CrossMark

Automated parameter tuning applied to sea ice in a global climate model

Lettie A. Roach[1,2] · Simon F. B. Tett[2] · Michael J. Mineter[2] · Kuniko Yamazaki[2,3] · Cameron D. Rae[2,4]

Abstract This study investigates the hypothesis that a significant portion of spread in climate model projections of sea ice is due to poorly-constrained model parameters. New automated methods for optimization are applied to historical sea ice in a global coupled climate model (HadCM3) in order to calculate the combination of parameters required to reduce the difference between simulation and observations to within the range of model noise. The optimized parameters result in a simulated sea-ice time series which is more consistent with Arctic observations throughout the satellite record (1980-present), particularly in the September minimum, than the standard configuration of HadCM3. Divergence from observed Antarctic trends and mean regional sea ice distribution reflects broader structural uncertainty in the climate model. We also find that the optimized parameters do not cause adverse effects on the model climatology. This simple approach provides evidence for the contribution of parameter uncertainty to spread in sea ice extent trends and could be customized to investigate uncertainties in other climate variables.

1 Introduction

Lack of consistency between observed sea ice extent trends and those simulated by complex climate models has garnered significant public and scientific attention in recent years. This interest is partly driven by the radiative properties of sea ice which amplify climate change as well as the implications of sea ice change for local communities and ecosystems. Yet simulations from the latest Coupled Model Intercomparison Project (CMIP5) continue to exhibit various biases in their mean state, particularly during winter months (Stroeve et al. 2012; Turner et al. 2013), as well as in the magnitude and even sign of hemispheric trends (Flato et al. 2013).

Such model biases in recent historical sea ice can broadly be attributed to two sources of uncertainty: firstly, internal variability arising due to the chaotic nature of climate, and secondly, structural uncertainty arising from incomplete knowledge of and inability to represent the complete physics of the climate system (Hawkins and Sutton 2009). Parametrizations of complex processes may reduce spatially or temporally evolving variables to a single constant value or approximate a number of processes by some non-physical constant. Uncertainty in parameter values has previously been quantified by constructing large perturbed physics ensembles, where parameters are varied within physically reasonable ranges to construct a range of climatologies (Murphy et al. 2004; Stainforth et al. 2005). Sea ice parameter uncertainty can be larger than that expected from natural variability (Ridley et al. 2007) and

✉ Lettie A. Roach
  lettie.roach@niwa.co.nz

[1] Present Address: National Institute of Water and Atmospheric Research, 301 Evans Bay Parade, Greta Point, Wellington, New Zealand

[2] School of GeoSciences, University of Edinburgh, Crew Building, Alexander Crum Brown Road, Edinburgh EH9 3FF, UK

[3] Present Address: Met Office, Fitzroy Road, Exeter, Devon EX1 3PB, UK

[4] Present Address: Centre for Atmospheric Science, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

from inter-model spread (Hodson et al. 2013), motivating its reduction.

Recent years have seen the development of methods for automated parameter optimization (Severijns and Hazeleger 2005; Tett et al. 2013; Zhang et al. 2015). These are motivated by the time-consuming and subjective nature of manual parameter tuning, whereby selected parameters are perturbed in a number of simulations to find a configuration which improves the specified focus of the modeller (Mauritsen et al. 2012; Voosen 2016; Hourdin et al. 2016). New automatic optimization software has been developed that uses cost function type methods to calculate the combination of parameters required to reduce the difference between model simulation and target observations for user-specified variables, whilst accounting for their non-linear dependence on parameters. Due to the high computational cost involved in many of the methods proposed, automated parameter optimization has only been tested on atmosphere-only [eg. Zhang et al. (2015)] or very low-resolution [eg. Gregoire et al. (2011)] models thus far.

The innovation of this paper is to apply an automated parameter tuning method to a widely-used global, fully-coupled climate model in order to investigate whether targeting sea ice extent can improve its consistency with observations. Changes to the model must use physically allowed parameter values, account for internal variability and not compromise other climate variables. We implement the Tett et al. (2013) method for automated tuning, discussing optimization input in Sect. 2.2 and parameters in Sect. 3 to ensure these criteria are satisfied, and evaluate the results in Sect. 4 by comparison to observations.

2 Methods

Most data in this study comes from global climate model simulations, which are discussed below. Observational data is from the HadISST set of satellite sea ice concentrations (Rayner et al. 2003) and the Pan-Arctic Ice Ocean Modeling and Assimilation System sea ice volume reanalysis [PIOMAS, Schweiger et al. (2011)]. We show HadISST sea ice extents from 1941, but note that data is derived from hand-drawn charts, with limited Antarctic sources, before the beginning of satellite observations in 1978 (Rayner et al. 2003).

2.1 Model

The model used in this study is the Hadley Centre Coupled Model version 3 (HadCM3) (Gordon et al. 2000), a global model with full coupling between the ocean, atmosphere and sea ice components. The ocean component has a horizontal resolution of $1.25° \times 1.25°$ with 20 vertical levels and the atmosphere component has a lower resolution of $3.75° \times 2.5°$ with 19 vertical levels.

The sea ice component is described in Cattle et al. (1995) and we summarize the main features here, noting that it is a simple model. Sea ice is advected with the top ocean current $\mathbf{v}$ and mixed laterally by unresolved wind and ocean drags, with magnitude determined by coefficient of turbulent diffusion A_H. Hence the time evolution of ice thickness g_i is given by

$$\frac{\partial g_i}{\partial t} = \nabla.(\mathbf{v}g_i\Theta(4 - g)) + A_H\nabla^2 g_i + \Delta_{\text{thermo}}g_i \qquad (1)$$

where the Heaviside step function prevents ice convergence once it has reached a depth of 4 m. Sea ice forms in the ocean when the sea surface temperature reaches a freezing temperature $T_f = -1.8°C$. For thermodynamic changes, ice and snow are treated as a single layer, though which heat conduction H_I is given by

$$H_I = \frac{\kappa_s(T_s - T_f)}{h_s + (h_i\kappa_s/\kappa_i)} \qquad (2)$$

where κ_i and κ_s are the thermal conductivities of ice and snow, h_i and h_s are the ice and snow layer thicknesses, and T_S is the temperature of the top surface. The heat flux from the ocean into the bottom of the ice H_o is given by

$$H_o = \frac{\rho c_w}{0.5d}K(T_o - T_f) \qquad (3)$$

where ρ and c_w are the density and specific heat capacity of sea water, the first ocean layer has temperature T_o and depth $d = 10$ m and K is the ocean-ice diffusion coefficient. To account for leads, sea ice concentration is not allowed to exceed $c_{\max}^{NH} = 99.5\%$ in the Arctic and $c_{\max}^{SH} = 98.0\%$ in the Antarctic. The model assumes constant salinity (0.6%) and a linear temperature profile in the ice. Albedo is linearly parametrized, with a warm ice albedo of $\alpha_w = 0.5$ at an ice temperature of 0 °C which increases linearly to a cold ice albedo of $\alpha_c = 0.8$ at the defined cold-ice temperature of $T_{\alpha_c} = -10$ °C and below. In the ocean, horizontal eddy mixing of tracers such as salinity and temperature is parametrized. In particular, along-isopycnal diffusion is parametrized by a single diffusion constant D.

In contrast to HadCM3, state-of-the-art sea ice models represent sea ice using multiple thickness categories and multiple ice and snow layers (Hunke et al. 2010). They use rheologies such as the non-linear viscous plastic model (Hibler 1979). The sea ice albedo may depend on snow and ice thickness and spectral properties, and melt ponds are calculated explicitly (Hunke et al. 2010). While the representation of sea ice, as well as other processes, in HadCM3 is much more limited, it was included in CMIP5 and is still widely used for research and impact studies. Its low computational cost makes it an ideal choice for this study.

In this study, 'control' refers to model experiments under pre-industrial conditions where all external forcings are as observed in 1850. All model runs started from initial conditions taken from a 5505 year-long control run. The first year after the 5505-year control simulation is model year one. For parameter perturbation experiments, the parameter change is implemented after model year twenty. While the adjustment period for HadCM3 can be around 350 years (Banks et al. 2007), this timescale is not feasible for us to gauge parameter sensitivity. We observe that initial drifts in sea ice extent following individual diffusion coefficient perturbations relax after approximately ten years (see Fig. 1). Therefore model years 31–40 are used to select parameters. Model years 31–170 are used to examine long-term model climatology and estimate internal variability. The choice of a 140-year period arose from CO_2 doubling experiments which will be written about elsewhere. We assume that internal variability is unaffected by forcings and calculate it from the standard deviation of relevant decadally-averaged observables from the 140 years.

'Forced' runs include a package of observationally-based forcings including greenhouse gases, aerosols and volcanoes from Schurer et al. (2014). The forced run began in model year zero with forcings starting as in AD800 and was branched into four ensemble members at AD1400 (model year 600). Post 2000, forcings follow the A2 emissions scenario. Any sea ice data from the forced standard model shown is the mean of four ensemble members corresponding to the forcing year.
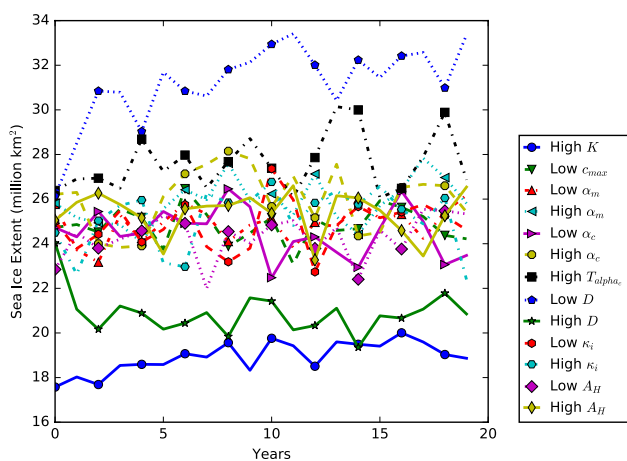
2.2 Optimization

Previous optimization studies [eg. Severijns and Hazeleger (2005); Zhang et al. (2015)] have used the downhill simplex method, which Wright (1996) has suggested is often unreliable. We instead employ a classical Gauss-Newton approach. A simple summary of this optimization technique is provided here and readers are referred to Tett et al. (2013) and Tett et al. (2017) for a fuller description. The aim is to reduce the root mean square error F between a vector of simulated observables $\mathbf{S}$ and a vector of target observables $\mathbf{O}$, with $F^2 = ||\mathbf{S} - \mathbf{O}||^2$. This is a problem in n-dimensional parameter space for n chosen parameters which are allowed to vary within some n-dimensional boundary N. Beginning with the model default parameter values, the algorithm automatically executes the following steps:

1. Runs the model n times in parallel. For each run, one parameter is perturbed individually by +10% of its allowed range away from the starting point.
2. Constructs the error F from each of the n runs, and then calculates a finite-difference Jacobian, which describes how the error evolves under parameter changes.
3. Uses the Gauss-Newton equations to find the sequence in parameter space which converges towards a stationary point for F and defines a search direction $\hat{\mathbf{k}}$.
4. Carries out a line-search at 20, 70 and 100% of the vector $\hat{\mathbf{k}}$ in parameter space and runs the model at these points in parallel.
5. Calculates the error and:

   (a) Uses the point with the smallest error as the start point for the next iteration, or
   (b) Terminates if it has been unable to reduce the error from the start point.

The resulting simulation therefore uses the configuration of constant parameter values which minimize the error F within the boundary of allowed parameter values. Computing resources will dictate the number of model runs that can be carried out in parallel and hence limit parameter dimensionality, which in our case is four.

The algorithm requires: (1) a vector of target observations; (2) parameters which have an impact on the simulated observables and their allowed range; and (3) a simulation design. Below we outline these choices.

Simulated variables for comparison with observations are calculated from sea ice concentration during the first decade where satellite observations are available: 1981 to 1990. Monthly sea ice concentration is averaged over these ten years and then a 15% ice extent cutoff is applied. Due

Fig. 1 The maximum sea ice extent each year in the Southern Hemisphere during the first twenty years after a parameter perturbation is individually perturbed to a high or low value. *Colours* and linestyles are defined in the legend to the *right* of the plot and parameters are described in Table 1

Legend:
- High K
- Low c_{max}
- Low α_m
- High α_m
- Low α_c
- High α_c
- High T_{alpha_c}
- Low D
- High D
- Low κ_i
- High κ_i
- Low A_H
- High A_H

to the high scientific interest in sea ice extent extrema, we tune the maximum and minimum of the monthly sea ice extent cycle in each hemisphere, giving four tuning variables. We choose the maximum and minimum regardless of which month they fall in, as HadCM3's Northern Hemisphere maximum does not always fall in the same month (see Fig. 2). Equivalent target observations are calculated from the HadISST1 set of satellite sea ice concentrations (Rayner et al. 2003) (inadvertently) over 1980 to 1990, but differences between the mean over 1981 to 1990 and over 1980 to 1990 are smaller than model internal variability (Fig. 3). Differences between different observational data sets are considerable (Bunzel et al. 2016) but we choose to neglect this for simplicity.

Tuning is only applicable to models where the difference between the standard model and observations is larger than that expected due to model noise (Notz 2015). We find
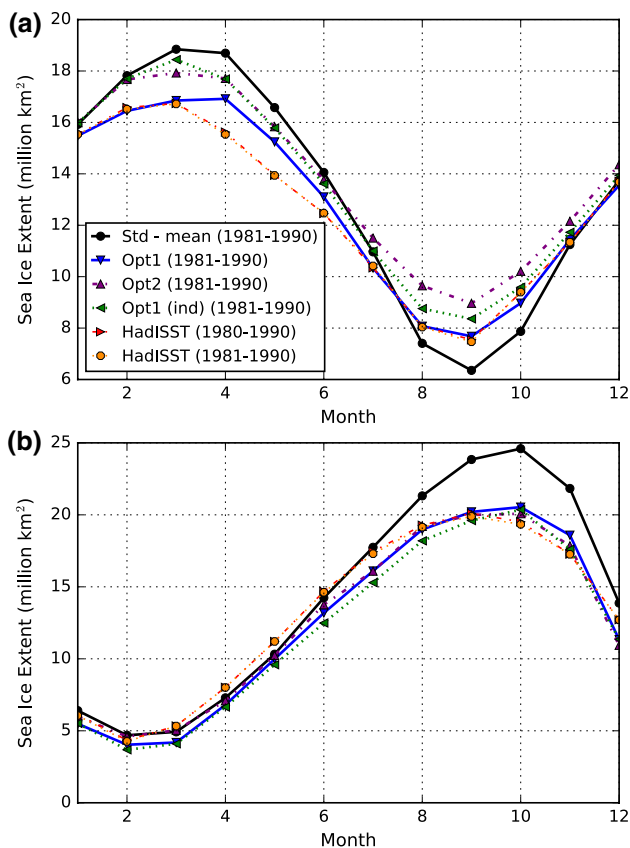
Fig. 2 Sea ice extent calculated using a 15% cut-off from monthly sea ice concentration averaged over 1981–1990 in **a** the Northern Hemisphere and **b** the Southern Hemisphere shown for: the standard model (*black solid line*); the resulting optimized simulation from OPT1 (*blue solid*); the resulting optimized simulation from OPT2 (*purple dashed*); the simulation with OPT1 parameters started in 1940 (*green dashed*) and the HadISST observations over 1981–1990 (*orange dashed*) and over 1980–1990 (*red dashed*). *Red* and *orange dashed lines* are overlying

that the differences between our target and observations are indeed larger than internal variability (Fig. 3).

Internal variability will magnify any minute differences between simulations, so multiple experiments are required to confirm that the stationary point located is not due to chance, thus avoiding selection bias. While ten years may be sufficient for fast adjustment of sea ice extent to individual parameter perturbations, as discussed above, variation of multiple parameters may require a longer adjustment period. We therefore trial two approaches: the first uses a spin-up period of ten years, requiring a total model run length of twenty years to calculate the observables in each of the two steps per iteration (case OPT1) and the second uses spin-up period of twenty years, requiring a total model run length of thirty years (case OPT2). Computational expense limits further trials. Model runs include observationally-based forcings from Schurer et al. (2014).

We expect that multiple stationary points exist, with different parameter combinations resulting in the same or similar error functions. Issues of equifinality are explored in Tett et al. (2013). We do not examine these here, and instead choose to start from the default model parameter values and investigate the first stationary point that is encountered.

3 Parameter selection

A literature review justified variation of eight of the parameters described in Sect. 2.1: K, D, α_m, α_c, T_{α_c}, $c_{max}^{NH}/c_{max}^{SH}$, κ_i and A_H. Many values are adopted from Brierley et al. (2010) and Murphy et al. (2004), who consulted with HadCM3 experts to identify physically reasonable ranges for K, D, α_m and T_{α_c}, but we make some changes as follows. See Table 1 for a summary of extreme values and justifications.

From four years of field measurements, Perovich and Polashenski (2012) find seven steps in albedo evolution. They find a cold snow albedo of 0.85 which lasts between freeze-up and melt onset and warm snow albedo below 0.5 and as low as 0.32 during melt pond drainage. Hence we extend the albedo range suggested by Murphy et al. (2004) to include a lower warm albedo α_m of 0.45 and a higher cold albedo α_c of 0.85.

c_{max} is a parameter unique to HadCM3 which is based on a minimum lead fraction of 0.005 in the Arctic and 0.02 in the Antarctic. However, Lindsay and Rothrock (1995) observe a seasonally-varying lead fraction in the central Arctic, between 0.02 in winter and 0.06 in summer, using a radiometer with a resolution of around 1km. Marcq and Weiss (2012) finds that most leads are very small (<20 m) and hence unresolvable in satellite datasets, which may further reduce sea ice concentrations.

(a) NH Max   **(b)** NH Min   **(c)** SH Max   **(d)** SH Min
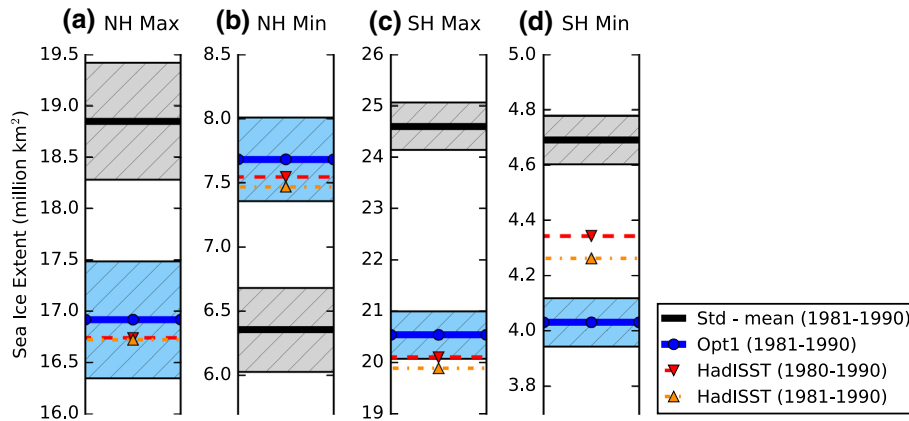
Fig. 3 Sea ice extent calculated using a 15% cut-off from monthly sea ice concentration averaged over 1981–1990 at **a** the maximum of the Northern Hemisphere annual cycle, **b** the minimum of the Northern Hemisphere annual cycle, **c** the maximum of the Southern Hemisphere annual cycle and **d*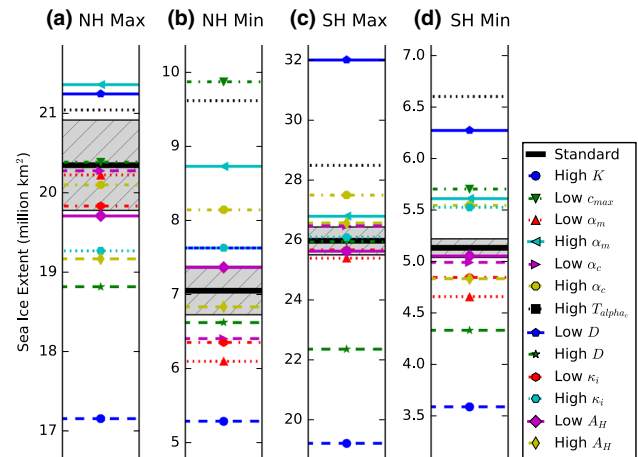* the minimum of the Southern Hemisphere annual cycle. *Solid lines* show the standard and optimized (OPT1) simulations with *hatched areas* to denote internal variability. Equivalent observations calculated from the HadISST data set are shown by *dashed lines*. We also show the observations averaged over 1980–1990

We use the conservative lower estimate of 0.04, so that $c_{max} = 0.96$, as our extreme value. We could not find a comparable observed limiting lead fraction in the literature for the Antarctic and thus assume the same figure as for the Arctic.

Rae et al. (2014) perturb ice conductivity κ_i in a sensitivity study for a later Hadley Centre model, HadGEM3, on the basis of the experimental range summarized by Pringle et al. (2007). They perturb only to the high end of this range; we additionally perturb to the low end of this range.

The very high uncertainty in the standard value of the turbulent diffusion coefficient A_H—with default values ranging from 675 to 2000 $m^2\ s^{-1}$—motivated us to reduce it to zero. We choose its upper limit according to the study where it was originally proposed, Bryan (1969).

Individually perturbing parameters to the extrema of their physically reasonable ranges in short control experiments provides a first-order estimate of parameter uncertainty. We select the four parameters that result in the largest changes in our simulated observables when individually perturbed, those changes being significantly larger than our estimate of internal variability (Fig. 4), and which span a variety of different physical processes. The four parameters are: the ocean-ice diffusion coefficient K, the along-isopycnal ocean diffusion coefficient D, the cold-ice temperature (which controls variation of sea ice albedo) T_{α_c}, and the maximum grid cell sea ice concentration c_{max}. Allowed ranges for these four parameters are further tested in long control experiments. Evidence of drift in the top of atmosphere net flux using the extreme values of D led us to refine its allowed range. The four parameters and their final allowed value ranges are shown in Table 2.

(a) NH Max   **(b)** NH Min   **(c)** SH Max   **(d)** SH Min

Fig. 4 Sea ice extent calculated using a 15% cut-off from monthly sea ice concentration averaged over model years 31–40 (ten years after parameter perturbations) of control simulations at **a** the maximum of the Northern Hemisphere annual cycle, **b** the minimum of the Northern Hemisphere annual cycle, **c** the maximum of the Southern Hemisphere annual cycle and **d** the minimum of the Southern Hemisphere annual cycle. *Lines with markers* show simulations where one parameter is perturbed to a high or low extreme value, the *solid line* shows the standard model and the *grey hatched area* shows our estimate of internal variability. See Table 1 for a description of parameter perturbations

4 Results

Both optimizations converged within four iterations, with four parameter perturbations and three line-searches per iteration plus one standard model run requiring $(4 \times 7 + 1) \times 20 = 580$ and $(4 \times 7 + 1) \times 30 = 870$ years of parallel model runs for OPT1 and OPT2 respectively. In

both cases, most of the error reduction occurred within the first iteration, suggesting the optimization could be run for just one iteration where computer resources are limited.

The parameter values which resulted in the simulation with the smallest error are shown in Table 2. In both trials, K increased, c_{max} decreased, D increased and T_{α_c} remained the same relative to the standard model. The magnitude of changes for K and c_{max} are slightly larger in OPT2, which had a longer spin-up period. The proximity of the final parameter configurations suggests that both experiments have located the same stationary point and provides evidence that the optimization is robust under changes in spin-up period.

The ocean-ice heat diffusion parameter K changed the most from the default value relative to its range. Examining the change in error per unit change in parameter for the Gauss-Newton step of the four iterations in OPT1 shows that perturbing K caused the largest change in error (see Table 3). Increasing K in the first iteration decreased sea ice extent in both hemispheres and seasons. This is not desirable for the Northern Hemisphere minimum, where the standard model has too little ice, but standard model difference from observations is largest in winter (Fig. 2) and so winter improvements are prioritized. We also note that in later Hadley Centre models, K is replaced by a function of friction velocity with a minimum value of $u_* = 0.005$ ms^{-2} [McLaren et al. (2006); McPhee (1992)]. This value would require a minimum value of $K = 1.5 \times 10^{-4}$ m^{2}s^{-1} in the HadCM3 scheme, hence the higher optimized value of K may have some physical basis. Turning to the other parameters, increasing T_{α_c} leads to higher-albedo ice and larger sea ice extent in all four tuning variables in the first iteration, increasing the error. The default value of T_{α_c} lies on the lower bound specified by Murphy et al. (2004) and so it stays constant during the line-search and remains at the default value. The change in error per unit parameter change is smallest for ocean isopycnal diffusion D. It causes decreases in sea ice extent in the Antarctic but increases in the Arctic, which cancel out changes in the error function F. The final value for the maximum allowed sea ice concentration c_{max} is the most different from its default value relative to its allowed range. This is the least physically-based parameter of the four, designed to cover a range of dynamical and thermodynamical processes involved in sea ice fracture and lead opening. As such, we suggest that this parameter is the most reasonable to manipulate to modeller requirements.

These parameter changes cause large error reductions for both optimizations. The root mean square error F between simulation and target reduces by 89% in OPT1 (from 5.3 million km^2 to 0.5 million km^2) and by 64% in OPT2 (to 1.9 million km^2). The difference in error between OPT1 and OPT2 is due to both model noise

and the small differences in parameter values for K, c_{max} and D. For simplicity, we consider only the lowest-error result, OPT1, in the following.

Firstly, we confirm that the control model climatology using OPT1 parameter values is stable. Over a 140 year model run, we find that the model does not exhibit a drift in the top of atmosphere net flux or in surface temperature (Fig. 5a, b). The differences in climatology between the optimized and standard configurations are consistent with there being significantly less global sea ice extent in the optimized case than the standard case (Fig. 5c). For example, relative to the standard control, the optimized control shows some warming (Fig. 6). Given that HadCM3 is biased slightly cold in Northern Hemisphere sea surface temperatures [Gordon et al. (2000)], this is a slight improvement in the model climate. Surface air temperature and ocean surface potential temperature are higher in the optimized model by up to 0.5 K away from the poles (Fig. 6). In the Arctic, surface air temperatures increase by up to 2.5 K in many areas, but also show a decrease of similar magnitude over Siberia and Northern Canada, which contrasts with a more uniform temperature increase of up to around 2 K in the Antarctic. We attribute the cooling in high-latitude ocean surface temperature to the decrease in ocean isopycnal diffusion, D, with a contribution from the increase in ocean-ice heat flux, K, extracting more heat from the ocean during the melt season. These cooler ocean surface temperatures
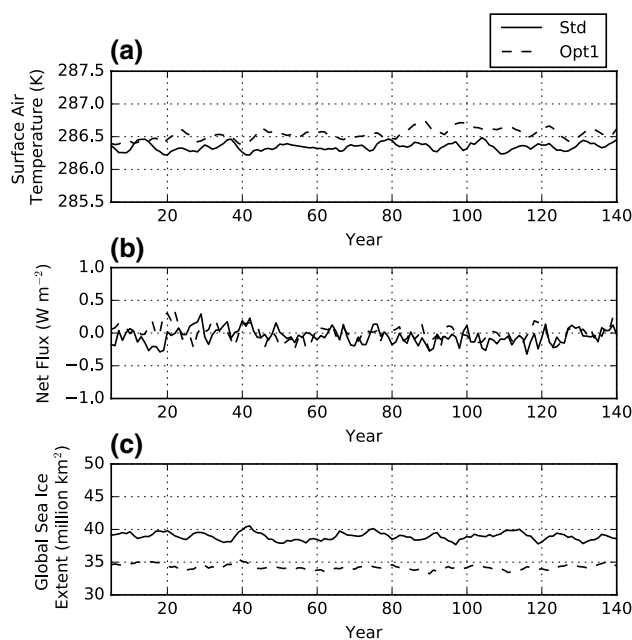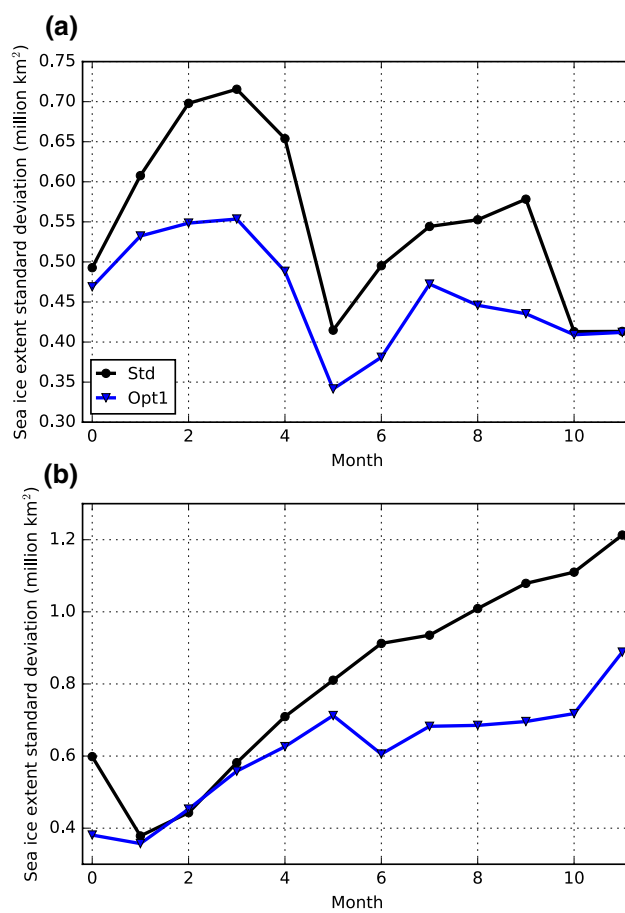

Fig. 5 The global annual average **a** surface air temperature, **b** net top-of-atmosphere flux, and **c** sea ice extent from control simulations using (*solid line*) the standard model parameters and (*dashed line*) the OPT1 parameters. Data are smoothed with a five year moving average

(a)

(b)

(c)
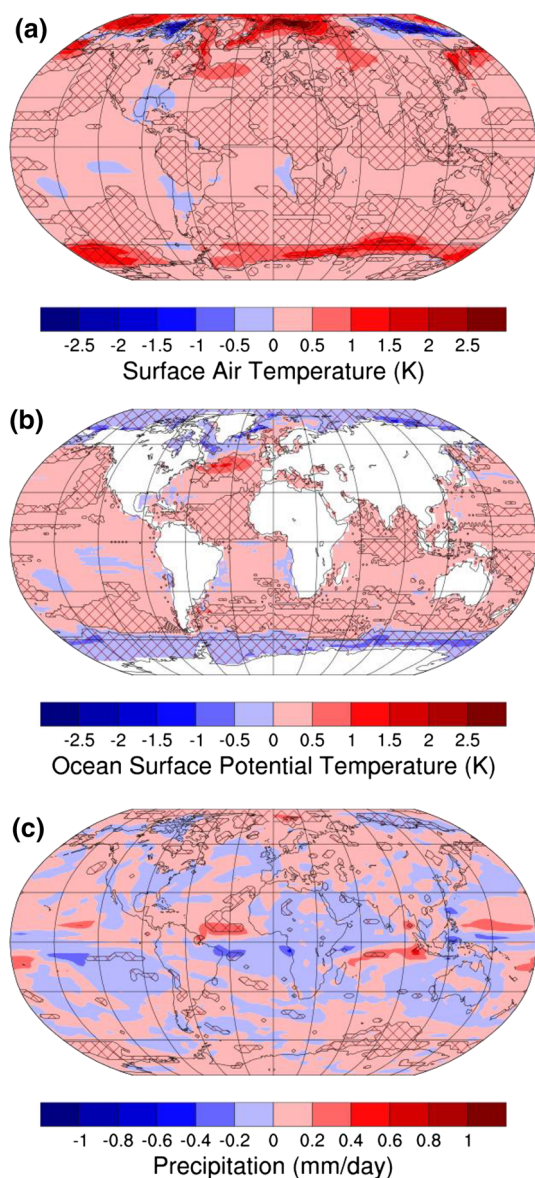
Fig. 6 The difference in 140-year averages between a control simulation using the standard parameters and a control simulation using the OPT1 parameters for **a** surface air temperature, **b** ocean surface potential temperate and **c** precipitation. Hatching shows statistically significant differences at the 95% confidence level

(a)

(b)

Fig. 7 The standard deviation of sea ice extent for (*black line, circular markers*) the standard model and (*blue line, triangular markers*) the optimized model OPT1 in **a** the Northern Hemisphere and **b** the Southern Hemisphere. The standard deviation is calculated for each month of the year over 140 years of control simulation. The large jump between December and January values in the Southern Hemisphere is consistent with the results from Zunz et al. (2013)

may be behind the small increase of 0.5 Sv in the strength of the Atlantic Meridional Overturning Circulation, although this is unlikely to be significant given its large variability (Banks et al. 2007). Precipitation changes are generally smaller than model noise except for small shifts in position of the Intertropical Convergence Zone (ITCZ). Chiang and Bitz (2005) propose that the ITCZ shifts away from an imposed increase in ice cover in a particular hemisphere; in the optimized model both hemispheres have reduced sea ice and there is no clear direction in the ITCZ shift. We conclude that changes in the

model climate are reasonable and suggest perturbations to sea-ice relevant parameters may even reduce other model biases.

Three of the four tuning variables in the optimized model (OPT1) are now consistent with observations (Fig. 3). The Southern Hemisphere maximum is still outside of the range of internal variability, but the difference between model and observations has reduced by around 3 million km^2. Over the full monthly sea ice cycle, the optimized model is much more consistent with observations over 1981 to 1990 than the standard model (despite being optimized to observations for 1980–1990) (Fig. 2). However, we find that the time lag between model and observations noted by Turner et al. (2006) persists. Over the long control experiment, variability is lower in the optimized model than the standard model for nearly all months of the year in both hemispheres (Fig. 7). Zunz et al. (2013) found that HadCM3,

and most CMIP5 models, overestimated Antarctic sea ice extent variance so this result is encouraging.

Having established the agreement between optimized model and observations in terms of the overall amount of sea ice, the tuned observable, we now consider regional differences in sea ice concentration. During winter in the Northern Hemisphere, default HadCM3 produces too much ice, particularly around the east coasts of Canada and Greenland, and the Barents and Okhotsk seas (Fig. 8a). The optimized model strongly reduces sea ice formation in
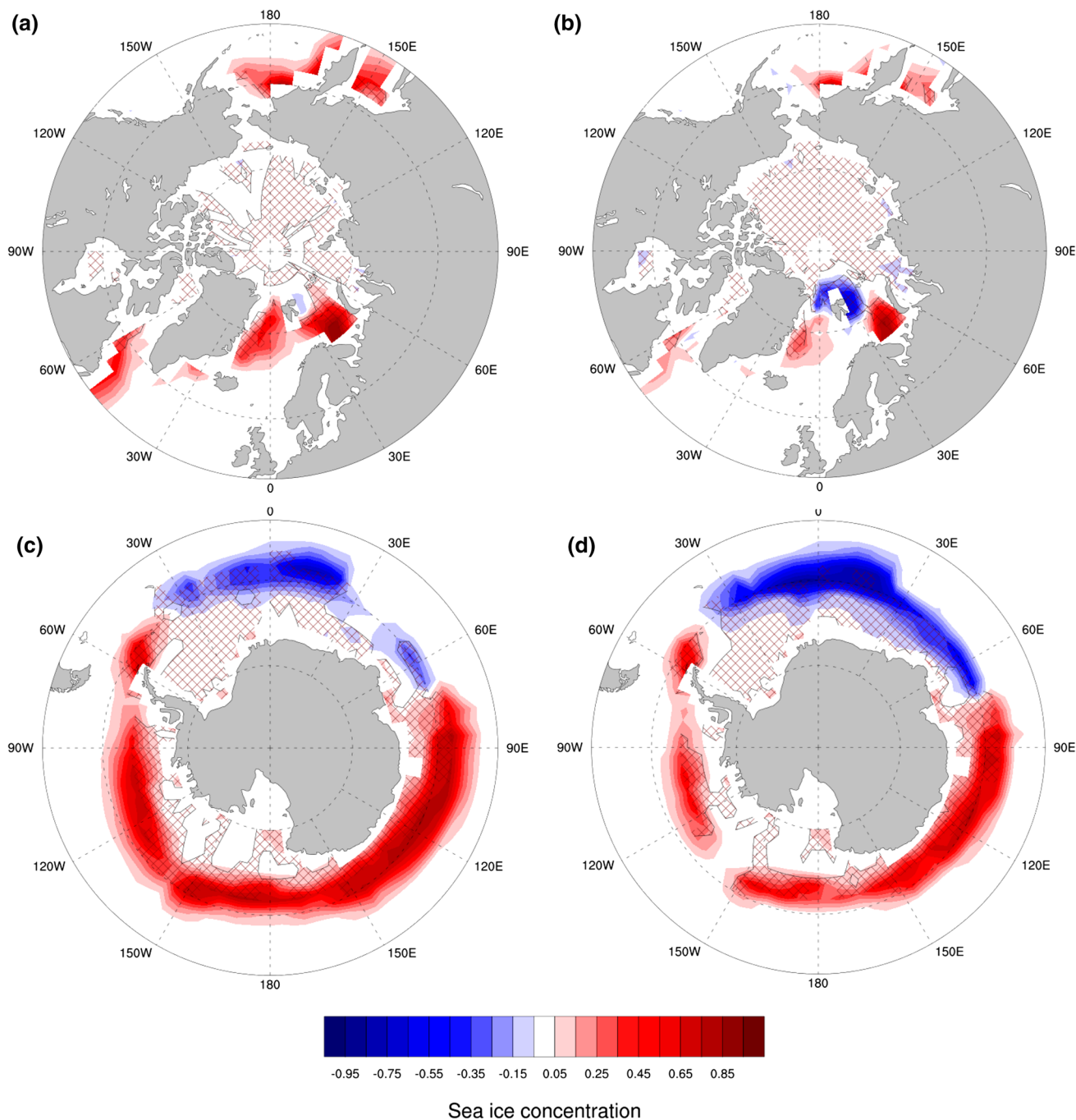
Fig. 8 The difference in sea ice concentration between **a** the standard model and the HadISST observations in March in the Northern Hemisphere, **b** the optimized model (OPT1) and the HadISST observations in March in the Northern Hemisphere, **c** the standard model and the HadISST observations in September in the Southern Hemisphere, and **d** the optimized model (OPT1) and the HadISST observations in September in the Southern Hemisphere. *Blue* shows model underestimation of observations and *red* overestimation. All data are averaged over 1981–1990. HadISST sea ice concentration was interpolated onto the HadCM3 grid. *Hatching* denotes statistically significant differences calculated from a Students t-test over the 10 year period at the 95% confidence level
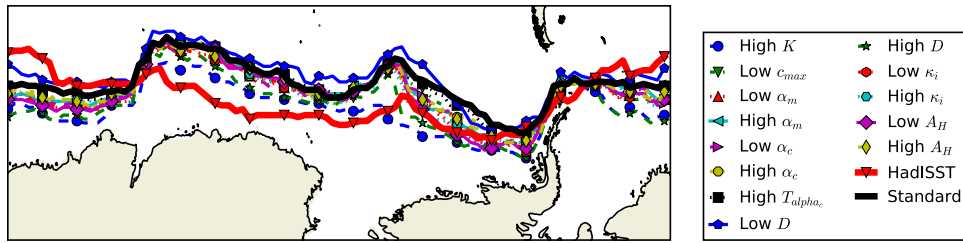
Fig. 9 The location of the 15% sea ice edge in the Antarctic in September for (*black solid line*) the standard model in the control averaged over model years 31–40, (*dashed lines with markers*) parameter perturbed runs in the control averaged over years 31–40 (ten years after parameter perturbations), and (*solid red line*) the HadISST observations averaged over years 1981–1990

those areas. The smallest improvement is seen in the Barents Sea, while an anomalously sea ice-free region appears northeast of Svalbard (Fig. 8b). We attribute this to the decline in magnitude of the north-south temperature gradient observed in the control optimized model relative to the standard model (Fig. 6), although internal variability may also contribute.

A Student's t-test reveals significant ($p < 5\%$) differences between the standard model and observations in the central Arctic, with the number of grid cells in the Northern Hemisphere in March which have statistically significant differences from observations increasing from 373 for the standard model to 491 in the optimized model (out of 3504 Northern Hemisphere grid cells). This is due to the decrease in the limiting parameter c_{max}, which biases the model towards lower concentrations in the central Arctic where observed concentrations are much closer to one. For example, along the latitude line of 85° N in March (mean 1981–1990), HadISST sea ice concentrations are all greater than 95% with one third of these also greater than 99%. The value of c_{max} in the optimized model means that it cannot simulate sea ice concentrations greater than 97%. While more recent models do not use this parameter (McLaren et al. 2006), this issue could be resolved for HadCM3 in future work by a different choice of error function which takes into account differences in concentrations, rather than simply sea ice extent. It should be noted again, however, that there is significant uncertainty in satellite observations (Bunzel et al. 2016). True observed concentrations may be lower due to large numbers of unresolvable leads (Marcq and Weiss 2012) and widely-used satellite datasets differ in their representation of high-concentration grid cells due to differing treatment of melt ponds (Notz 2014).

In the Southern Hemisphere, the optimized simulation does significantly reduce the high bias in sea ice concentration in the Ross and Amundsen seas (Fig. 8d). This mainly occurs because of the increase in ocean-ice heat diffusion; the extreme value of this parameter has the largest impact of of all individual parameter perturbations on the September Antarctic sea ice edge in this area

(Fig. 9). However, the ocean-ice heat diffusion increase simultaneously drives a loss of sea ice around King Haakon VII sea (Fig. 9), which degrades the simulation (Fig. 8d). The reduction in the number of grid cells which are significantly different ($p < 5\%$) from observations
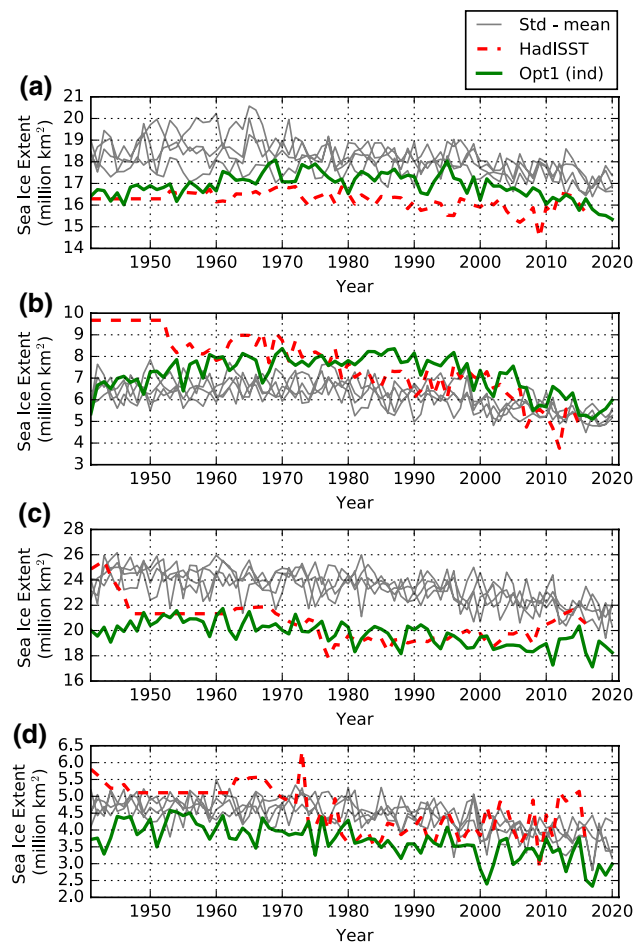

Fig. 10 Sea ice extent each year at **a** the Northern Hemisphere maximum, **b** the Northern Hemisphere minimum, **c** the Southern Hemisphere maximum and **d** the Southern Hemisphere minimum from the standard model ensemble (*thin grey solid lines*), HadISST observations (*red dashed line*) and the simulation with OPT1 parameters started in 1940 (*green solid line*)

between the standard and optimized models (from 172 to 137 in March and from 314 to 286 in September, out of 3504 Southern Hemisphere grid cells) shows that there is an overall improvement. However, both models poorly capture the observed distribution of Antarctic sea ice (Fig. 8c, d), a feature common to a number of the CMIP5 models (Turner et al. 2013). None of the individual parameter perturbations for HadCM3 are able to replicate the shape of the observed sea ice edge, in particular the more extensive area around the King Haakon VII Sea (Fig. 9). This points to uncertainties other than simply that in parameter values (see Discussion), although much of the difference from observations on a regional scale may be due to decadal climate variability.

Finally, we ran the model using the optimized parameters from 1940 under observationally-based forcings.

Tuning parameters to conditions in the early satellite record will not guarantee agreement to observations in other years. The longer spin-up period causes these parameters to perform less well against observations in this experiment even during the tuning period (Fig. 2) compared to the direct optimization output, but still significantly better than the standard model. The mean state from this independent run and root mean square error of 2 million km^2 over 1981–1990 is very similar to that of OPT2, which had a longer spin-up period (Fig. 2). This suggests that a longer spin-up period reduces selection bias in results. The differences in simulated tuning variables between OPT1 and the independent run begun in 1940 are statistically significant ($p < 5\%$) only in the Northern Hemisphere. Only the Northern Hemisphere maximum is additionally outside the range of variability calculated from the standard model,
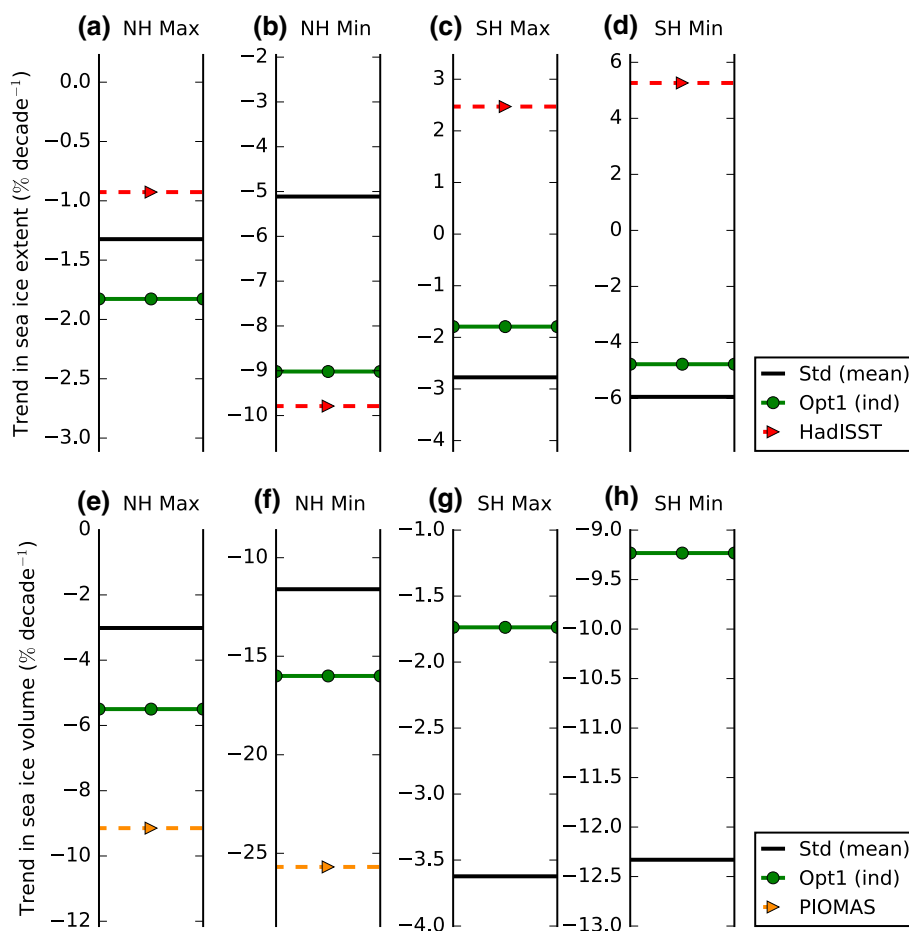

Fig. 11 *Top* the linear trend in sea ice extent from 1980–2014 relative to the 1980–2009 average in **a** the maximum of the Northern Hemisphere annual cycle, **b** the minimum of the Northern Hemisphere annual cycle, **c** the maximum of the Southern Hemisphere annual cycle and **d** the minimum of the Southern Hemisphere annual cycle from the mean of the standard model ensemble (*black solid line*), HadISST observations (*red dashed line*) and the simulation with OPT1 parameters started in 1940 (*green solid line*). *Bottom* the linear trend in sea ice volume from 1980–2014 relative to the 1980–2009 average in **e** the maximum of the Northern Hemisphere annual cycle, **f** the minimum of the Northern Hemisphere annual cycle, **g** the maximum of the Southern Hemisphere annual cycle and **h** the minimum of the Southern Hemisphere annual cycle from the mean of the standard model ensemble (*black solid line*), PIOMAS reanalysis (*orange dashed line*) and the simulation with OPT1 parameters started in 1940 (*green solid line*)

with both OPT2 and the independent run being closer to the standard model in the Northern Hemisphere maximum than OPT1. Examining the intermediate steps in the optimization shows that the Northern Hemisphere maximum took the longest to converge. The Northern Hemisphere maximum may be more difficult to optimize because of the limitations due to surrounding landmasses. It is also likely to have a larger inter-annual correlation than the Southern Hemisphere, as it has a larger proportion of multi-year ice.

The improvement in the mean sea ice state due to optimized parameters causes HadCM3 to more skilfully predict sea ice extent up to the present day (Fig. 10). Linear trends from 1980–2014 relative to the 1980–2009 average in all four variables are statistically significant ($p < 5\%$) and are displayed in Fig. 11a–d. The optimized model fails to capture the observed positive trend in the Antarctic, although trends are less negative compared to the standard model (for example, Southern Hemisphere maximum: standard model −2.8%/decade, OPT1 −1.8%/decade and HadISST +2.5%/decade). This is consistent with the hypothesis that there are significant process-related uncertainties in the Antarctic. Trends in the Northern Hemisphere are more negative in the optimized case (OPT1) than in the standard model, with an improvement towards consistency with observations seen in the Arctic minimum (standard model −5.2%/decade, OPT1 −9.1%/decade and HadISST −9.8%/decade). The absolute trend in the Arctic minimum over 1980–2011 is −0.71 million km²/decade. Comparison to Fig. 3 in Stroeve et al. (2012) reveals that optimization moves the Arctic minimum simulated by HadCM3 to within their calculated observational range, where only 7 of the 18 CMIP5 models examined by that paper lie.

We also examine sea volume and compare time series from the standard and optimized model to PIOMAS reanalysis data in the Arctic (Fig. 12). Given the significant uncertainties in observed sea ice volume, we follow Shu et al. (2015) and show a ±15% error around the PIOMAS data. This reveals that the standard model has a high bias in maximum Northern Hemisphere sea ice volume, which is amplified in the optimized case due to cooler sea surface temperatures (Fig. 6). In the Northern Hemisphere minimum, the PIOMAS reanalysis lies in between the standard and optimized models, and given the large uncertainties in this data, suggests that neither simulation has an unreasonable minimum. The Antarctic minimum is very similar in both the standard and optimized models; the Antarctic maximum is lower in the optimized case but there are no available observations which would confirm whether this is a move towards or away from the observed state. The reduction in Antarctic volume is likely due to the increase in K without a similar decrease in sea surface temperatures. Interestingly,
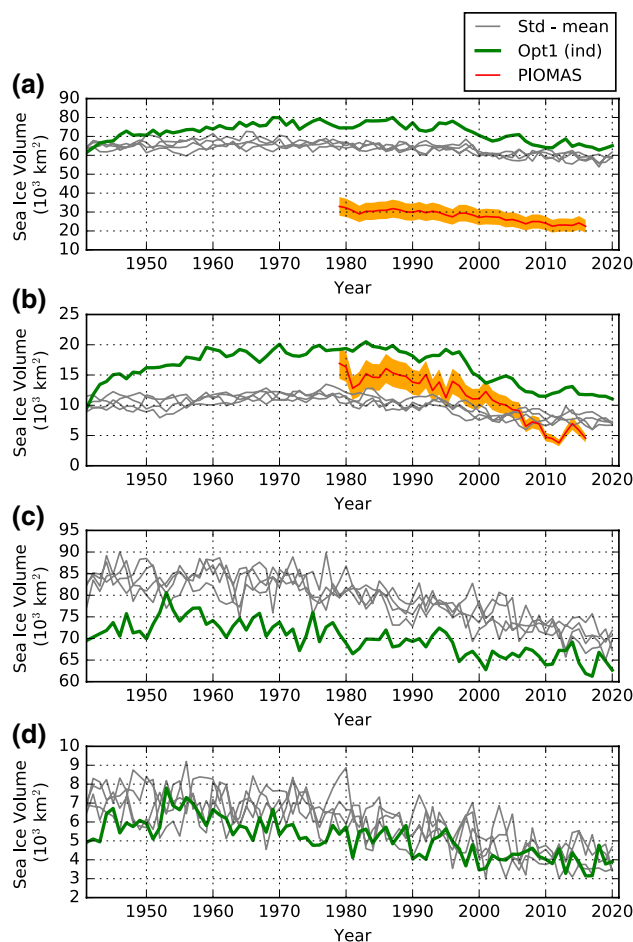
Fig. 12 Sea ice volume each year at **a** the Northern Hemisphere maximum, **b** the Northern Hemisphere minimum, **c** the Southern Hemisphere maximum and **d** the Southern Hemisphere minimum from the standard model ensemble (*thin grey solid lines*), the simulation with OPT1 parameters started in 1940 (*green solid line*) and the PIOMAS reanalysis (*red solid line*, with orange shading denoting ±15% error)

although optimization has worsened the high bias in volume in the Arctic winter, trends in Arctic sea ice volume from the optimized case show a larger decline and therefore agree better with the PIOMAS data than the standard model (Fig. 11e, f). This may be because, once thick ice is established, it is better insulated from the cold ocean beneath and will melt more easily from surface heat fluxes than thinner ice (Hodson et al. 2013), resulting in larger losses. Conversely, Antarctic ice is thinner in the optimized case (Fig. 12), resulting in smaller losses compared to the standard model (Fig. 11g, h). Lastly, we note that sea ice volume takes around two decades to adjust to parameter changes (Fig. 12). If future optimizations were to include sea ice volume as a tuning variable, they would require a longer adjustment period.

Table 1 All sea-ice-relevant parameters initially investigated in this study, with default values, extreme values and justifications for these perturbations

Parameter	Default value		Perturbation	Justification
Ocean-ice diffusion K (m^{2}s^{-1})	2.50×10^{-5a}	Low	Standard	
		High	3.75×10^{-4}	Murphy et al. (2004)
Ocean isopycnal diffusion D (m^{2}s^{-1})	1×10^3	Low	2×10^2	Brierley et al. (2010)
		High	2×10^3	Brierley et al. (2010)
Warm ice albedo α_m	0.50	Low	0.45	Perovich and Polashenski (2012)
		High	0.55	Murphy et al. (2004)
Cold ice albedo α_c	0.80	Low	0.75	Perovich and Polashenski (2012)
		High	0.85	Perovich and Polashenski (2012)
Cold ice albedo temperature T_{α_c} (°C)	-10	Low	Standard	
		High	-2	Murphy et al. (2004)
Maximum allowed SIC $c_{\max}$ NH	0.995	Low	0.960	Lindsay and Rothrock (1995)
		High	Standard	
SH	0.980	Low	0.960	[b]
		High	Standard	
Ice thermal conductivity κ_i (W m^{-1} K^{-1})	2.09	Low	1.69	Pringle et al. (2007)
		High	2.63	Rae et al. (2014)
Ice turbulent diffusion A_H (m^{2}s^{-1})	675^a	Low	0	[c]
		High	2500	Bryan (1969)

In some cases, the default value is the extreme end of the allowed range; we mark these perturbations as 'standard'

[a]Literature default values do not agree, so we select the default in UM4.5 as supplied to the University of Edinburgh

[b]Because of paucity of Antarctic observations, we assume the same limiting lead fraction as the Arctic

[c]Due to the high disagreement in the literature as to the default value, we set A_H to zero to gauge sensitivity

Table 2 (Columns 1–3): parameters, their default values and allowed value ranges used in the optimization

Parameter	STD	Range	OPT1	OPT2
Ocean-ice diffusion K (10^{-5} m^2 s^{-1})	2.5	[2.4, 37.5]	9.1	10.1
Ocean isopycnal diffusion D (10^3 m^2 s^{-1})	1	[0.75, 1.75]	0.86	0.88
Cold ice albedo temperature T_{α_c} (°C)	-10	[$-10, -2$]	-10	-10
Maximum allowed SIC $c^{\max}$ (%) NH	99.5	[96.0, 99.5]	97.0	96.5
SH	98.0	[96.0, 98.0]	97.0	96.5

(Columns 4–5): resulting parameter values from the two optimization trials

Table 3 The change in root mean square error F per parameter step from the Gauss-Newton individual parameter perturbations in each iteration of OPT1

Parameter	Change in error F per parameter step (million km^2)			
	Iteration 1	Iteration 2	Iteration 3	Iteration 4
K	-3.11×10^6	-7.56×10^5	-3.53×10^5	3.16×10^5
D	-2.48×10^5	4.10×10^5	1.00×10^5	1.15×10^5
T_{α_c}	3.15×10^5	7.03×10^5	2.16×10^5	4.78×10^5
$c_{\max}$	2.38×10^5	1.07×10^6	-4.69×10^5	-4.62×10^5

5 Discussion

This paper reports on the application of automated parameter tuning to uncertainty in climate model simulations of sea ice. We find that manipulating poorly-constrained parameter values within physically reasonable ranges can reduce model error by up to an order of magnitude in decadal mean historical sea ice extent. The improved total sea ice extent in the mean state corresponds to more consistency with observations in longer-term trends, most notably in the Arctic minimum. A key achievement is that these parameter values apply to both hemispheres and both sea ice growth and melt seasons.

Parameter tuning has been criticized for artificially improving models beyond their capabilities (Eisenman et al. 2007), but a manual approach is widely used amongst climate modellers (Mauritsen et al. 2012) and, like Massonnet et al. (2014), we argue that we can justifiably alter poorly-known or poorly-defined parameters to fit modelling needs. Using a 140-year control model experiment, we find no adverse effects of the parameter changes on the model climatology. We note an increase in Northern Hemisphere winter sea ice volume, in disagreement with reanalysis data, but this reduces later in the simulation and sea ice volume decreases faster than in the standard model after 1990.

Examining the spatial distribution of sea ice concentration, we note that, while improvements occur in the Arctic, Antarctic improvements are minimal. The inability to improve some regional aspects may be due to missing or poorly parametrized processes relevant to the Antarctic. HadCM3 does not include ice sheet melt, its coarse resolution is not suited to accurate Southern Ocean representation (Zunz et al. 2013), Antarctic topography for HadCM3 is lacking (Turner et al. 2006) and none of the model's sea ice parametrizations were specifically designed for the Antarctic (Turner et al. 2006). Indeed, our literature review showed that the experimental values behind various parameters are often measured in the Arctic only. In the face of these other issues, only limited improvement from parameter optimization using hemispheric sea ice extent error functions can be expected. We would expect to achieve greater consistency with observations using separate parameter values for each hemisphere, which would further point to deficiency in model processes, as the physics should be independent of hemisphere. We conclude that improvements from global parameter optimization are limited when other structural uncertainties dominate.

The high error reduction found here may not be replicated in other models that have better-constrained parameters or more advanced representation of sea ice processes. Yet even state-of-the-art models contain a large number of parameters and other studies have found improvements in uncoupled runs with advanced sea ice models such as CICE (Kim et al. 2006) and LIM3 (Massonnet et al. 2014). While more efficient than manual model tuning, automated parameter optimization is still computationally costly which limits ability to repeat experiments. The near-replication of resulting parameters in the second optimization experiment suggests that the successful error reduction was not due to chance, but future work should test this more robustly.

We stress that this is the first use of such a method targeting only one climate variable in a coupled climate model and as such there are areas that can be developed further. For sea ice, more advanced metrics than sea ice extent that take into account sea ice concentration differences could be investigated. Recent improvements in sea ice thickness observations mean that this could be used as an optimization metric for near present-day conditions. Similarly, more advanced models simulate variables such as meltpond area (Schroder et al. 2014), which can be observed by satellite, and therefore may provide another tuning variable. In our optimization method, a number of tuning metrics may be used without increasing computational cost. Another area for development would be to take into account observational uncertainty and weight the error function accordingly. Nevertheless, we have shown that automated parameter optimization significantly reduces difference from observations in the sea ice simulation in HadCM3. Our approach could be customized for other poorly-simulated climate observables and other models to investigate whether similar uncertainty reductions are possible.

References

Banks HT, Stark S, Keen AB (2007) The adjustment of the coupled climate Model HadGEM1 toward Equilibrium and the Impact on Global Climate. Journal of Climate 20(23):5815–5826. doi:10.1175/2007jcli1688.1. http://journals.ametsoc.org/doi/abs/10.1175/2007JCLI1688.1

Brierley CM, Collins M, Thorpe AJ (2010) The impact of perturbations to ocean-model parameters on climate and climate change in a coupled model. Clim Dyn 34(2-3):325–343. ISSN 0930-7575

Bryan K (1969) Climate and the ocean circulation: III. The ocean model. Monthly Weather Review 97(11):806–827. ISSN 1520-0493

Bunzel F, Notz D, Baehr J, Muller WA, Frohlich K (2016) Seasonal climate forecasts significantly affected by observational uncertainty of Arctic sea ice concentration. Geophys Res Lett 43(2):852–859. ISSN 1944-8007. doi:10.1002/2015GL066928

Cattle H, Crossley J, Drewry DJ (1699) Modelling Arctic climate change [and discussion]. Philos Trans R Soc London A: Math Phys Eng Sci 352:201–213, 1995. ISSN 1364-503X

Chiang JCH, Bitz CM (2005) Influence of high latitude ice cover on the marine Intertropical Convergence Zone. Clim Dyn 25(5):477–496. ISSN 1432-0894. doi:10.1007/s00382-005-0040-5

Eisenman I, Untersteiner N, Wettlaufer JS (2007) On the reliability of simulated Arctic sea ice in global climate models. Geophys Res Lett 34(10):n/a–n/a. ISSN 1944-8007. doi:10.1029/2007GL029914

Flato G, Marotzke J, Abiodun B, Braconnot P, Chou SC, Collins W, Cox P, Driouech F, Emori S, Eyring V, Forest C, Gleckler P, Guilyardi E, Jakob C, Kattsov V, Reason C, Rummukainen M (2013) Evaluation of Climate Models, book section 9, p 741866. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013. ISBN 978-1-107-66182-0. doi:10.1017/CBO9781107415324.020

Gordon C, Cooper C, Senior CA, Banks H, Gregory JM, Johns TC, Mitchell JFB, Wood RA (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. Clim Dyn 16(2-3):147–168. ISSN 0930-7575

Gregoire LJ, Valdes PJ, Payne AJ, Kahana R (2011) Optimal tuning of a GCM using modern and glacial constraints. Clim Dyn 37(3):705–719. ISSN 1432-0894. doi:10.1007/s00382-010-0934-8

Hawkins E, Rowan S (2009) The Potential to Narrow Uncertainty in Regional Climate Predictions. Bull Am Meteorol Soc 90(8):1095–1107. doi:10.1175/2009BAMS2607.1

Hibler WD (1979) A dynamic thermodynamic sea ice model. J Phys Oceanogr 9(4):815–846. ISSN 0022-3670. doi:10.1175/1520-0485(1979)009

Hodson DLR, Keeley SPE, West A, Ridley J, Hawkins E, Hewitt HT (2013) Identifying uncertainties in Arctic climate change projections. Clim Dyn 40(11):2849–2865. ISSN 1432-0894. doi:10.1007/s00382-012-1512-z

Hourdin F, Mauritsen T, Gettelman A, Golaz JC, Balaji V, Duan Q, Folini D, Ji D, Klocke D, Qian Y, Rauser F, Rio C, Tomassini L, Watanabe M, Williamson D (2016) The art and science of climate model tuning. Bull Am Meteorol Soc. doi:10.1175/bams-d-15-00135.1

Hunke EC, Lipscomb WH, Turner AK (2010) Sea-ice models for climate study: retrospective and new directions. J Glaciol 56(200):1162–1172. ISSN 0022-1430

Kim JG, Hunke EC, Lipscomb WH (2006) Sensitivity analysis and parameter tuning scheme for global sea-ice modeling. Ocean Model 14(1):61–80. ISSN 1463-5003

Lindsay RW, Rothrock DA (1995) Arctic sea ice leads from advanced very high resolution radiometer images. J Geophys Res Oceans 100(C3):4533–4544. ISSN 2156-2202. doi:10.1029/94JC02393

Marcq S, Weiss J (2012) Influence of sea ice lead-width distribution on turbulent heat transfer between the ocean and the atmosphere. The Cryosphere 6(1):143–156. ISSN 1994-0424. doi:10.5194/tc-6-143-2012. http://www.the-cryosphere.net/6/143/2012/

Massonnet F, Goosse H, Fichefet T, Counillon F (2014) Calibration of sea ice dynamic parameters in an ocean-sea ice model using an ensemble Kalman filter. J Geophys Res Oceans 119(7):4168–4184. ISSN 2169-9291

Mauritsen T, Stevens B, Roeckner E, Crueger T, Esch M, Giorgetta M, Haak H, Jungclaus J, Klocke D, Matei D, Mikolajewicz U, Notz D, Pincus R, Schmidt H, Tomassini L (2012) Tuning the climate of a global model. J Adv Model Earth Syst 4(3):n/a–n/a. ISSN 1942-2466. doi:10.1029/2012MS000154

McLaren AJ, Banks HT, Durman CF, Gregory JM, Johns TC, Keen AB, JK Ridley, Roberts MJ, Lipscomb WH, Connolley WM, Laxon SW (2006) Evaluation of the sea ice simulation in a new coupled atmosphere-ocean climate model (HadGEM1). J Geophys Res Oceans 111(C12):n/a–n/a, 2006. ISSN 2156-2202. doi:10.1029/2005JC003033

McPhee MG (1992) Turbulent heat flux in the upper ocean under sea ice. J Geophys Res Oceans 97(C4):5365–5379. ISSN 2156-2202. doi:10.1029/92JC00239

Murphy JM, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, Stainforth DA (2004) Quantification of modelling uncertainties in a large ensemble of climate change simulations. Nature 430(7001):768–772. ISSN 0028-0836

Notz D (2014) Sea-ice extent and its trend provide limited metrics of model performance. The Cryosphere 8:229–243

Notz D (2015) How well must climate models agree with observations? Philos Trans R Soc London A: Math Phys Eng Sci. 373(2052). doi:10.1098/rsta.2014.0164

Perovich DK, Polashenki C (2012) Albedo evolution of seasonal Arctic sea ice. Geophys Res Lett 39(8). ISSN 1944-8007

Pringle DJ, Eicken H, Trodahl HJ, Backstrom LGE (2007) Thermal conductivity of landfast Antarctic and Arctic sea ice. J Geophys Res Oceans 112(C4):n/a–n/a. ISSN 2156-2202. doi:10.1029/2006JC003641

Rae JGL, Hewitt HT, Keen AB, Ridley JK, Edwards JM, Harris CM (2014) A sensitivity study of the sea ice simulation in the global coupled climate model, HadGEM3. Ocean Modelling 74:60–76. ISSN 1463-5003

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J Geophys Res Atmos 108(D14):n/a–n/a. ISSN 2156-2202. doi:10.1029/2002JD002670

Ridley J, Lowe J, Brierly C, Harris G (2007) Uncertainty in the sensitivity of Arctic sea ice to global warming in a perturbed parameter climate model ensemble. Geophys Res Lett 34(19):n/a–n/a. ISSN 1944-8007. doi:10.1029/2007GL031209

Schröder D, Feltham DL, Flocco D, Tsamados M (2014) September Arctic sea-ice minimum predicted by spring melt-pond fraction. Nat Clim Change 4(5):353–357

Schurer AP, Tett SFB, Hegerl GC (2014) Small influence of solar variability on climate over the past millennium. Nat Geosci 7(2):104–108. ISSN 1752-0894

Schweiger A, Lindsay R, Zhang J, Steele M, Stern H, Kwok R (2011) Uncertainty in modeled Arctic sea ice volume. J Geophys Res Oceans 116(C8):n/a–n/a. ISSN 2156-2202. doi:10.1029/2011JC007084

Severijns CA, Hazeleger W (2005) Optimizing parameters in an atmospheric general circulation model. J Clim 18(17):3527–3535. ISSN 1520-0442

Shu Q, Song Z, Qiao F (2015) Assessment of sea ice simulations in the CMIP5 models. The Cryosphere 9(1):399–409. ISSN 1994-0416

Stainforth DA, Aina T, Christensen C, Collins M, Faull N, Frame DJ, Kettleborough JA, Knight S, Martin A, Murphy JM (2005) Uncertainty in predictions of the climate response to rising levels of greenhouse gases. Nature 433(7024):403–406. ISSN 0028-0836

Stroeve JC, Kattso V, Barrett A, Serreze M, Pavlova T, Holland M, Meier WN (2012) Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. Geophys Res Lett 39(16). ISSN 1944-8007. doi:10.1029/2012GL052676

Tett SFB, Mineter MJ, Cartis C, Rowlands DJ, Liu P (2013) Can Top-of-Atmosphere Radiation Measurements Constrain Climate Predictions? Part I: Tuning. J Clim 26(23):9348–9366. ISSN 0894-8755

Tett SFB, Yamazaki K, Mineter MJ, Cartis C, and Eizenberg N (2017) Calibrating Climate Models Using Inverse Methods: Case studies with HadAM3, HadAM3P and HadCM3, Geosci Model Dev Discuss. doi:10.5194/gmd-2016-305 (**in review**)

Turner J, Connolley WM, Marshall GJ, Lachlan-Cope TA (2006) The performance of the Hadley Centre Climate Model (HadCM3) in high southern latitudes. Int J Climatol 26(1):91–112. ISSN 1097-0088. doi:10.1002/joc.1260

Turner J, Bracegirdle TJ, Phillips T, Marshall GJ, Hosking JS (2013) An Initial Assessment of Antarctic Sea Ice Extent in the CMIP5 Models. J Clim 26(5):1473–1484. doi:10.1175/JCLI-D-12-00068.1

Voosen P (2016) Climate scientists open up their black boxes to scrutiny. Science 354(6311):401–402. doi:10.1126/science.354.6311.401

Wright MH (1996) Direct search methods: once scorned, now respectable. Addison-Wesley, pp 191–208

Zhang T, Li L, Lin Y, Xue W, Xie F, Xu H, Huang X (2015) An automatic and effective parameter optimization method for model tuning. Geosci Model Dev 8(11):3579–3591. ISSN 1991-9603. doi:10.5194/gmd-8-3579-2015

Zunz V, Goosse H, Massonnet F (2013) How does internal variability influence the ability of CMIP5 models to reproduce the recent trend in Southern Ocean sea ice extent? The Cryosphere 7(2):451–468. ISSN 1994-0416