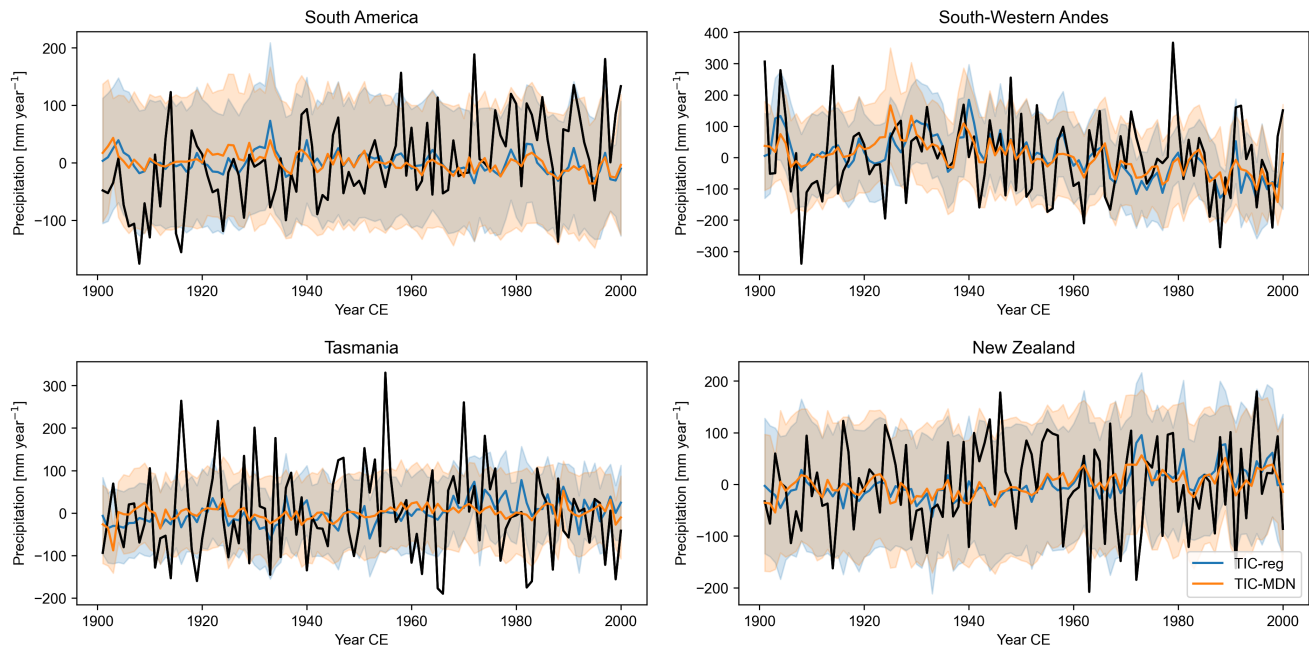
Figure 4. Reconstructed (in color) and GMF (in black) precipitation anomalies (in mm year^{-1}; relative to 1901–2000 CE) over 1901–2000 CE for different DA experiments with both ice core and tree-ring proxies (Table 1) in different regions (Sect. 3). The uncertainty (i.e., the ensemble spread; Sect. 2.1) associated with each reconstruction is represented by the shaded areas.

have retrieved the monthly climate fields from the ensemble mean of the TIC-MDN reconstruction. Then, the monthly climate inputs have been converted to daily inputs as in Sect. 2.4.1. For consistency, the tree-growth time series of the MAIDEN model have been normalized with the mean and the standard deviation of the reference run used in the DA procedure (Sect. 2.4.1). If we only look at the correlation coefficients between the observed and simulated tree-growth indexes (Fig. S10), MAIDEN
415 has been successfully run over 1900–2000 CE with the TIC-MDN reanalysis of near-surface air temperature and precipitation as inputs at most assimilated sites (with a mean correlation of 0.59). The TRW sites displaying lower correlation coefficients between the observed and simulated tree-ring indexes are the ones associated with higher PSM errors (Sect. 2.5), i.e., Aus_002 and SAm_~~024.~~ 024 (i.e., the name of the tree-ring sites, as in the PAGES2k database; Aus_* stands for a Tasmanian or New Zealand site, while SAm_* stands for a South American site). This result is logical as a lower confidence is given to the latter
420 in the particle filter (Sect. 2.1), so that the climate reconstructed at the local scale is less influenced by those records. However, the simulated variance is strongly underestimated compared to the observations, as seen on Fig. S10. This underestimation of the variance of tree growth directly comes from the underestimation of the variance in the reconstructed near-surface air temperature and precipitation, as highlighted above. The same pattern of small simulated variance is observed with the tree-ring indexes simulated with the linear regression.

425   Finally, the reconstructed SAM index is well and significantly correlated with the SAM index from Marshall (2003) for TIC-MDN (0.46) and TIC-reg (0.50) over the 1957–2000 CE period (Fig. 5). However, the TIC-reg experiment slightly better captures the significant observed trend of 0.28 standard deviation per decade over 1957-2000 CE, with a significant reconstructed trend of 0.34 and non-significant reconstructed trend of 0.20 standard deviation per decade for TIC-reg and TIC-MDN, respectively. However, both reconstructions display negative values for the CE, which indicates that they are unable to accurately

430 reproduce the observed amplitude of the SAM variations.
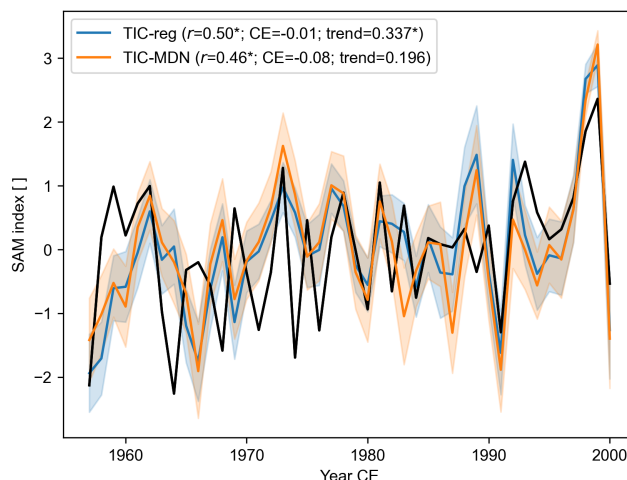
Figure 5. Z-scored (relative to 1957-2000 CE) Southern Annular mode (SAM) index of Marshall (2003) (in black) and reconstructed SAM index (in color) over 1957–2000 CE for different DA experiments with both ice core and tree-ring proxies (Table 1); Pearson correlation coefficient (r), coefficient of efficiency (CE), and decadal trend of the reconstructed SAM index (trend). Asterisks stand for a significant correlation coefficient or trend at the 95% confidence level. The observed trend of the SAM index is 0.28* standard deviation per decade over 1957–2000 CE. The uncertainty (i.e., the ensemble spread; Sect. 2.1) associated with each reconstruction is represented by the shaded areas.

     Overall, the DA experiment using MAIDEN as a PSM for assimilating the TRW records does not outperform the experiment using the regression model. However, its skill is always as good as the TIC-reg experiment with a potential better stability of MAIDEN through time, since it has been independently validated at the assimilated sites. It is also worth noting that the DA experiments with MAIDEN or the regression model have some characteristics that can influence their respective skills. Firstly,

435 the DA experiment with the regression model includes twice the number of TRW records compared to the experiment with MAIDEN (12 and six, respectively). The TRW records can also be located at different places which may result in different local skills of the DA experiments (e.g., in Tasmania). Secondly, the regression model was calibrated over the whole century (1901–2000 CE), while MAIDEN was calibrated over the second half of the century. This may confer a comparative advantage to the regression-based DA experiment, particularly when evaluated against the GMF climate dataset over 1901–2000 CE.

4.2 Contribution of tree-ring proxies to the skill of the reconstructions over the last century

The experiments discussed in Sect. 4.1 include different types of proxies. ~~Evaluating~~ For evaluating the specific contribution of tree-ring proxies using a regression- or process-based PSM, we focus on the regions and climate variables for which tree-ring data could likely provide a strong ~~constrain~~ constraint (Sect. 4.1). This evaluation is essential to ensure that the few tree-ring proxies assimilated – compared to the ice core data – are efficiently taken into account in a realistic framework including proxies in different regions using both MAIDEN or the regression-based PSM, and ~~if~~ whether differences arise between the experiments with these PSMs. Complementary information can be obtained by analyzing experiments in which only a fraction of the proxies is assimilated.

In South America (SA; Table 4 and Fig. 6), the correlations with observations are low for the experiments with tree-ring proxies alone, whatever the tree-ring PSM used (TREES-reg and TREES-MDN). Ice core proxies alone (IC) perform better than the tree-ring proxies alone (TREES-reg and TREES-MDN) or when combined with the tree-ring proxies (TIC-reg and TIC-MDN) in terms of correlation coefficients, but not in terms of reconstructed variance (standard deviation for TIC-reg, TIC-MDN, TREES-reg or TREES-MDN higher than IC). However, as previously noted, it is clear that, even though the reconstructions with tree-ring proxies display a higher variance, none of the DA experiments accurately reconstruct the observed variance of precipitation (Fig. 6).

At the more local scale, in the South-Western Andes (SAmSW), the tree-ring proxies alone provide a better reconstruction skill than for the South America (TREES-reg and TREES-MDN; Table 4). The correlation coefficients are, however, slightly lower than in the experiments using both tree-rings and ice core proxies or ice core proxies alone, in particular for TREES-MDN. Nevertheless, the TREES-reg and TREES-MDN reconstructions have a higher reconstructed variance compared to IC, in particular for TREES-reg (Fig. 6). The combination of both tree-ring and ice core proxies (TIC-reg and TIC-MDN) provides the best reconstruction in the South-Western Andes, in terms of correlation coefficients and of reconstructed variance.

Only one proxy record is used in both TREES-reg and TREES-MDN in South America (i.e, SAm_11). For this site (i.e., the grid cell closest to the proxy site), the correlation coefficients are similar for all the experiments, so that the different proxy records all bring some information, but when tree-ring records are included, an improvement of the reconstructed variance is obtained. More generally, at all spatial scales in South America the performance of TREES-reg in terms of correlation coefficients and reconstructed variance is slightly better than TREES-MDN.

In New Zealand, the correlations with observations are similar in all the experiments (NZ; Table 4). In other words, tree-ring proxies only are able to provide some skill at the scale of the New Zealand, even without assimilating ice core proxies, and with a slightly higher reconstructed variance (especially for TREES-reg). The contrast between the reconstructed variance in IC and TREES-reg or TREES-MDN experiments is stronger at the scale of the proxy site (Aus_005 and Aus_030, the tree-ring proxy sites assimilated both with MAIDEN and the regression model). At the Aus_030 proxy site the correlation coefficients are even higher for the experiments assimilating tree-ring proxies only compared to ice core proxies only. However, the reconstructed variance is still underestimated compared to observations (Fig. 6).
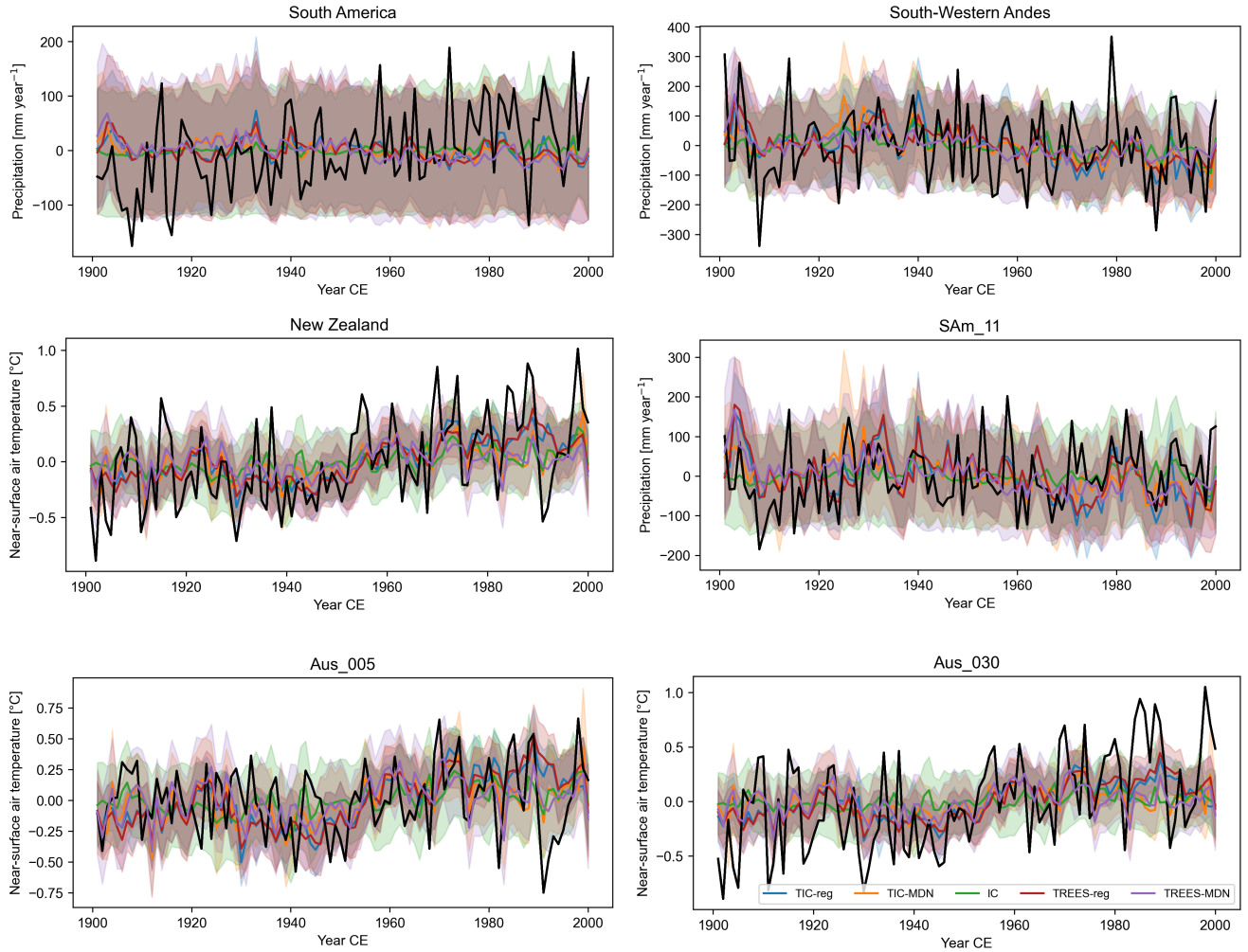
Figure 6. Reconstructed (in color) and GMF (in black) temperature (in °C; relative to 1901–2000 CE) or precipitation anomalies (in mm year^{-1}; relative to 1901–2000 CE) over 1901–2000 CE for different DA experiments (Table 1, with ice core proxies only, with tree-ring proxies only or with both) in different regions (Sect. 3) and proxy sites (Fig. 1). *Aus* refers to a site in New Zealand and *SAm*, in South America. The uncertainty (i.e., the ensemble spread; Sect. 2.1) associated with each reconstruction is represented by the shaded areas.
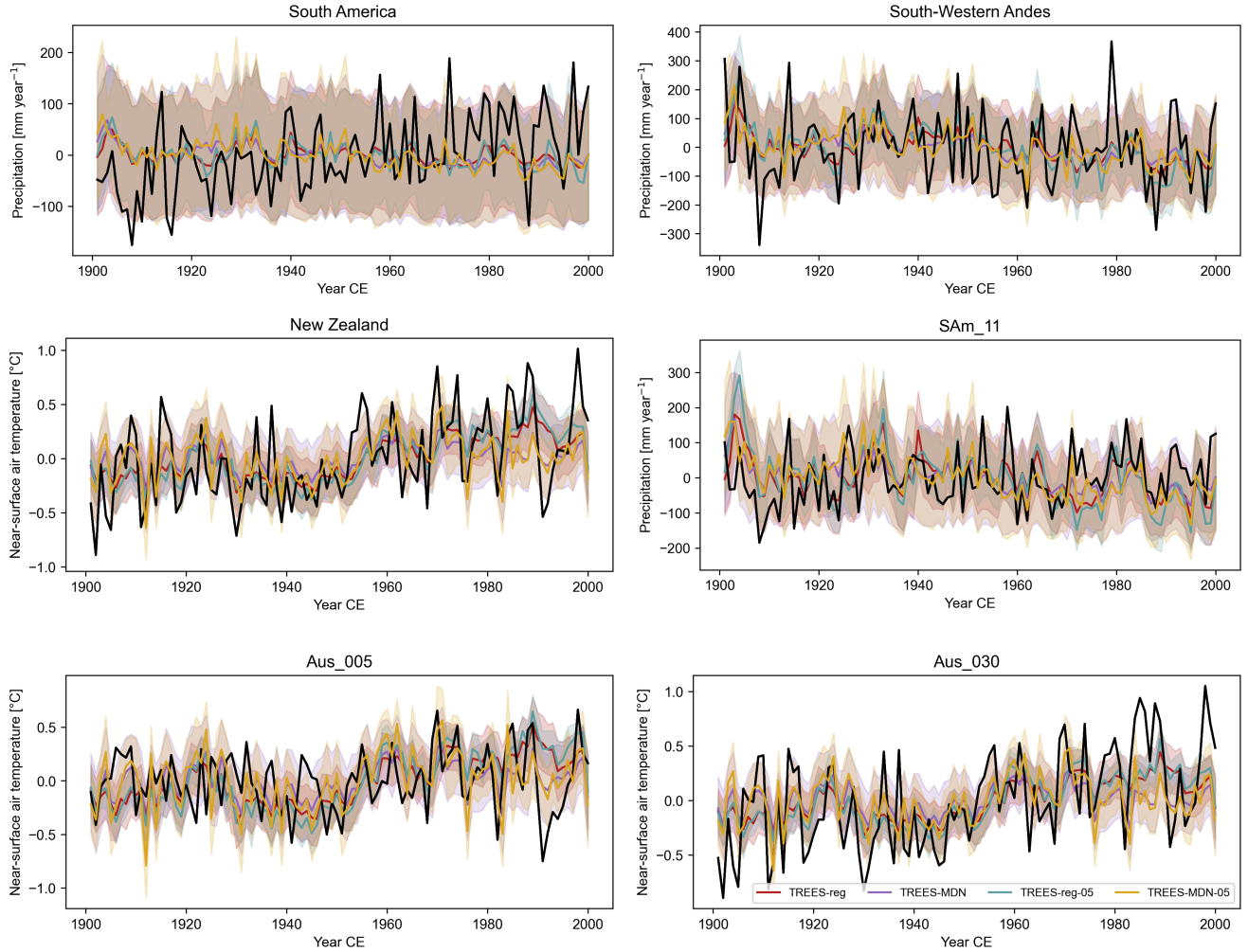
Figure 7. Reconstructed (in color) and GMF (in black) temperature (in °C; relative to 1901–2000 CE) or precipitation anomalies (in mm year⁻¹; relative to 1901–2000 CE) over 1901–2000 CE for different DA experiments (Table 1, with tree-ring proxies only, with the full observation error or the half of the observation error; see Sect. 3 for more details) in different regions (Sect. 3) and proxy sites (Fig. 1). *Aus* refers to a site in New Zealand and *SAm*, in South America. The uncertainty (i.e., the ensemble spread; Sect. 2.1) associated with each reconstruction is represented by the shaded areas.

Table 4. Pearson correlation coefficient between the reconstructed and GMF near-surface air temperature or precipitation over 1901–2000 CE for different DA experiments (Table 1, with ice core proxies only, with tree-ring proxies only or with both) in different regions (Sect. 3) and proxy sites (Fig. 1). Standard deviation of the reconstruction in °C (near-surface air temperature) or in mm year^{-1} (precipitation). *SA* stands for South America; *SAmSW* for the South-Western border of South America in the Andes where the TRW proxy sites are located; *NZ* for New Zealand. *Aus* refers to a site in New Zealand and *SAm*, in South America. Asterisks stand for significant correlation coefficients at the 95% confidence level.

	Correlation				
	TIC-reg	TIC-MDN	IC	TREES-reg	TREES-MDN
Precipitation					
SA	0.12	-0.05	0.28*	0.02	-0.20*
SAmSW	0.37*	0.35*	0.34*	0.33*	0.26*
SAm_11	0.26*	0.22*	0.22*	0.22*	0.14
Near-surface air temperature					
NZ	0.55*	0.54*	0.50*	0.53*	0.52*
Aus_005	0.20*	0.26*	0.29*	0.19	0.24*
Aus_030	0.49*	0.41*	0.30*	0.57*	0.48*

	Standard deviation				
	TIC-reg	TIC-MDN	IC	TREES-reg	TREES-MDN
Precipitation					
SA	17.82	15.29	8.31	17.22	18.53
SAmSW	63.59	51.79	33.79	52.00	38.50
SAm_11	59.51	41.84	17.63	57.74	41.38
Near-surface air temperature					
NZ	0.19	0.15	0.11	0.18	0.13
Aus_005	0.21	0.17	0.11	0.21	0.16
Aus_030	0.17	0.13	0.07	0.18	0.13

The underestimation of the variance compared to the observations could arise from a too low confidence given to the assimilated tree-ring records in the data assimilation procedure, which is directly related to the error associated with the PSMs (Sect. 2.5). In order to test this hypothesis, we have reduced this error by two. This leads to an increase of the reconstructed variance at all spatial scales in both South America (precipitation) and New Zealand (near-surface air temperature; Table 5 and Fig. 7), and in particular for TREES-MDN in New Zealand. This does not translate, however, into higher correlation coefficients, since they virtually remain unchanged (Table 5). In an ideal experiment where the relationship between climate and tree-ring growth would be perfectly reproduced, we could thus expect the observed variance to be more skillfully reconstructed by the ensemble mean reconstruction.

More generally, this sensitivity analysis shows that tree-ring proxies have a positive effect on the reconstructed mid-latitude climate changes in our framework with both the regression- or process-based dendroclimatic PSM, especially at the regional to local scale. The fact that the reconstructed variance can increase at the more local scale when assimilating tree-ring proxies

Table 5. Pearson correlation coefficient between the reconstructed and GMF near-surface air temperature or precipitation over 1901–2000 CE for different DA experiments (Table 1, with tree-ring proxies only, with the full observation error or the half of the observation error; see Sect. 3 for more details) in different regions (Sect. 3) and proxy sites (Fig. 1). Standard deviation of the reconstruction in °C (near-surface air temperature) or in mm year^{-1} (precipitation). *SA* stands for South America; *SAmSW* for the South-Western border of South America in the Andes where the TRW proxy sites are located; *NZ* for New Zealand. *Aus* refers to a site in New Zealand and *SAm*, in South America. Asterisks stand for significant correlation coefficients at the 95% confidence level.

| | **Correlation** | | | |
	TREES-reg	TREES-MDN	TREES-reg-05	TREES-MDN-05
Precipitation				
SA	0.02	-0.20	-0.05	-0.19
SAmSW	0.33*	0.26*	0.37*	0.24*
SAm_11	0.22*	0.14	0.21*	0.16
Near-surface air temperature				
NZ	0.53*	0.52*	0.51*	0.50*
Aus_005	0.19	0.24*	0.23*	0.25*
Aus_030	0.57*	0.48*	0.57*	0.48*

| | **Standard deviation** | | | |
	TREES-reg	TREES-MDN	TREES-reg-05	TREES-MDN-05
Precipitation				
SA	17.22	18.53	25.12	26.73
SAmSW	52.00	38.50	75.51	60.58
SAm_11	57.74	41.38	79.44	60.13
Near-surface air temperature				
NZ	0.18	0.13	0.23	0.21
Aus_005	0.21	0.16	0.27	0.26
Aus_030	0.18	0.13	0.22	0.22

alone indicates that the number of tree-ring proxies assimilated in our experiments likely has an impact on the reconstructed variance. If more proxies were available, we could expect that this local increase in the variance would lead to an increase at larger spatial scale. The definition of the error related to the PSMs likely plays a role as well. In particular, the computation of the error for MAIDEN is stricter than for the regression model (i.e., computed on the verification period and on the calibration period, respectively; Sect. 2.5) and the number of proxy records assimilated with MAIDEN is lower as well, which may advantage the regression-based PSM.

4.3 Data assimilation based reconstruction of the Southern Hemisphere climate over the last 400 years

In this section, we compare the DA-based reconstructions with the process- or regression-based PSM over the past four centuries, to expand to longer timescales the conclusions obtained from the comparison of the two methodologies over the 20th century. If we first look at the regional reconstructions of near-surface air temperature (Fig. 8) and precipitation (Fig. 9) for

the four regions, we observe that the ensemble spread is in general large, contrasting with an interannual variability of the
ensemble mean reconstructions that is generally small. This could indicate that we overestimate the error associated with the
PSM, and therefore underestimate the confidence of TRW data in the DA procedure. Besides, as highlighted before, this is at
least partly due to the small number of assimilated TRW proxies in the mid-latitudes that prevents the prior distribution given
by the climate model to be effectively constrained by the available observations. This pattern is stronger when the spatial scale
is large (e.g., South America), and weaker when the spatial scale is small (e.g., Tasmania) and closer to the proxy sites. This is
particularly true for MAIDEN, for which we assimilated half the number of TRW proxy records compared with the regression
model.
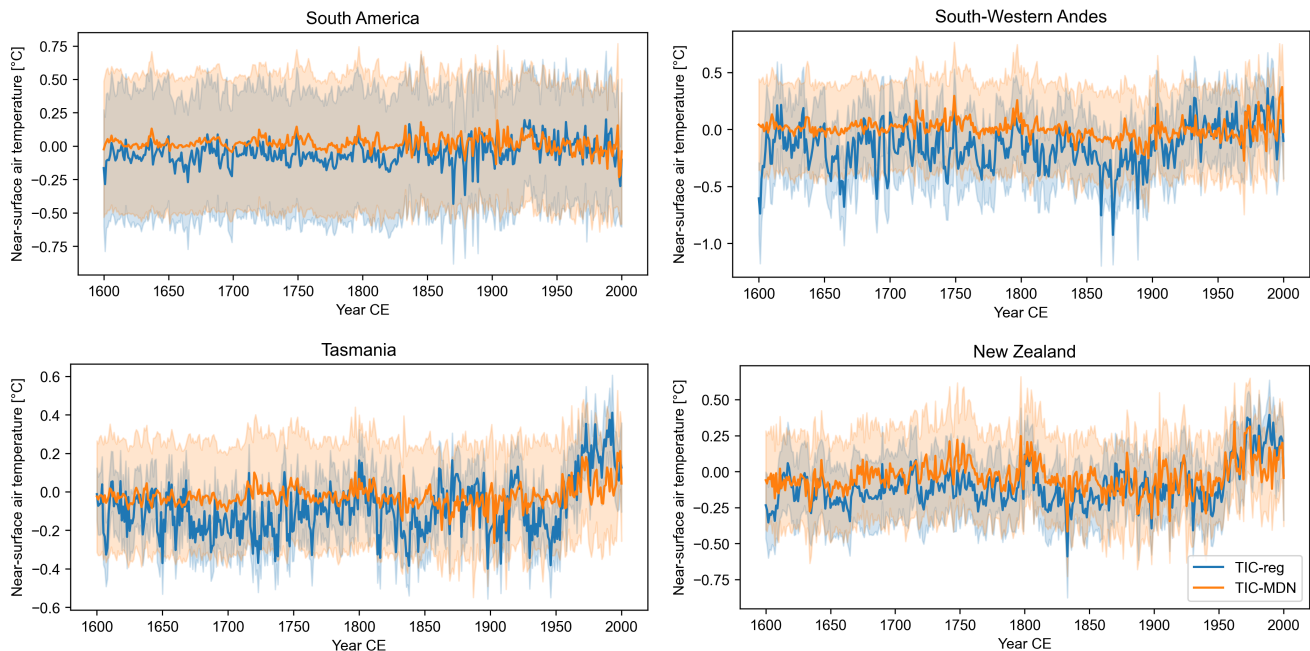
Figure 8. Reconstructed TIC-reg (in blue) and TIC-MDN (in orange) near-surface air temperature anomalies (in °C; relative to 1901–2000
CE) over 1600–2000 CE (Table 1) in different regions (Sect. 3). The uncertainty (i.e., the ensemble spread; Sect. 2.1) associated with each
reconstruction is represented by the shaded areas.

In New Zealand, both TIC-reg and TIC-MDN display a smaller spread and larger variability of the reconstructed near-surface
air temperature compared to the other regions (Fig. 8), which suggests that the DA has been more effective in updating the
climate model simulations with the available proxy records. The reconstructed near-surface air temperature in New Zealand
was also well evaluated against GMF observations over 1901–2000 CE in Sect. 4.1 (Table 3). The low-frequency variability of
the two reconstructions over the last 400 years is relatively similar overall, with a very small trend before the industrial era (until
around 1850-1900 CE) followed by an increase in temperature. TIC-reg and TIC-MDN display a significant trend of 0.047
and 0.023°C per decade during the 20th century (against 0.062°C per decade in the GMF observations) and a significant trend

of 0.001 and non-significant trend of 0°C per decade before 1900 CE, respectively. The correlation coefficient between both reconstructions is 0.63 (significant at the 95% confidence level), which indicates that the trend and timing of the reconstructions are close to each other. However, TIC-MDN displays higher near-surface air temperature anomalies before 1850 CE than TIC-reg (-0.05°C in average for TIC-MDN, against -0.14°C for TIC-reg, relative to 1901-2000 CE). In Tasmania, there are no assimilated sites for MAIDEN (Figs. S2 and S6) and four assimilated sites for the regression over a very small area (Figs. S3 and S9), which may explain the difference between the reconstructed near-surface air temperature of TIC-reg and TIC-MDN (Fig. 8). This is consistent with the results obtained over the 20th century. Finally, we observe a difference in the reconstructed low-frequency variability between TIC-reg and TIC-MDN. Compared with TIC-reg, TIC-MDN tends to reconstruct a weaker temperature difference between the preindustrial and industrial era, particularly in South America, South-Western Andes and Tasmania.
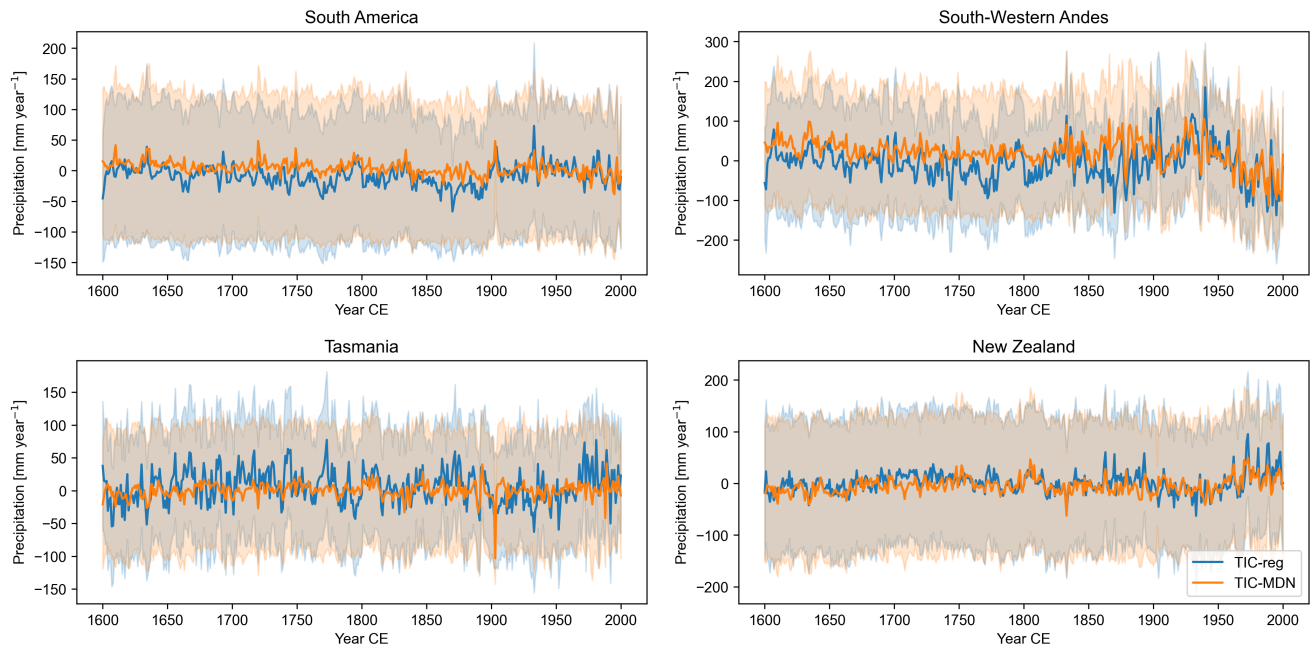
Figure 9. Reconstructed TIC-reg (in blue) and TIC-MDN (in orange) precipitation anomalies (in mm year^{-1}; relative to 1901–2000 CE) over 1600–2000 CE (Table 1) in different regions (Sect. 3). The uncertainty (i.e., the ensemble spread; Sect. 2.1) associated with each reconstruction is represented by the shaded areas.

For precipitation (Fig. 9), the South-Western Andes is the region for which the ensemble spread is the smallest and the interannual variability the largest for both TIC-reg and TIC-MDN. Both experiments are also skillful for reconstructing precipitation in this region when compared with the state-of-the-art climate datasets over the last century (Table 3; Sect. 4.1). The correlation coefficient between both reconstructions is 0.58 and is statistically significant, so that the trend and timing of the reconstructions are close to each other. Both reconstructions display a very small trend before 1900 CE (a significant trend of

-0.586 and a non-significant trend of -0.042 mm year^{-1} per decade for TIC-reg and TIC-MDN, respectively) and a negative trend after 1900 CE (a significant trend of -12.79 and -8.89 mm year^{-1} per decade for TIC-reg and TIC-MDN, respectively; against a non-significant trend of -1.26 mm year^{-1} per decade in the GMF observations). However, TIC-MDN displays higher anomalies before 1900 CE than TIC-reg (28.23 mm year^{-1} in average for TIC-MDN against -5.70 mm year^{-1} for TIC-reg, relative to 1901-2000 CE).

At the regional scale, we have highlighted from the reconstructed near-surface air temperature and precipitation that the experiments with the statistical or process-based PSM show different amplitudes of climate change over the past 400 years, even if close in terms of correlation (i.e, the trend and timing of change). It is not possible, of course, to determine which one is the most accurate. This difference in the reconstructions could be related to the number of proxies included in each of the DA experiments, and their associated error (Sect. 2.5). This also points out that the commonly used linear regression may also not be fully able to always reflect properly the full spectrum of the climate variability over the past centuries. Alternatively, MAIDEN may underestimate low-frequency trends, for instance because it overestimates the effect of CO_2 concentration on tree growth (Rezsöhazy et al., 2021) which leads to an underestimation of the near-surface temperature low-frequency variability. It may also be related to the calibration of MAIDEN that may not be able to properly capture the low-frequency signal in tree-ring series.

Finally, we observe an increase of the reconstructed variability in the precipitation of the South-Western Andes (Fig. 9), particularly for TIC-MDN, from around 1800–1850 CE, i.e., when most of the ice core records start to become available (Fig. S11). Before 1830 CE, the standard deviation of the reconstruction is 30.7 and 21.0 mm year^{-1} and, after 1830 CE, 56.3 and 44.2 mm year^{-1} for TIC-reg and TIC-MDN, respectively.
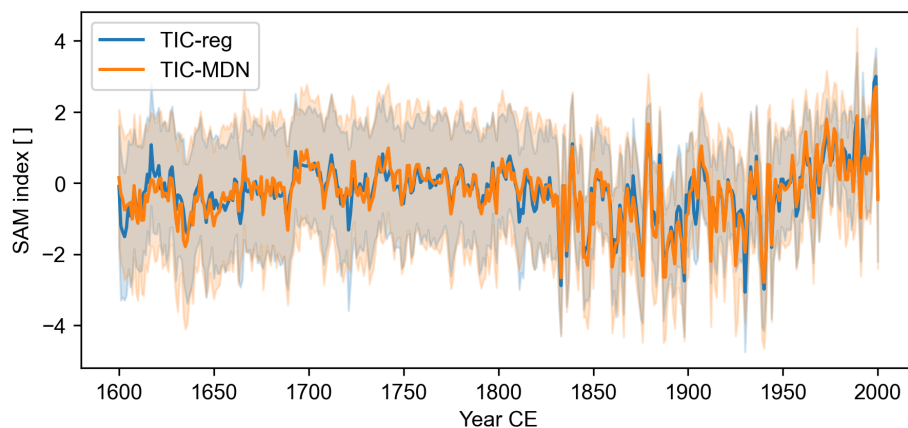
Figure 10. Z-scored (relative to 1901–2000 CE) TIC-reg (in blue) and TIC-MDN (in orange) reconstructed SAM index over 1600–2000 CE (Table 1). The uncertainty (i.e., the ensemble spread; Sect. 2.1) associated with each reconstruction is represented by the shaded areas.

This increase of interannual variability is even more striking in the reconstructed SAM index for both reconstructions (Fig. 10). Before around 1830 CE particularly, too few TRW proxy sites have been assimilated in the mid-latitudes to be able to

reconstruct the large-scale variability adequately (Fig. S11). The standard deviation of the z-scored reconstructed SAM index is 0.52 and 0.51 before 1830 CE and 1.05 and 1.11 after 1830 CE for TIC-reg and TIC-MDN, respectively. However, the SAM reconstructions are still highly and significantly correlated with each other ($r = 0.91$), and look essentially the same over the whole 1600-2000 CE period (Fig. 10). Before 1830 CE, when the number of ice core records is more limited (Fig. S11), the reconstructions are less correlated ($r = 0.77$; and $r = 0.95$ after 1830 CE).

Lastly, over the last 400 years, the MAIDEN model generally reproduces the observed tree growth well in terms of correlation coefficients at the six assimilated sites (with a mean correlation of 0.53; Fig. S12). However, as previously highlighted in Sect. 4.1, the simulated variance is largely underestimated compared to the observations, which is not surprising regarding the reconstructions of near-surface air temperature and precipitation of the past 400 years.

5   Conclusions

In this study, we have included for the first time a dendroclimatic process-based model into a data assimilation procedure to reconstruct the past climate variability of the Southern Hemisphere. All the results from our DA experiments have been compared with the commonly used regression-based PSM. First, we have shown that including the ecophysiological model MAIDEN as a proxy system model into a data assimilation procedure is technically possible. The assimilation of actual TRW records with a tree-growth mechanistic model has never been published before. Knowing that it is at least achievable is promising.

The general pattern of near-surface air temperature, precipitation and the atmospheric circulation of the mid-to-high latitudes of the SH is well reproduced by the DA experiment with MAIDEN over the recent period. Over the last century, the DA experiment with MAIDEN is also able to produce reconstructions that are well correlated with different climate fields, such as near-surface air temperature, precipitation and atmospheric circulation in some regions of the Southern Hemisphere, with only a few TRW proxy records assimilated. The DA-based reconstructions with the regression have a similar skill as with MAIDEN.

Over the instrumental era, the skill of MAIDEN as a PSM in the present DA framework is not better but it is at least as good as the regression-based PSM. Given the advantages that a process-based PSM can still bring beyond the instrumental period in terms of non-linearity of the relationship between climate and tree growth, but also in terms of climate variables included as inputs of the model (e.g. atmospheric CO_2 concentration), having a comparable skill to a regression-based PSM is encouraging. There are still opportunities for improvement with the MAIDEN model while for the regression, the perspectives are more limited. For instance, we could improve the model-data comparison by using an observed tree growth index that is closer to the variable simulated by MAIDEN, i.e., forest carbon accumulation, by combining density and ring width measurements. Different shortcomings related to the calibration of the MAIDEN model have also been addressed in Rezsöhazy et al. (2020) and Rezsöhazy et al. (2021). If solved, this may likely help to increase the number of proxy sites where MAIDEN could be validated. It could also improve the performance of MAIDEN in simulating the relationship between climate and tree growth, and so enhance the skill of the DA-based reconstruction using MAIDEN as a PSM. In particular, the calibration procedure could be more informative in the future, for instance in the definition of the prior ranges of the calibrated parameters, that could be

set by species or biome. Different data sources, such as satellite data and the associated vegetation indexes, could also be used to better inform the different biological processes in the model (e.g. related to photosynthesis or phenology). Moreover, in this study, the criteria for including a TRW proxy record with the regression had to be less strict than for MAIDEN. MAIDEN was indeed validated at all the assimilated sites while the regression was not. This may provide a better stability of the MAIDEN model.

It appears clearly that our DA framework does not ~~include enough TRW proxy records in the mid-latitudes to~~ provide reconstructions that properly reproduce the observed variance of near-surface air temperature and precipitation with both TRW PSMs, particularly at larger spatial scale (e.g., South America), likely due to the low number of TRW proxy records assimilated in the mid-latitudes. The Southern Hemisphere counts indeed few TRW proxy records (PAGES 2k Consortium, 2017) but offers an interesting framework as few reconstructions of the Southern Hemisphere climate exist to date (Neukom et al., 2011, 2014). ~~This~~ The underestimation of variance is a fundamental issue with the particle filter, but also potentially with other data assimilation methods (e.g. Tardif et al., 2019). In the ideal case, individual members of ensemble climate simulations have realistic temporal variability on all spatial scales. Any averaging of random, non-forced variability will reduce the variability. In other words, any ensemble mean will always have a too low variability, regardless of how it is constructed. However, the extent of the reduction depends on whether members with more similar or more different variability are used for calculating the weighted ensemble mean. In contrast to unweighted ensemble means, the particle filter will give high weights to ensemble members that match the observations. If the empirical information strongly constrains the particle selection, the particles will be more similar than in a less constrained case, and there will be less underestimation of variance. Besides, the similarity measure is determined at the locations of the proxies, but the weights given to each particle are independent of location. Particles that have similar states at the proxy locations do not necessarily have similar states at other locations, and the reduction of variance in unconstrained locations is therefore likely to be larger than in constrained locations.

In our reconstructions, this pattern of variance becomes even more striking before 1830 CE, where we observe a smaller interannual variability of the reconstructed near-surface air temperature and precipitation in some regions of the mid-latitudes, and of the reconstructed SAM index, compared to the last 150 years. However, the timing of the climate changes is well constrained in some localized areas of the mid-latitudes, e.g., the South-Western Andes, New Zealand and Tasmania, as shown by the significant and relatively high correlations with the observations. As a consequence, the MAIDEN model simulates tree-ring growth series that are generally well correlated with the observations over the last 400 years, but which display a small variance compared to the observations.

Despite this, at the regional scale, we have highlighted a difference between the two DA experiments in the reconstructed amplitude of the climate change in the preindustrial era relative to the 20th century. Using MAIDEN as a PSM for TRW records tends to reconstruct an average warmer and wetter climate than the regression-based PSM. This questions the almost exclusive use of linear statistical PSMs for tree-ring proxies in data assimilation to reconstruct the past climate changes. However, this would need an in-depth investigation with more TRW proxy records in order to be confirmed, as well as additional analysis to ensure the validity and robustness of MAIDEN and its calibration in our framework, and more generally of sophisticated tree-growth PSMs.

For now, our framework only offers a limited skill of the reconstructions, and we are far from providing a complete hemispheric reconstruction of atmospheric fields in the Southern Hemisphere. However, our results are encouraging for MAIDEN,
and more generally for the use of dendroclimatic process-based models in DA. In the future, mechanistic models like MAIDEN
will likely become an important tool of the DA framework that would help overcome the potential bias associated with regression models. However, this would require to continue to improve those models, including their calibration, so that they could
be applied more efficiently and more skillfully at large spatial scale. Calibrating a model like MAIDEN at large spatial scale
remains a challenge, and incorporating it into a data assimilation procedure is still more technically time consuming than the
traditional statistical approach. Finally, more investigation of the performance of the MAIDEN model as a PSM in the DA
procedure compared with the commonly used regression model would also be needed using more TRW proxy records in order
to fully document the added value of using a dendroclimatic process-based model in the data assimilation framework.

Author contributions. This study was part of JR's thesis under the supervision of HG and JG. JR, HG and JG designed the study. JR
performed the analysis and made the figures. QD contributed to the analysis, by providing help, advice and analysis scripts. FK and QD
provided the code for the offline assimilation. FK modified the original code to include MAIDEN in the data assimilation, and helped for
running it. JR led the writing of the manuscript with contributions from all authors.

Competing interests. The authors declare that there is no conflict of interest.

Acknowledgements. JR was F.R.S-FNRS research fellow, Belgium (grant no. 1.A841.18); QD is F.R.S-FNRS postdoctoral researcher, Belgium; HG is research director at F.R.S.-FNRS, Belgium; JG is emeritus research director at CNRS, France. This research has been supported
by the Belgian Research Action through Interdisciplinary Networks (BRAIN.be) from the Belgian Science Policy Office in the framework
of the "East Antarctic surface mass balance in the Anthropocene: observations and multiscale modelling (Mass2Ant)" project (contract no.
BR/165/A2/Mass2Ant). Computational resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL) and the Consortium des Équipements de Calcul Intensif en Fédération Wallonie Bruxelles (CÉCI) funded by the Fond de
la Recherche Scientifique de Belgique (F.R.S.-FNRS) under convention 2.5020.11 and by the Walloon Region.

References

Acevedo, W., Fallah, B., Reich, S., and Cubasch, U.: Assimilation of pseudo-Tree-ring-width observations into an atmospheric general circulation model, Climate of the Past, 13, 545–557, https://doi.org/10.5194/cp-13-545-2017, 2017.

Agosta, C., Amory, C., Kittel, C., Orsi, A., Favier, V., Gallée, H., Van Den Broeke, M. R., Lenaerts, J. T., Van Wessem, J. M., Van De Berg, W. J., and Fettweis, X.: Estimation of the Antarctic surface mass balance using the regional climate model MAR (1979-2015) and identification of dominant processes, Cryosphere, 13, 281–296, https://doi.org/10.5194/tc-13-281-2019, 2019.

Anchukaitis, K. J., Wilson, R., Briffa, K. R., Büntgen, U., Cook, E. R., D'Arrigo, R., Davi, N., Esper, J., Frank, D., Gunnarson, B. E., Hegerl, G., Helama, S., Klesse, S., Krusic, P. J., Linderholm, H. W., Myglan, V., Osborn, T. J., Zhang, P., Rydval, M., Schneider, L., Schurer, A., Wiles, G., and Zorita, E.: Last millennium Northern Hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions, Quaternary Science Reviews, 163, 1–22, https://doi.org/10.1016/j.quascirev.2017.02.020, 2017.

Babst, F., Bodesheim, P., Charney, N., Friend, A. D., Girardin, M. P., Klesse, S., Moore, D. J., Seftigen, K., Björklund, J., Bouriaud, O., Dawson, A., DeRose, R. J., Dietze, M. C., Eckes, A. H., Enquist, B., Frank, D. C., Mahecha, M. D., Poulter, B., Record, S., Trouet, V., Turton, R. H., Zhang, Z., and Evans, M. E.: When tree rings go global: Challenges and opportunities for retro- and prospective insight, Quaternary Science Reviews, 197, 1–20, https://doi.org/10.1016/j.quascirev.2018.07.009, 2018.

Badgeley, J. A., Steig, E. J., Hakim, G. J., and Fudge, T. J.: Greenland temperature and precipitation over the last 20 000 years using data assimilation, Climate of the Past, 16, 1325–1346, https://doi.org/10.5194/cp-16-1325-2020, 2020.

Boucher, E., Guiot, J., Hatté, C., Daux, V., Danis, P. A., and Dussouillez, P.: An inverse modeling approach for tree-ring-based climate reconstructions under changing atmospheric CO2 concentrations, Biogeosciences, 11, 3245–3258, https://doi.org/10.5194/bg-11-3245-2014, 2014.

Brady, E., Stevenson, S., Bailey, D., Liu, Z., Noone, D., Nusbaumer, J., Otto-Bliesner, B. L., Tabor, C., Tomas, R., Wong, T., Zhang, J., and Zhu, J.: The Connected Isotopic Water Cycle in the Community Earth System Model Version 1, Journal of Advances in Modeling Earth Systems, 11, 2547–2566, https://doi.org/10.1029/2019MS001663, 2019.

Briffa, K. R., Schweingruber, F. H., Jones, P. D., Osborn, T. J., Harris, I. C., Shiyatov, S. G., Vaganov, E. A., and Grudd, H.: Trees tell of past climates: but are they speaking less clearly today?, Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353, 65–73, https://doi.org/10.1098/rstb.1998.0191, https://doi.org/10.1098/rstb.1998.0191, 1998.

Büntgen, U., Tegel, W., Nicolussi, K., McCormick, M., Frank, D., Trouet, V., Kaplan, J., Herzig, F., Heussner, K.-U., Wanner, H., Luterbacher, J., and Esper, J.: 2500 Years of European Climate Variability and Human Susceptibility, Science, 331, 578–582, 2011.

Cook, E. R., Meko, D. M., Stahle, D. W., and Cleaveland, M. K.: Drought reconstructions for the continental United States, Journal of Climate, 12, 1145–1163, https://doi.org/10.1175/1520-0442(1999)012<1145:drftcu>2.0.co;2, 1999.

Cook, E. R., Anchukaitis, K. J., Buckley, B. M., D'Arrigo, R. D., Jacoby, G. C., and Wright, W. E.: Asian monsoon failure and megadrought during the last millennium, Science, 328, 486–489, https://doi.org/10.1126/science.1185188, 2010.

Cook, E. R., Krusic, P. J., Anchukaitis, K. J., Buckley, B. M., Nakatsuka, T., and Sano, M.: Tree-ring reconstructed summer temperature anomalies for temperate East Asia since 800 C.E, Climate Dynamics, 41, 2957–2972, https://doi.org/10.1007/s00382-012-1611-x, 2013.

Crespin, E., Goosse, H., Fichefet, T., and Mann, M. E.: The 15th century Arctic warming in coupled model simulations with data assimilation, Climate of the Past, 5, 389–401, https://doi.org/10.5194/cp-5-389-2009, 2009.

670 Dalaiden, Q., Goosse, H., Rezsöhazy, J., and Thomas, E.: Reconstructing atmospheric circulation and sea-ice extent in the West Antarctic over the past 200 years using data assimilation, Climate Dynamics, https://doi.org/10.1007/s00382-021-05879-6, https://doi.org/10.1007/s00382-021-05879-6, 2021.

D'Arrigo, R., Wilson, R., Liepert, B., and Cherubini, P.: On the 'Divergence Problem' in Northern Forests: A review of the tree-ring evidence and possible causes, Global and Planetary Change, 60, 289–305, https://doi.org/10.1016/j.gloplacha.2007.03.004, 2008.

675 Dee, S. G., Steiger, N. J., Emile-Geay, J., and Hakim, G. J.: On the utility of proxy system models for estimating climate states over common era, Journal of Advances in Modeling Earth Systems, 8, 1164–1179, https://doi.org/https://doi.org/10.1002/2016MS000677, 2016.

Drew, D. M., Downes, G. M., and Battaglia, M.: CAMBIUM, a process-based model of daily xylem development in Eucalyptus, Journal of Theoretical Biology, 264, 395–406, https://doi.org/10.1016/j.jtbi.2010.02.013, http://dx.doi.org/10.1016/j.jtbi.2010.02.013, 2010.

Dubinkina, S., Goosse, H., Sallaz-Damaz, Y., Crespin, E., and Crucifix, M.: Testing a Particle Filter To Reconstruct Climate Changes
680 Over the Past Centuries, International Journal of Bifurcation and Chaos, 21, 3611–3618, https://doi.org/10.1142/S0218127411030763, http://www.worldscientific.com/doi/abs/10.1142/S0218127411030763, 2011.

Dufrêne, E., Davi, H., François, C., Le Maire, G., Le Dantec, V., and Granier, A.: Modelling carbon and water cycles in a beech forest. Part I: Model description and uncertainty analysis on modelled NEE, Ecological Modelling, 185, 407–436, https://doi.org/10.1016/j.ecolmodel.2005.01.004, 2005.

685 Esper, J., Cook, E. R., and Schweingruber, F. H.: Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability, Science, 295, 2250–2253, https://doi.org/10.1126/science.1066208, 2002.

Esper, J., George, S. S., Anchukaitis, K., D'Arrigo, R., Ljungqvist, F., Luterbacher, J., Schneider, L., Stoffel, M., Wilson, R., and Büntgen, U.: Large-scale, millennial-length temperature reconstructions from tree-rings, Dendrochronologia, 50, 81–90, https://doi.org/10.1016/j.dendro.2018.06.001, 2018.

690 Evans, M. N., Tolwinski-Ward, S. E., Thompson, D. M., and Anchukaitis, K. J.: Applications of proxy system modeling in high resolution paleoclimatology, Quaternary Science Reviews, 76, 16–28, https://doi.org/10.1016/j.quascirev.2013.05.024, http://dx.doi.org/10.1016/j.quascirev.2013.05.024, 2013.

Franke, J., Brönnimann, S., Bhend, J., and Brugnara, Y.: A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past climatic variations, Scientific Data, 4, 1–19, https://doi.org/10.1038/sdata.2017.76, 2017.

695 Fritts, H., Blasing, T., Hayden, B. P., and Kutzbach, J.: Multivariate Techniques for Specifying Tree-Growth and Climate Relationships and for Reconstructing Anomalies in Paleoclimate, Journal of Applied Meteorology, 10, 845–864, 1971.

Fritts, H. C.: Tree rings and climate, Academic Press, London, 1976.

Gea-Izquierdo, G., Guibal, F., Joffre, R., Ourcival, J. M., Simioni, G., and Guiot, J.: Modelling the climatic drivers determining photosynthesis and carbon allocation in evergreen Mediterranean forests using multiproxy long time series, Biogeosciences, 12, 3695–3712,
700 https://doi.org/10.5194/bg-12-3695-2015, 2015.

Gennaretti, F., Gea-Izquierdo, G., Boucher, E., Berninger, F., Arseneault, D., and Guiot, J.: Ecophysiological modeling of the climate imprint on photosynthesis and carbon allocation to the tree stem in the North American boreal forest, Biogeosciences, 14, 4851–4866, https://doi.org/10.5194/bg-14-4851-2017, 2017.

Gennaretti, F., Boucher, E., Nicault, A., Gea-Izquierdo, G., Arseneault, D., Berninger, F., Savard, M. M., Bégin, C., and Guiot, J.: Underes-
705 timation of the Tambora effects in North American taiga ecosystems, Environmental Research Letters, 13, https://doi.org/10.1088/1748-9326/aaac0c, 2018.

Gong, D. and Wang, S.: Definition of Antarctic oscillation index, Geophysical Research Letters, 26, 459–462, https://doi.org/10.1029/1999GL900003, 1999.

Goosse, H., Crespin, E., de Montety, A., Mann, M. E., Renssen, H., and Timmermann, A.: Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation, Journal of Geophysical Research, 115, 1–17, https://doi.org/10.1029/2009jd012737, 2010.

Goosse, H., Crespin, E., Dubinkina, S., Loutre, M. F., Mann, M. E., Renssen, H., Sallaz-Damaz, Y., and Shindell, D.: The role of forcing and internal dynamics in explaining the "Medieval Climate Anomaly", Climate Dynamics, 39, 2847–2866, https://doi.org/10.1007/s00382-012-1297-0, 2012.

Guiot, J., Boucher, E., and Gea-Izquierdo, G.: Process models and model-data fusion in dendroecology, Frontiers in Ecology and Evolution, 2, 1–12, https://doi.org/10.3389/fevo.2014.00052, http://journal.frontiersin.org/article/10.3389/fevo.2014.00052/abstract, 2014.

Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., Steiger, N., and Perkins, W. A.: The last millennium climate reanalysis project: Framework and first results, Journal of Geophysical Research, 121, 6745–6764, https://doi.org/10.1002/2016JD024751, 2016.

Hartig, F., Minunno, F., and Paul, S.: BayesianTools: General-Purpose MCMC and SMC Samplers and Tools for Bayesian Statistics. R package version 0.1.6., https://github.com/florianhartig/BayesianTools, 2019.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hunter, J. D.: Matplotlib: A 2D graphics environment, Computing in Science and Engineering, 9, 90–95, https://doi.org/10.1109/MCSE.2007.55, 2007.

Jones, P. D., Briffa, K. R., Osborn, T. J., Lough, J. M., Van Ommen, T. D., Vinther, B. M., Luterbacher, J., Wahl, E. R., Zwiers, F. W., Mann, M. E., Schmidt, G. A., Ammann, C. M., Buckley, B. M., Cobb, K. M., Esper, J., Goosse, H., Graham, N., Jansen, E., Kiefer, T., Kull, C., Küttel, M., Mosley-Thompson, E., Overpeck, J. T., Riedwyl, N., Schulz, M., Tudhope, A. W., Villalba, R., Wanner, H., Wolff, E., and Xoplaki, E.: High-resolution palaeoclimatology of the last millennium: A review of current status and future prospects, Holocene, 19, 3–49, https://doi.org/10.1177/0959683608098952, 2009.

Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 2003.

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., and Wanner, H.: European Seasonal and Annual Temperature Variability, Trends, and Extremes since 1500, Science, 303, 1499–1503, https://doi.org/10.1126/science.1093877, 2004.

Mann, M. E., Bradley, R. S., and Hughes, M. K.: Global-scale temperature patterns and climate forcing over the past six centuries, Nature, 392, 779–787, https://doi.org/10.1038/nature02478, 1998.

Mann, M. E., Bradley, R. S., and Hughes, M. K.: Northern Hemisphere temperatures during the past millennium, Geophysical Research Letters, 26, 759–762, 1999.

Mann, M. E., Rutherford, S., Wahl, E., and Ammann, C.: Robustness of proxy-based climate field reconstruction methods, Journal of Geophysical Research Atmospheres, 112, 1–18, https://doi.org/10.1029/2006JD008272, 2007.

Mann, M. E., Zhang, Z., Hughes, M. K., Bradley, R. S., Miller, S. K., Rutherford, S., and Ni, F.: Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia, Proceedings of the National Academy of Sciences, 105, 13 252–13 257, https://doi.org/10.1073/pnas.0805721105, 2008.

Marshall, G. J.: Trends in the Southern Annular Mode from observations and reanalyses, Journal of Climate, 16, 4134–4143, 2003.

Matsikaris, A., Widmann, M., and Jungclaus, J.: On-line and off-line data assimilation in palaeoclimatology: A case study, Climate of the Past, 11, 81–93, https://doi.org/10.5194/cp-11-81-2015, 2015.

Medley, B. and Thomas, E. R.: Increased snowfall over the Antarctic Ice Sheet mitigated twentieth-century sea-level rise, Nature Climate Change, 9, 34–39, https://doi.org/10.1038/s41558-018-0356-x, http://dx.doi.org/10.1038/s41558-018-0356-x, 2019.

Medley, B., McConnell, J. R., Neumann, T. A., Reijmer, C. H., Chellman, N., Sigl, M., and Kipfstuhl, S.: Temperature and Snowfall in Western Queen Maud Land Increasing Faster Than Climate Model Projections, Geophysical Research Letters, 45, 1472–1480, https://doi.org/10.1002/2017GL075992, 2018.

Misson, L.: MAIDEN: a model for analyzing ecosystem processes in dendroecology, Canadian Journal of Forest Research, 34, 874–887, https://doi.org/10.1139/x03-252, 2004.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H.: Anthropogenic and Natural Radiative Forcing, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, USA., https://doi.org/10.1017/CBO9781107415324.018, 2013.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I - A discussion of principles, Journal of Hydrology, 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Neukom, R., Luterbacher, J., Villalba, R., Küttel, M., Frank, D., Jones, P. D., Grosjean, M., Wanner, H., Aravena, J. C., Black, D. E., Christie, D. A., D'Arrigo, R., Lara, A., Morales, M., Soliz-Gamboa, C., Srur, A., Urrutia, R., and von Gunten, L.: Multiproxy summer and winter surface air temperature field reconstructions for southern South America covering the past centuries, Climate Dynamics, 37, 35–51, https://doi.org/10.1007/s00382-010-0793-3, 2011.

Neukom, R., Gergis, J., Karoly, D. J., Wanner, H., Curran, M., Elbert, J., González-Rouco, F., Linsley, B. K., Moy, A. D., Mundo, I., Raible, C. C., Steig, E. J., Van Ommen, T., Vance, T., Villalba, R., Zinke, J., and Frank, D.: Inter-hemispheric temperature variability over the past millennium, Nature Climate Change, 4, 362–367, https://doi.org/10.1038/nclimate2174, 2014.

Okazaki, A., Miyoshi, T., Yoshimura, K., Greybush, S. J., and Zhang, F.: Revisiting Online and Offline Data Assimilation Comparison for Paleoclimate Reconstruction: An Idealized OSSE Study, Journal of Geophysical Research: Atmospheres, 126, 1–17, https://doi.org/10.1029/2020JD034214, 2021.

Oke, P. R. and Sakov, P.: Representation error of oceanic observations for data assimilation, Journal of Atmospheric and Oceanic Technology, 25, 1004–1017, https://doi.org/10.1175/2007JTECHO558.1, 2008.

PAGES 2k Consortium: Continental-scale temperature variability during the past two millennia, Nature Geoscience, 6, 339–346, https://doi.org/10.1038/ngeo1797, 2013.

PAGES 2k Consortium: A global multiproxy database for temperature reconstructions of the Common Era, Scientific Data, 4, 170 088, https://doi.org/10.1038/sdata.2017.88, 2017.

Rezsöhazy, J.: Data-model fusion approach using tree-ring width series, a mechanistic growth model and an Earth system model to reconstruct the large-scale climate variability, Ph.D. thesis, Université catholique de Louvain - Aix Marseille Université, 2021.

Rezsöhazy, J., Goosse, H., Guiot, J., Gennaretti, F., Boucher, E., André, F., and Jonard, M.: Application and evaluation of the dendroclimatic process-based model MAIDEN during the last century in Canada and Europe, Climate of The Past, 16, 1043–1059, https://doi.org/10.5194/cp-16-1043-2020, 2020.

785 Rezsöhazy, J., Gennaretti, F., Goosse, H., and Guiot, J.: Testing the performance of dendroclimatic process-based models at global scale with the PAGES2k tree-ring width database, Climate Dynamics, 57, 2005–2020, https://doi.org/10.1007/s00382-021-05789-7, 2021.

Rubino, M., Etheridge, D. M., Thornton, D. P., Howden, R., Allison, C. E., Francey, R. J., Langenfelds, R. L., Steele, P. L., Trudinger, C. M., Spencer, D. A., Curran, M. A. J., Van Ommen, T. D., and Smith, A. M.: Revised records of atmospheric trace gases CO2, CH4, N2O and d13C-O2 over the last 2000 years from Law Dome, Antarctica, Earth System Science Data, pp. 473–492, https://doi.org/10.5194/essd-11-473-2019, 2019.

790

Schwarz, G.: Estimating the Dimension of a Model, The Annals of Statistics, 6, 461–464, https://doi.org/10.1214/aos/1176344136, 1978.

Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, Journal of Climate, 19, 3088–3111, https://doi.org/10.1175/JCLI3790.1, 2006.

Souverijns, N., Gossart, A., Gorodetskaya, I. V., Lhermitte, S., Mangold, A., Laffineur, Q., Delcloo, A., and Van Lipzig, N. P.: How does

795 the ice sheet surface mass balance relate to snowfall? Insights from a ground-based precipitation radar in East Antarctica, Cryosphere, 12, 1987–2003, https://doi.org/10.5194/tc-12-1987-2018, 2018.

Steiger, N. J. and Smerdon, J. E.: A pseudoproxy assessment of data assimilation for reconstructing the atmosphere-ocean dynamics of hydroclimate extremes, Climate of the Past, 13, 1435–1449, https://doi.org/10.5194/cp-13-1435-2017, 2017.

Steiger, N. J., Hakim, G. J., Steig, E. J., Battisti, D. S., and Roe, G. H.: Assimilation of time-averaged pseudoproxies for climate reconstruc-

800 tion, Journal of Climate, 27, 426–441, https://doi.org/10.1175/JCLI-D-12-00693.1, 2014.

Steiger, N. J., Smerdon, J. E., Cook, E. R., and Cook, B. I.: A reconstruction of global hydroclimate and dynamical variables over the Common Era, Scientific Data, 5, 1–15, https://doi.org/10.1038/sdata.2018.86, 2018.

Stenni, B., Curran, M. A., Abram, N. J., Orsi, A., Goursaud, S., Masson-Delmotte, V., Neukom, R., Goosse, H., Divine, D., Van Ommen, T., Steig, E. J., Dixon, D. A., Thomas, E. R., Bertler, N. A., Isaksson, E., Ekaykin, A., Werner, M., and Frezzotti, M.: Antarctic climate

805 variability on regional and continental scales over the last 2000 years, Climate of the Past, 13, 1609–1634, https://doi.org/10.5194/cp-13-1609-2017, 2017.

Stevenson, S., Otto-Bliesner, B. L., Brady, E. C., Nusbaumer, J., Tabor, C., Tomas, R., Noone, D. C., and Liu, Z.: Volcanic Eruption Signatures in the Isotope-Enabled Last Millennium Ensemble, Paleoceanography and Paleoclimatology, 34, 1534–1552, https://doi.org/10.1029/2019PA003625, 2019.

810 Tardif, R., Hakim, G. J., Perkins, W. A., Horlick, K. A., Erb, M. P., Emile-Geay, J., Anderson, D. M., Steig, E. J., and Noone, D.: Last Millennium Reanalysis with an expanded proxy database and seasonal proxy modeling, Climate of the Past, 15, 1251–1273, https://doi.org/10.5194/cp-15-1251-2019, 2019.

Thomas, E. R., Melchior Van Wessem, J., Roberts, J., Isaksson, E., Schlosser, E., Fudge, T. J., Vallelonga, P., Medley, B., Lenaerts, J., Bertler, N., Van Den Broeke, M. R., Dixon, D. A., Frezzotti, M., Stenni, B., Curran, M., and Ekaykin, A. A.: Regional Antarctic snow

815 accumulation over the past 1000 years, Climate of the Past, 13, 1491–1513, https://doi.org/10.5194/cp-13-1491-2017, 2017.

Tingley, M. P. and Huybers, P.: Recent temperature extremes at high northern latitudes unprecedented in the past 600 years, Nature, 496, 201–205, https://doi.org/10.1038/nature11969, 2013.

Tolwinski-Ward, S. E., Evans, M. N., Hughes, M. K., and Anchukaitis, K. J.: An efficient forward model of the climate controls on interannual variation in tree-ring width, Climate Dynamics, 36, 2419–2439, https://doi.org/10.1007/s00382-010-0945-5, 2011.

Vaganov, E. A., Hughes, M. K., and Shashkin, A.: Growth dynamics of conifer tree rings, New York, springer edn., 2006.

Valler, V., Brugnara, Y., and Bronnimann, S.: Assimilating monthly precipitation data in a paleoclimate data assimilation framework, Climate of the Past, 16, 1309–1323, https://doi.org/10.5194/cp-16-1309-2020, 2020.

van Leeuwen, P. J.: Particle Filtering in Geophysical Systems, Monthly Weather Review, 137, 4089–4114, https://doi.org/10.1175/2009MWR2835.1, http://journals.ametsoc.org/doi/abs/10.1175/2009MWR2835.1, 2009.

van Wessem, J. M., van de Berg, W. J., Noël, B. P. Y., van Meijgaard, E., 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, 1479–1498, https://doi.org/10.5194/tc-12-1479-2018, 2018.

Wang, J., Yang, B., and Ljungqvist, F. C.: A millennial summer temperature reconstruction for the eastern tibetan plateau from tree-ring width, Journal of Climate, 28, 5289–5304, https://doi.org/10.1175/JCLI-D-14-00738.1, 2015.

Widmann, M., Goosse, H., Van Der Schrier, G., Schnur, R., and Barkmeijer, J.: Using data assimilation to study extratropical Northern Hemisphere climate over the last millennium, Climate of the Past, 6, 627–644, https://doi.org/10.5194/cp-6-627-2010, 2010.

Wilmking, M., van der Maaten-Theunissen, M., van der Maaten, E., Scharnweber, T., Buras, A., Biermann, C., Gurskaya, M., Hallinger, M., Lange, J., Shetti, R., Smiljanic, M., and Trouillier, M.: Global assessment of relationships between climate and tree growth, Global Change Biology, 26, 3212–3220, https://doi.org/10.1111/gcb.15057, 2020.

Wilson, R. and Elling, W.: Temporal instability in tree-growth/climate response in the Lower Bavarian Forest region: Implications for dendroclimatic reconstruction, Trees - Structure and Function, 18, 19–28, https://doi.org/10.1007/s00468-003-0273-z, 2004.

Wilson, R., D'Arrigo, R., Buckley, B., Büntgen, U., Esper, J., Frank, D., Luckman, B., Payette, S., Vose, R., and Youngblut, D.: A matter of divergence: Tracking recent warming at hemispheric scales using tree ring data, Journal of Geophysical Research Atmospheres, 112, 1–17, https://doi.org/10.1029/2006JD008318, 2007.

Wilson, R., Anchukaitis, K., Briffa, K. R., Büntgen, U., Cook, E., D'Arrigo, R., Davi, N., Esper, J., Frank, D., Gunnarson, B., Hegerl, G., Helama, S., Klesse, S., Krusic, P. J., Linderholm, H. W., Myglan, V., Osborn, T. J., Rydval, M., Schneider, L., Schurer, A., Wiles, G., Zhang, P., and Zorita, E.: Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context, Quaternary Science Reviews, 134, 1–18, https://doi.org/10.1016/j.quascirev.2015.12.005, 2016.

S1: Supplementary materials

Table S1. Pearson correlation coefficient between the observed tree-ring width (TRW) time series and the tree-ring index time series simulated by the regression model (Sect. 2.4.2) over 1950-2000 (calibration period) and 1901-1949 (verification period) at the 21 TRW sites of the PAGES2k databases (PAGES 2k Consortium, 2013, 2017). The site names are from the PAGES2k database: *Aus* for Tasmanian or New Zealand sites; *SAm* for South American sites. Asterisks stand for significant correlation coefficients at the 95% confidence level.

Site	Calibration	Verification
Aus_001	0.42*	-0.09
Aus_002	0.31*	-0.08
Aus_004	0.24	0.15
Aus_005	0.52*	-0.03
Aus_007	0.65*	0.35*
Aus_009	0.43*	-0.06
Aus_030	0.16	0.31*
Aus_031	0.25	-0.05
SAm_006	0.65	0.01
SAm_024	0.09	0.03
SAm_025	030*	0.50*
SAm_029	0.25	0.12
SAm_9	0.29*	-0.06
SAm_10	0.17	0.07
SAm_11	0.57*	0.22
SAm_12	0.31*	0.09
SAm_15	0.18	-0.04
SAm_16	0.18	0.21
SAm_17	0.20	0.22
SAm_18	0.09	-0.07
SAm_19	0.13	-0.15
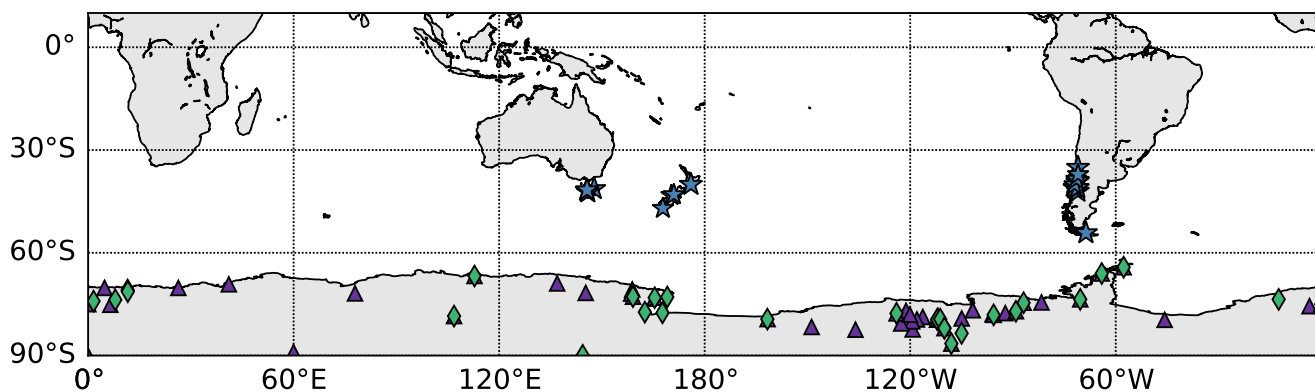
Figure S1. Location of the 21 tree-ring width (TRW) records (blue stars) from the PAGES2k databases (PAGES 2k Consortium, 2013, 2017); the 49 ice core snow accumulation records (purple triangles) from Thomas et al. (2017) and Medley et al. (2018); and the 29 $\delta^{18}O$ records (green diamonds) from Stenni et al. (2017) (Sect. 2.3). Background map from Hunter (2007).
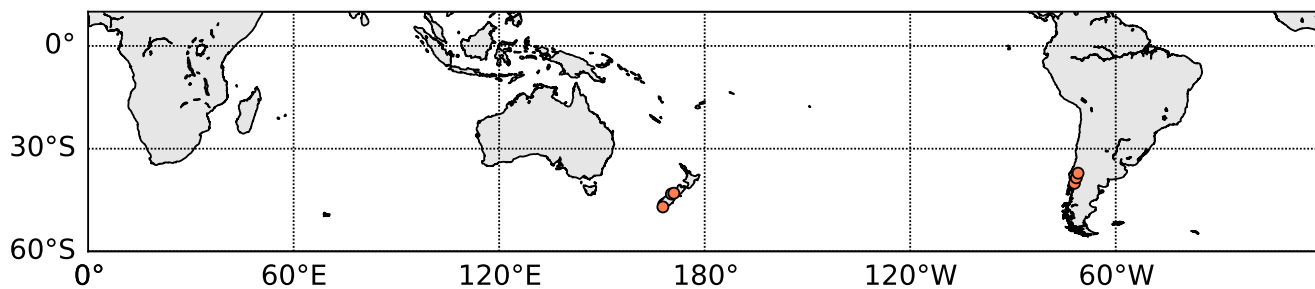
Figure S2. Location of the six tree-ring width (TRW) records (orange circles) from the PAGES2k databases (PAGES 2k Consortium, 2013, 2017) assimilated with the MAIDEN model calibrated over the 1950-2000 CE time period (see Sect. 2.4.1 for more details). Background map from Hunter (2007).
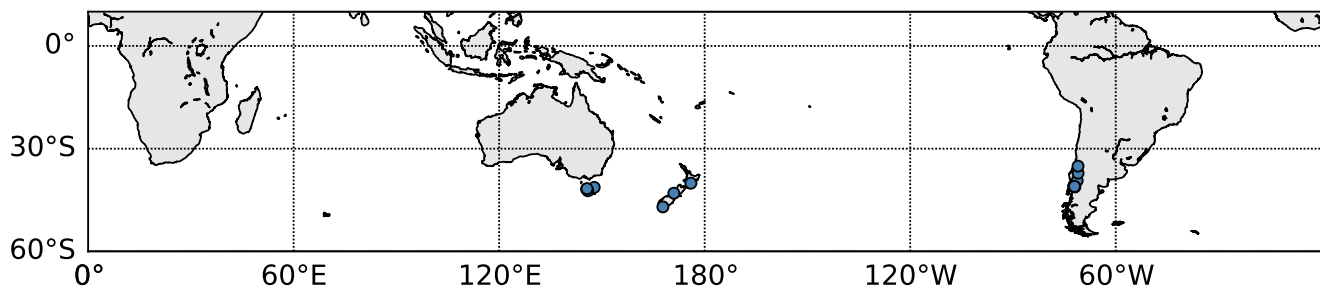
Figure S3. Location of the 12 tree-ring width (TRW) records (blue circles) from the PAGES2k databases (PAGES 2k Consortium, 2013, 2017) assimilated with the regression-based model calibrated over the 1901-2000 CE time period (see Sect. 2.4.2 for more details). Background map from Hunter (2007).
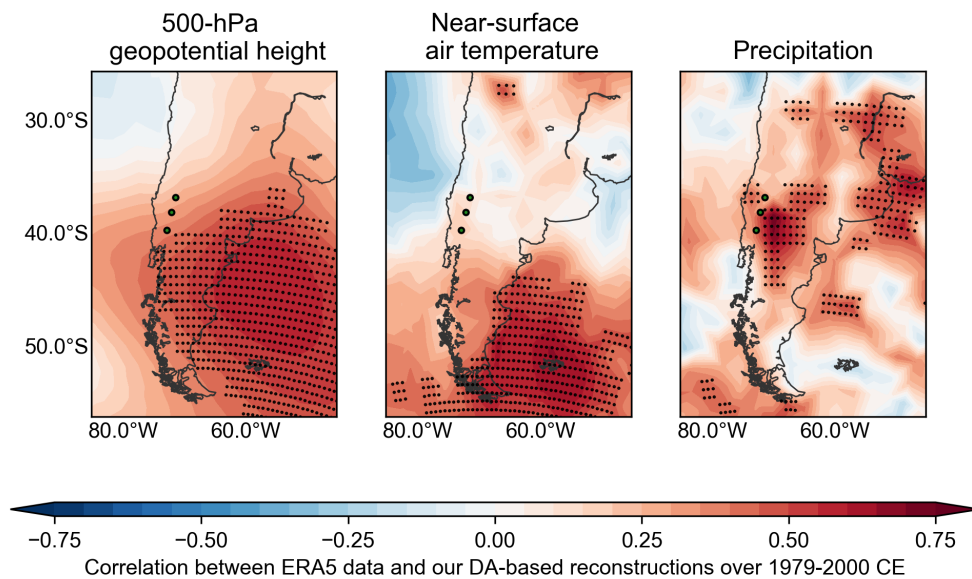
Figure S4. Pearson correlation coefficient between the reconstructed and ERA5 geopotential height at 500 hPa, near-surface air temperature or cumulative precipitation in South America over 1979-2000 CE for the TIC-MDN experiment (see Sect. 3 and Table 1 for details on the experiment). The green dots indicate the localization of the assimilated tree-ring width records. The black dots indicate a significant correlation coefficient at the 95% confidence level.
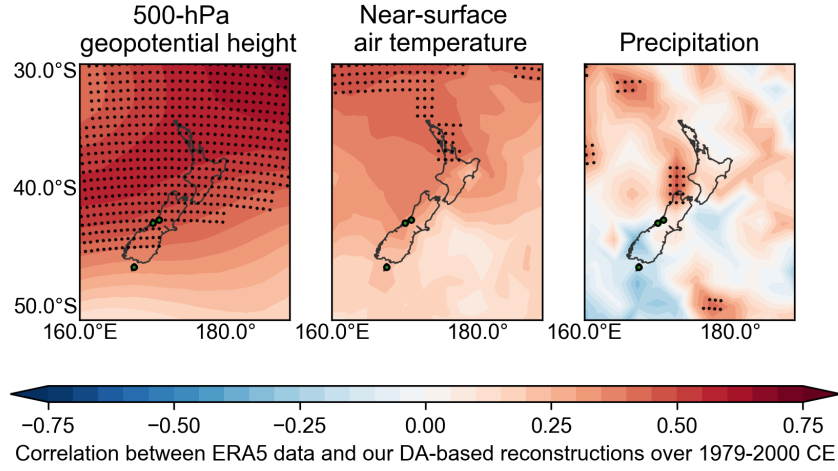
Figure S5. Pearson correlation coefficient between the reconstructed and ERA5 geopotential height at 500 hPa, near-surface air temperature or cumulative precipitation in New Zealand over 1979-2000 CE for the TIC-MDN experiment (see Sect. 3 and Table 1 for details on the experiment). The green dots indicate the localization of the assimilated tree-ring width records. The black dots indicate a significant correlation coefficient at the 95% confidence level.
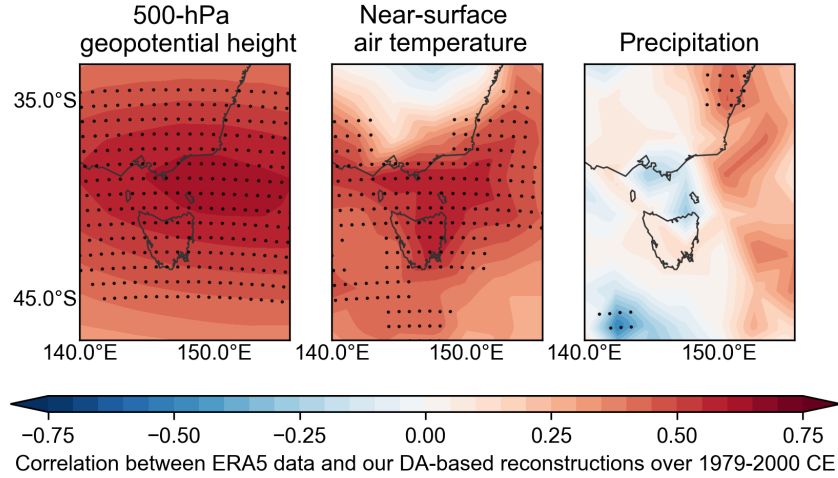
Figure S6. Pearson correlation coefficient between the reconstructed and ERA5 geopotential height at 500 hPa, near-surface air temperature or cumulative precipitation in Tasmania over 1979-2000 CE for the TIC-MDN experiment (see Sect. 3 and Table 1 for details on the experiment). The green dots indicate the localization of the assimilated tree-ring width records. The black dots indicate a significant correlation coefficient at the 95% confidence level.
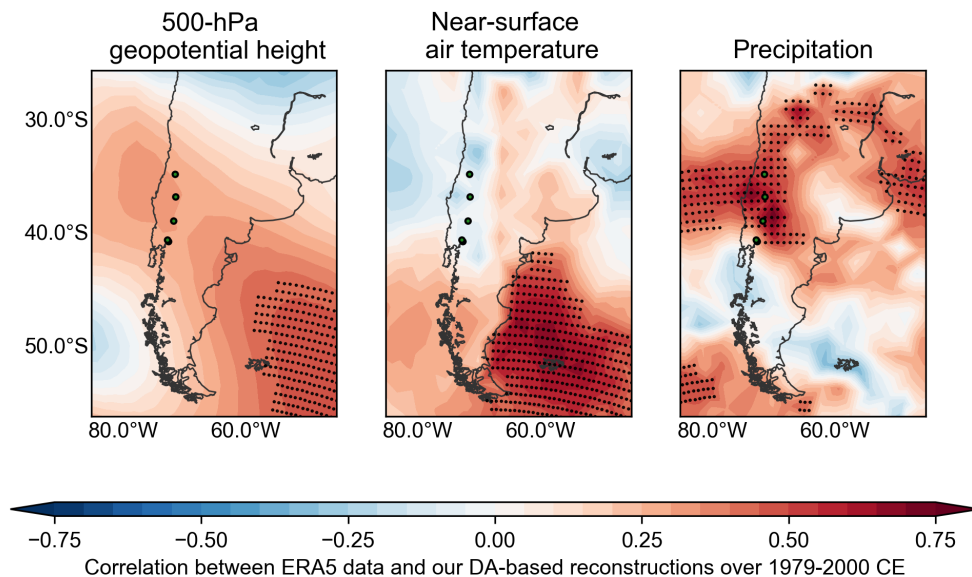
Figure S7. Pearson correlation coefficient between the reconstructed and ERA5 geopotential height at 500 hPa, near-surface air temperature or cumulative precipitation in South America over 1979-2000 CE for the TIC-reg experiment (see Sect. 3 and Table 1 for details on the experiment). The green dots indicate the localization of the assimilated tree-ring width records. The black dots indicate a significant correlation coefficient at the 95% confidence level.
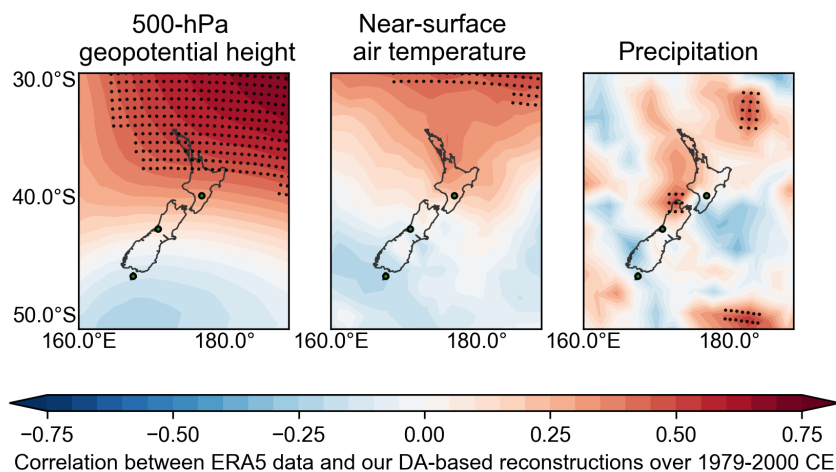
Figure S8. Pearson correlation coefficient between the reconstructed and ERA5 geopotential height at 500 hPa, near-surface air temperature or cumulative precipitation in New Zealand over 1979-2000 CE for the TIC-reg experiment (see Sect. 3 and Table 1 for details on the experiment). The green dots indicate the localization of the assimilated tree-ring width records. The black dots indicate a significant correlation coefficient at the 95% confidence level.
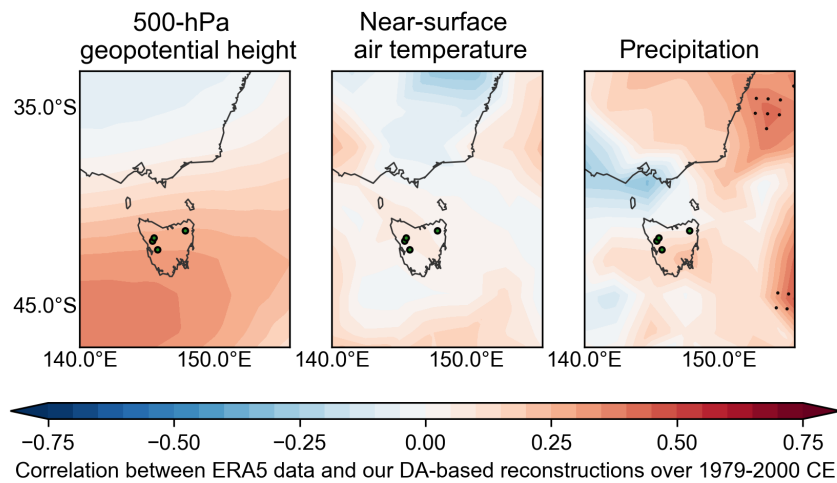
Figure S9. Pearson correlation coefficient between the reconstructed and ERA5 geopotential height at 500 hPa, near-surface air temperature or cumulative precipitation in Tasmania over 1979-2000 CE for the TIC-reg experiment (see Sect. 3 and Table 1 for details on the experiment). The green dots indicate the localization of the assimilated tree-ring width records. The black dots indicate a significant correlation coefficient at the 95% confidence level.
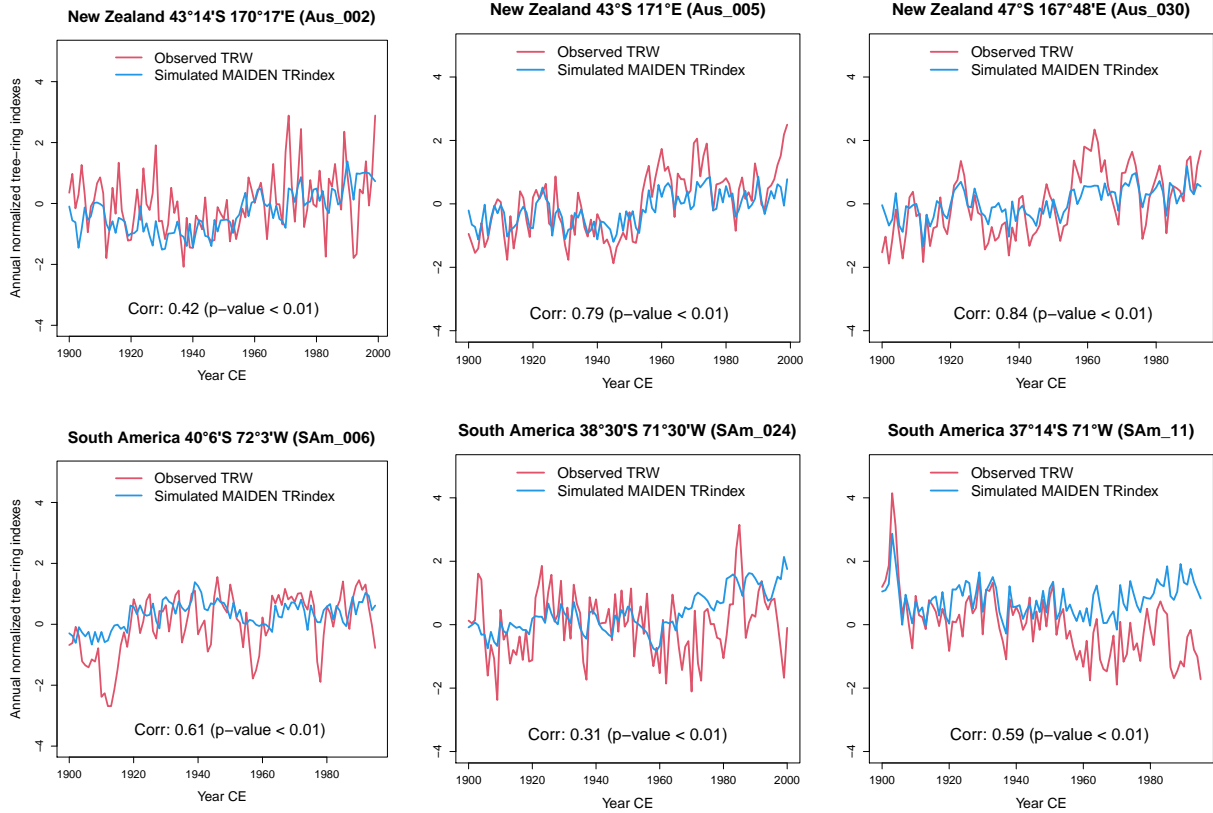
Figure S10. Observed tree-ring width and simulated tree-ring index by MAIDEN over 1900-2000 CE at the six assimilated tree-ring width sites (Sect. 2.4.1). Pearson correlation coefficient (Corr) between observed and simulated tree-ring index and associated p-value. The observations are normalized relative to 1900-2000 CE. The simulations are normalized relative to 1900-2000 CE using the mean and the standard deviation of the reference run of MAIDEN with the iCESM1 data used in the data assimilation procedure (Sect. 2.4.1). The TIC-MDN reanalysis is used to run MAIDEN (see Sect. 3 for details on the data assimilation experiment).
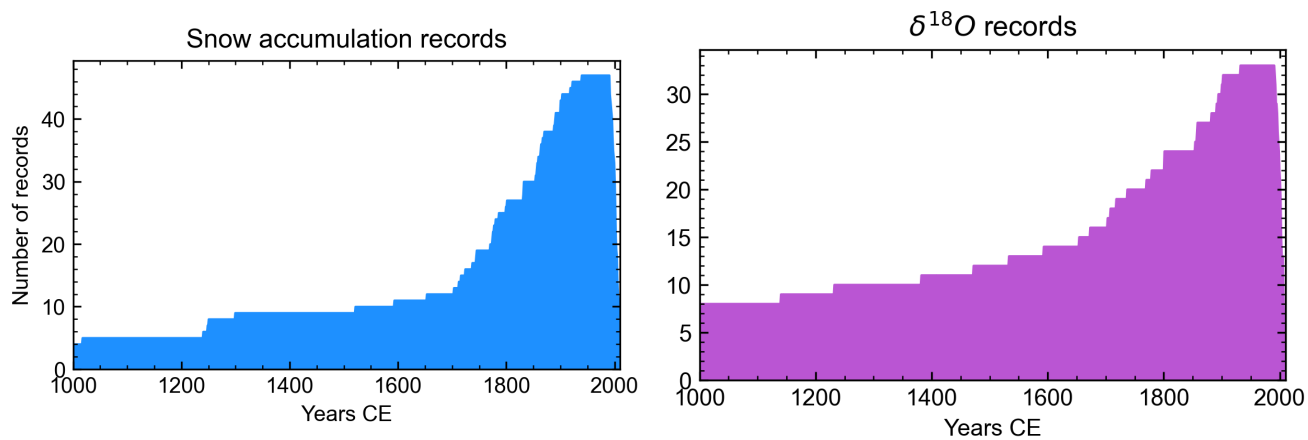
Figure S11. Number of snow accumulation and δ^{18}O records used in this study (Sect. 2.3.2) over 1000-2000 CE. Figures from Dalaiden et al. (2021).
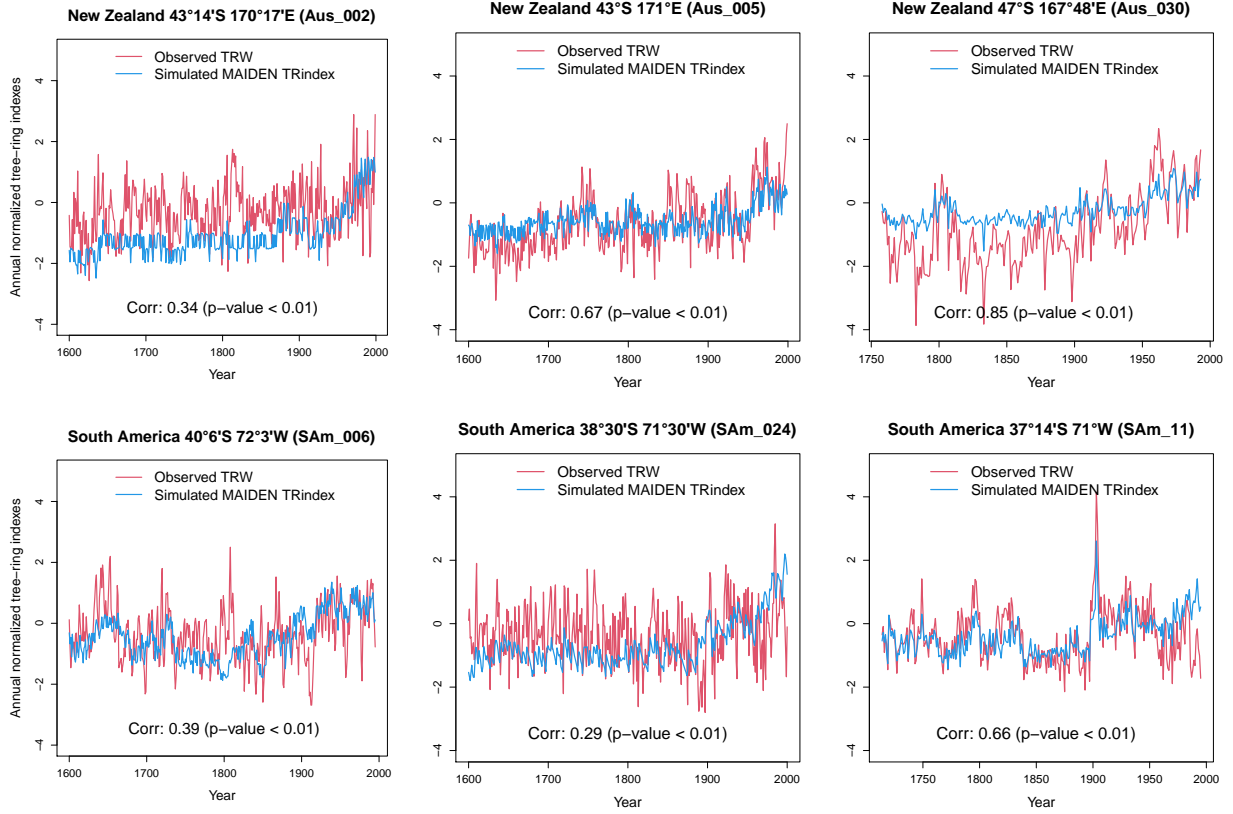
Figure S12. Observed tree-ring width and simulated tree-ring index by MAIDEN over 1600-2000 CE at the six assimilated tree-ring width sites (Sect. 2.4.1). Pearson correlation coefficient (Corr) between observed and simulated tree-ring index and associated p-value. The observations are normalized relative to 1900-2000 CE. The simulations are normalized relative to 1900-2000 CE using the mean and the standard deviation of the reference run of MAIDEN with the iCESM1 data used in the data assimilation procedure (Sect. 2.4.1). The TIC-MDN reanalysis is used to run MAIDEN (see Sect. 3 for details on the data assimilation experiment).