

This discussion paper is/has been under review for the journal The Cryosphere (TC).  
Please refer to the corresponding final paper in TC if available.

# Sea-ice extent provides a limited metric of model performance

**D. Notz**

Max Planck Institute for Meteorology, Hamburg, Germany

Received: 13 May 2013 – Accepted: 7 June 2013 – Published: 27 June 2013

Correspondence to: D. Notz (dirk.notz@zmaw.de)

Published by Copernicus Publications on behalf of the European Geosciences Union.

**TCD**

7, 3095–3131, 2013

**Evaluation of  
modeled sea-ice  
concentration**

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Abstract

We examine the common practice of using sea-ice extent as the primary metric to evaluate modeled sea-ice coverage. Based on this analysis, we recommend a possible best practice for model evaluation. We find that for Arctic summer sea ice, model biases in sea-ice extent can be qualitatively different compared to biases in the geophysically more meaningful sea-ice area. These differences come about by a different frequency distribution of high-concentration sea-ice: while in summer about half of the CMIP5 models and satellite retrievals based on the Bootstrap and the ASI algorithm show a compact ice cover with large areas of high concentration sea ice, the other half of the CMIP5 models and satellite retrievals based on the NASA Team algorithm show a loose ice cover. The different behaviour of the CMIP5 models can be explained by their different distribution of excess heat between lateral melt and sea-ice thinning. Differences in grid geometry and round-off errors during interpolation only have a minor impact on the different biases in sea-ice extent and sea-ice area. Because of regional cancellation of biases in the integrative measures sea-ice extent and sea-ice area, these measures show little correlation with the more meaningful mean absolute bias in sea-ice concentration. Comparing the uncertainty arising directly from the satellite retrievals with those that arise from internal variability, we find that the latter by far dominates the uncertainty estimate for trends in sea-ice extent and area: much of the differences between modeled and observed trends can simply be explained by internal variability. Only for the absolute value of sea-ice area, differences between observations and models are so large that they cannot be explained by either observational uncertainty nor internal variability.

## 1 Introduction

Individual satellite retrievals of sea-ice concentration are often taken as the truth for model-evaluation purposes. However, significant differences exist between estimates

TCD

7, 3095–3131, 2013

## Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



of sea-ice concentration based on different satellite algorithms. These differences imply some uncertainty in our knowledge of the “true” sea-ice coverage. In this contribution, we examine in as how much this uncertainty limits the reliable assessment of the quality of simulated sea-ice coverage.

We focus primarily on the two integrative measures that are most widely used for the quantitative assessment of model quality: first, *sea-ice area*, which is just the total area of sea ice. And second, *sea-ice extent*, which is the total area of the ocean surface in which significant amounts of sea ice exist. To calculate sea-ice extent in gridded data, one usually adds the area of all grid cells with an ice concentration of more than 15%. While sea-ice extent was initially only used to assess the observed long-term evolution of the sea-ice cover (e.g. Zwally et al., 1983; Parkinson et al., 1987), it has now become common practice to use sea-ice extent also as the primary (and often sole) variable to assess the quality of modeled sea-ice coverage (e.g. Stroeve et al., 2007, 2012; Massonnet et al., 2012).

For all aspects of air–ice–sea interaction, sea-ice extent is less important than sea-ice area. Such interaction includes for example the surface albedo, the response of the ice cover to wind forcing, and the heat exchange between the ocean and the atmosphere. This was already acknowledged by early works on satellite remote sensing (c.f. Zwally et al., 1983). The focus on sea-ice extent is, nevertheless, understandable since this parameter can be more reliably observed from ships, airplanes and satellites than sea-ice area. This then allows both for a better assessment of the long-term (including pre-satellite) evolution of the ice cover and reduces the uncertainty of the observational data against which model simulations are compared.

This reduction in uncertainty in the observational data comes, however, at a price, in that sea-ice extent can give misleading results regarding model quality. Consider the trivial, fictitious observed sea-ice cover in three grid cells shown in Fig. 1a. Compared to these observations, a model could simulate a *smaller* sea-ice area that nevertheless results in a *larger* sea-ice extent because of a slight shift in the location of the sea-ice cover (Fig. 1b). A model could also simulate a *larger* sea-ice area with a *smaller* sea-

## Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



ice extent (Fig. 1c). Hence, small shifts in the location of the modeled sea-ice pack, in particular in the marginal ice zone with its strong gradients in sea-ice concentration, can result in misleading results regarding the actual bias in modeled sea-ice cover.

In addition to these grid-independent issues, there is also a grid-dependent issue related to the usage of sea-ice extent vs. sea-ice area. Generally, sea-ice extent is the smaller the higher the grid resolution is. At very high resolution, it converges to the same value as sea-ice area, since then almost all grid cells will either be fully ice covered or fully ice free.

We became aware of these issues when we analyzed results from the Max-Planck-Institute for Meteorology Earth System Model MPI-ESM: compared to observations, this model has about 6% too small a September Arctic sea-ice extent, but 20% too small a sea-ice area (Notz et al., 2013). In contrast, this model's predecessor ECHAM5/MPIOM had about 20% too large a September Arctic sea-ice extent, but only about 7% too large a sea-ice area. This gave rise to the question if too strong a focus on sea-ice extent with its smaller observational uncertainty can give misleading results regarding the quality of modeled sea-ice coverage, and which implications this has for quantitative model evaluation.

In this contribution, we examine these questions by analyzing output from models that have contributed to the Coupled Model Intercomparison Project, phase 5 (CMIP5). Our aim is to give the reader a quantitative assessment, and an explanation, for the different outcome in model–data comparison studies based on sea-ice extent vs. sea-ice area. Because positive and negative regional biases cancel in the calculation of either sea-ice extent or sea-ice area, we additionally analyse these measures' relationship to the mean absolute bias in sea-ice concentration, which avoids such cancellation of errors. We also touch upon the issue of local biases in sea-ice concentration, which are relevant for a more detailed analysis of model quality. Our aim is to allow the reader an informed assessment of which parameter to use for a specific purpose and how to handle the related observational uncertainty. In particular, we put our findings into the

Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



context of uncertainty that arises because of the internal variability of the Arctic climate system.

The satellite products and the model data that we use are introduced in Sect. 2. In Sect. 3.1, we analyse the compactness of the modelled and satellite-retrieved sea-ice cover and explain in Sect. 3.2 why about half of the CMIP5 models simulate a compact ice cover in summer time, while the other half does not. Based on these insights, in Sect. 3.3 we analyse the different biases in sea-ice extent and area, and in their trends. In Sect. 3.4 we examine the impact of grid resolution, followed by an analysis of cancelling negative and positive biases in Sect. 3.5. Section 3.6 then contains an analysis of the impact of internal variability on the assessment of model quality. In Sect. 3.7 we briefly touch upon some issues related to the non-linearity of sea-ice extent. We discuss the implications of these findings for model-evaluation purposes in Sect. 4. Our main findings are finally summarized in Sect. 5.

## 2 Models and data

For our analysis, we focus on the period 1979–2005, which is the overlapping period of the most-widely used satellite records of sea ice coverage and the “historical” simulations of the CMIP5 protocol (Taylor et al., 2012). These “historical” simulations are forced by the observed evolution of greenhouse gases, solar radiation, etc. For all 117 historical simulations that we consider here, time series of monthly sea-ice extent and area are calculated from their monthly sea-ice concentration fields. The sea-ice extent is calculated as the total area of all grid cells with at least 15 % sea-ice concentration. For sea-ice area, the area of all grid cells is multiplied by their sea-ice concentration and then added. For sea-ice area and extent, linear trends are calculated as a least-squares fit to the time series. Ensemble-mean and multi-model mean time series of sea-ice extent and sea-ice area are calculated as the ensemble-mean and the multi-model mean of the individual simulations’ time-series of these two parameters, and not from the ensemble-mean or multi-model mean concentration fields (compare Sect. 3.7).

### Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Evaluation of  
modeled sea-ice  
concentration**

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



The model results are compared against satellite retrievals of sea-ice concentration. We here discuss primarily comparisons against the two satellite algorithms for sea ice concentration that are most widely used for model-data intercomparison studies: the Bootstrap algorithm (Comiso, 1986) and the NASA Team algorithm (Cavalieri et al., 1984) that forms the basis for the NSIDC Sea-Ice Index (Fetterer et al., 2002, updated 2012). Additionally, we consider retrievals based on the ASI algorithm (Kaleschke et al., 2001; Spreen et al., 2008). The Bootstrap algorithm is probably more reliable than the NASA Team algorithm, because the latter has been found to be biased low compared to independent observations (e.g Agnew and Howell, 2003; Partington et al., 2003). The Bootstrap algorithm, in contrast, results in estimates of sea-ice concentration that are very close to the “Climate Data Record of Passive Microwave Sea Ice Concentration” (CDR, Meier et al., 2011) that is a merged product of different algorithms with the aim to provide a consistent time series of sea-ice concentration. In summer, estimates of sea-ice area of the Bootstrap algorithm also agree favourably with estimates based on the ASI algorithm from SSMI satellite data and the higher resolved AMSR-E satellite data, while estimates of sea-ice area based on the NASA team algorithm are significantly lower (Fig. 2). For these reasons, we tentatively trust that estimates of sea-ice coverage based on the Bootstrap algorithm are closer to the real sea-ice cover than those of the NASA team algorithm. Unless noted otherwise, we will therefore use the term “observations” to refer to retrievals based on the Bootstrap algorithm. For our discussion of the impact of sea-ice area vs. sea-ice extent in model–data intercomparison studies, we will take both the estimates of sea-ice area and sea-ice extent from the Bootstrap algorithm as the “truth” and will only get back to the issue of the larger uncertainty in estimates of sea-ice area in our discussion in Sect. 4. There we will also discuss in more detail the differences between the various algorithms shown in Fig. 2.







**Evaluation of  
modeled sea-ice  
concentration**

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



in sea-ice volume, thickness and concentration in the central Arctic (north of 85° N) between May and December. In this region, the two satellite algorithms that we examine here show only small seasonal changes in sea-ice concentration: excluding the satellite pole hole, the NASA Team algorithm shows in the area 85° N–87.5° N from May to August a mean reduction in sea-ice concentration of about 7 %, and from September to December an increase in sea-ice concentration of about 10 %. The Bootstrap algorithm shows changes of less than 1 % for both periods.

In the models with a generally loose ice cover, the concentration in this region decreases during the spring–summer transition much more than in the satellite record (Fig. 5a). In these models, most of the volume loss in this region is therefore governed by changes in ice concentration. In simulations with a compact ice cover, the decrease in sea-ice concentration is more similar to that of the satellite record. In these simulations, the volume loss is primarily governed by changes in ice thickness. The distribution is similar for the transition from summer to winter (Fig. 5b): from September to December, thickness increases comparably little (or even slightly decreases) in models with a loose summer ice pack, while it increases more in models with a compact summer ice pack.

These differences point to differences in the parameterisation of thinning and thickening versus lateral melting and growth as the main responsible for the emergence of two classes of models. During periods of melting, these parameterisations must capture the two main processes that can contribute to a reduction in mean sea-ice concentration: first, the total melting of thin ice within a certain grid cell, and second the lateral melt of the ice pack. In models with a sub-grid scale ice-thickness distribution, the first process is explicitly resolved and only the second one needs to be parameterised. In models with only a single ice class in a specific grid cell, both processes must be parameterised (e.g Hibler, 1979; Notz et al., 2013). In more complex parameterisations, and in all models with an ice-thickness distribution, the two processes are treated independently. The description of lateral melt is then most realistically based on a floe-size distribution within a certain grid cell (Hunke and Lipscomb, 2010; Notz, 2012).

## Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



It is likely that differences in these parameterisations lie at the heart of the emergence of the two model classes, since the more advanced models with an ice-thickness distribution usually show a compact ice cover throughout summer with comparably small seasonal changes in sea-ice concentration, while models that only simulate a single ice class primarily show a loosened ice cover in summer time and large changes in ice concentration.

However, there are exceptions to this rule: CCSM4, HadGEM2 and IPSL-CM5A do have a sub-grid scale ice thickness distribution but produce a loose sea-ice cover. In contrast, the MIROC models and CSIRO-Mk3-6-0 do not have a sub-grid scale ice-thickness distribution but produce a compact sea-ice cover. This points to the fact that not only thermodynamic factors influence the seasonal cycle of sea-ice concentration, but also dynamic factors: a divergent wind field, for example, will cause a loosening of the ice pack with no direct change in sea-ice volume. Hence, even a model with a very advanced parameterisation of thermodynamic growth and melt can result in unrealistic seasonal cycles in sea-ice concentration if the atmospheric circulation is too divergent. Analysing the contribution of such dynamical effects is beyond the scope of this study.

### 3.3 Extent versus area

In the context of the present study, one difference between a loose ice cover and a compact one is most relevant: differences between sea-ice extent and sea-ice area are comparably small for compact sea ice, because of the large number of grid cells with a very high ice concentration. In contrast, the difference between extent and area is much larger for a loose ice cover (see Fig. 6a–c).

This obviously has direct consequences for the analysis of model biases based on these two measures. For example, we find that biases relative to the Bootstrap retrieval are for simulations with a compact sea-ice cover similar for sea-ice area and for sea-ice extent (red dots are close to red line in Fig. 7a). In particular, all simulations with a compact sea-ice cover that are within  $\pm 10\%$  of the retrieved sea-ice extent are also within  $\pm 10\%$  of the retrieved sea-ice area. For the simulations with a loose ice cover, we find

that those models that underestimate sea-ice extent relative to the Bootstrap retrieval have a stronger percentage bias in sea-ice area than they have in sea-ice extent, while those simulations that overestimate sea-ice extent have a smaller percentage bias in sea-ice area than in extent (Fig. 7a). A number of simulations with a loose ice cover that fall within  $\pm 10\%$  of the retrieved sea-ice extent are clearly outside the  $\pm 10\%$  range of the retrieved sea-ice area, and vice versa. Hence, a focus on sea-ice extent can give misleading results regarding model quality compared to a focus on sea-ice area.

Relative to the satellite-retrieved estimates based on the NASA Team algorithm, we find that biases for sea-ice extent are similar to biases for sea-ice area for simulations with a loose ice cover (blue dots close to green line in Fig. 7a). Simulations with a compact ice cover that overestimate the mean sea-ice extent compared to the NASA Team algorithm in contrast have a stronger percentage bias in sea-ice area, and vice versa.

For March, all simulations and both satellite retrievals have a compact ice cover. Hence, percentage biases in sea-ice area are for all simulations almost identical to the biases in sea-ice extent (Fig. 7b).

To understand this behaviour of simulations with a compact ice cover versus those with a loose ice cover, we need to consider that the former have a small difference between sea-ice extent and sea-ice area, while the latter have a larger difference. Figure 8 illustrates how this explains the different behaviour of the two model families: if any of the loose-ice simulations with their comparably large difference between sea-ice extent and sea-ice area results in too small a mean sea-ice extent, this simulations' bias in sea-ice area will be comparably large. If, however, the simulation resulted in too large a sea-ice extent, its bias in sea-ice area would be comparably smaller – simply because the difference between extent and area is larger in the simulations than in the observations. For simulations with a compact ice cover, biases in extent and area relative to the Bootstrap algorithm are very similar, because these simulations' difference between sea-ice extent and sea-ice area is similar to that of the Bootstrap observations. Compared to observations based on the NASA Team algorithm, the simulations

**Evaluation of modeled sea-ice concentration**

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



with a compact ice cover have generally a lower difference between extent and area, which explains their contrasting behaviour relative to the NASA Team algorithm.

In winter, all simulations result in a compact sea ice cover. Therefore, in winter they have a similar difference between sea-ice extent and sea-ice area as the Bootstrap observations, which explains their consistent behaviour compared to winter-time observations.

Examining trends in sea-ice area and sea-ice extent, we find that the Bootstrap retrieval gives almost the same number for both these measures, namely an average loss of 0.56 million km<sup>2</sup> decade<sup>-1</sup> in sea-ice extent and a loss of 0.58 million km<sup>2</sup> per decade in sea-ice area during the period 1979–2005. The models, in contrast, show inconsistent behaviour, with both smaller and larger trends in sea-ice area than in extent (Fig. 6d–f). The consistent trends in the satellite retrieval can be understood by analysing the individual trends for different ice-concentration ranges (Fig. 9). Almost all the ice loss in the Bootstrap retrievals happens within the ice-concentration range 90 to 100 %, with no compensating increase in lower ice-concentration ranges (second to last panel in Fig. 9). An ice loss at these high concentrations will have roughly the same impact on sea-ice area and on sea-ice extent. For most models, in contrast, the ice loss is spread over a wider range of sea-ice concentrations. In addition, the grid cells with high ice concentration often only lose some of their ice, which then causes an increase of the number of grid cells with intermediate ice concentration. This compensation then causes a smaller loss of sea-ice extent than of sea-ice area. Some models, however, also show a faster loss in sea-ice extent than in sea-ice area. This behaviour can be understood if a significant amount of grid cells with intermediate sea-ice concentration become ice free in a simulation. The entire area of these grid cells is then lost in terms of sea-ice extent, while only the fraction of these grid cells that was ice-covered is lost from sea-ice area.

The different biases in trends of area and extent in models versus the satellite retrievals obviously again have consequences for the assessment of model quality (Fig. 10a). A number of simulations result in trends that lie within  $\pm 20\%$  of the Boot-

## Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



strap retrieved trends in sea-ice extent, while they lie outside the  $\pm 20\%$  range for the simulated trends in sea-ice area. In particular, models that have too fast a loss in sea-ice extent compared to Bootstrap retrievals sometimes have too slow a loss in sea-ice area compared to the Bootstrap retrievals. The same holds for the trends in winter sea-ice coverage (Fig. 10b). Hence, again, an assessment of model quality based on an analysis of trends in sea-ice extent can give misleading results.

### 3.4 Grid resolution

While the different histograms of sea-ice concentration explain most of the findings discussed so far, also differences in model grids might be relevant for different biases in sea-ice extent and in sea-ice area. As discussed in the introduction, one would generally expect a smaller difference between sea-ice extent and sea-ice area for higher grid resolution. The comparably high resolution of the satellite data set might therefore have contributed to the comparably small difference between sea-ice area and sea-ice extent for the Bootstrap algorithm (Fig. 6c).

To examine this possibility, we interpolated the gridded Bootstrap-derived sea-ice concentration field of September 2007 from the original 25 km EASE grid to each individual model grid. For each resulting sea-ice coverage on the original model grids, we then counted the number of ice-covered grid cells and calculated sea-ice extent and sea-ice area. We find that the calculated extent on all model grids is slightly larger than that obtained on the original EASE grid (Fig. 11a). For sea-ice area, in contrast, round-off errors during the interpolation result on some model grids in a slightly larger area than on the original EASE grid, while other grids show a slightly smaller ice-covered area (Fig. 11b). Both for extent and for area, the error that is introduced by the re-gridding is, however, so small that the difference in observational vs. model grids is largely irrelevant.

## Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



### 3.5 Canceling biases

So far, we have examined possible misinterpretations that can arise when using sea-ice extent instead of sea-ice area for model-evaluation purposes. However, a number of issues actually occur through the usage of any of these two integrated measures for model evaluation purposes. One of these issues is related to the possibility of canceling biases: a model that has a large positive bias in sea-ice concentration in one region and a large negative bias in another region might simulate a better overall sea-ice area than a model that has weak negative biases in both regions. Therefore, an analysis of the mean absolute bias in sea-ice concentration gives a better indication of model performance compared to either sea-ice extent or sea-ice area.

We calculated for the period 1979 until 2005 the area-weighted, monthly mean bias and the area-weighted, monthly mean absolute bias in sea-ice concentration in the CMIP5 simulations relative to the Bootstrap retrievals. Doing so, we find obviously very good correlation between the mean percentage bias in the integrative measures extent or area and the mean bias in sea-ice concentration (compare Fig. 12a/b versus c): For the mean bias in concentration, regional errors cancel in a similar way as they do for extent and area. Therefore, a linear regression of the biases in area versus biases in mean concentration results in a high value of  $R^2 = 0.93$ . Because of the non-linearity of sea-ice extent, the linear regression of sea-ice extent on sea-ice concentration gives a slightly lower value of  $R^2 = 0.85$ . The fact that  $R^2$  is not 1 for the linear regression of area versus mean concentration is primarily related to interpolation issues during the calculation of mean biases.

For the absolute biases in sea-ice concentration that prevent the cancellation of regional biases, however, correlation with the absolute percentage bias in the integrative measures sea-ice extent and sea-ice area is low, giving  $R^2 \approx 0.5$  for both measures: some models with almost no bias in sea-ice extent or area still have comparably large mean absolute concentration biases, and vice versa. This casts some doubt on the relevance of the integrative measures sea-ice extent and sea-ice area for model-

## Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





caused by the uncertainty in the satellite estimate. Only 10 % of CMIP5 simulations lie below the estimated reasonable range, while about 25 % lie above that range. Mean September sea-ice extent of all simulations is 7.1 million km<sup>2</sup>, a value that is between the 7.3 million km<sup>2</sup> of the Bootstrap algorithm and the 6.9 million km<sup>2</sup> for the NASA Team algorithm.

For trends of sea-ice area and sea-ice extent, finally, internal variability is by far dominating any uncertainty that arises from uncertainty in satellite retrievals (Fig. 6d, e): for the period 1979–2005, many models which generated one simulation with a sea-ice trend very close to the observed one simulate for identical forcing and slightly different initial conditions trends that are twice as strongly negative, or trends that even are positive. Hence, any trend that falls within this range might be the consequence of internal variability of the downward trend rather than a model deficiency. Using such criterion, all simulations that we consider here show a “reasonable” trend for the period 1979–2005. Since 2005, Arctic sea ice coverage in summer has decreased rapidly. The trend in September sea-ice coverage for the extended period 1979–2012, however, remains below 1 million km<sup>2</sup> ice loss per decade both for extent and area. As such, the trend remains comfortably within our estimated range of internal variability of modeled trends. Hence, also for the extended temporal range until 2012, we cannot positively identify any of the modeled trends as unreasonable.

Despite the fact that the range of crudely estimated modeled internal variability is comparable for sea-ice area and sea-ice extent, much more models fall outside that range for sea-ice area than for sea-ice extent. This, again, is obviously of relevance for any model evaluation study that focuses primarily on sea-ice extent.

### 3.7 Non-linearity

For completing our discussion of the usage of sea-ice extent for model evaluation, we should finally note that for any such comparison of modeled mean sea-ice extent with observations, the non-linearity of sea-ice extent must carefully be taken into account. Mean sea-ice extent should normally be calculated as the mean of the sea-ice extents

## Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





---

**Evaluation of  
modeled sea-ice  
concentration**D. Notz

---

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

of the individual simulations, and not as the sea-ice extent of the mean concentration of the simulations. Consider, for example, two simulations, one with 0 % ice concentration in a certain region and the other with 35 % ice concentration in that same region. The mean ice concentration of these simulations is larger than 15 %, and the sea-ice extent of the mean of the two simulations will be identical to the sea-ice extent of the simulation with the higher sea-ice concentration. The same issue arises when directly comparing sea-ice extent from daily observations with monthly mean fields of model output: the monthly-mean sea-ice extent as derived from a monthly-mean sea-ice concentration field will usually be larger than the monthly mean of daily estimates of sea-ice extent. Since sea-ice area scales linearly with ice coverage, these issues do not apply for any study using sea-ice area as a metric for model quality.

## 4 Discussion

In the previous section, we have shown that for a number of reasons the sole consideration of sea-ice extent for the evaluation of model quality can give misleading results. We therefore recommend that future studies that aim at evaluating the performance of sea ice move away from the sole consideration of sea-ice extent and also consider the model performance for the more meaningful integrative quantity sea-ice area.

In doing so, differences between different satellite algorithms will play a more prominent role than for sea-ice extent (see Fig. 2). Hence, such comparison will need to take more the form of a comparison of observational data with a specific uncertainty versus model simulations with a specific internal variability. To quantify the uncertainty of the satellite data, we compared in more detail the four satellite algorithms shown in Fig. 2. We find that despite their large difference in retrieved sea-ice area, these algorithms have a similar year-to-year variability, which becomes apparent if anomalies of all satellite algorithms relative to the retrieved area in 2010 are plotted together (see Fig. 13a, b). Hence, the difference between the satellite products is largely caused by a constant offset and there is larger certainty in anomalies in sea-ice area than there is

in its absolute value. This is important for any model simulation with assimilated sea-ice concentration fields: one should expect such model to at least retrieve the anomaly structure of the satellite time series, which can be very reliably estimated.

To quantify the uncertainty of sea-ice area retrievals and of the retrieved trends, we calculated the mean seasonal cycle of sea-ice area and of trends in sea-ice area for the period 2003–2010, for which all four satellite products contain data (see Fig. 13c, d). We then for each month simply subtracted the maximum value from the minimum value to obtain a time series of uncertainties. Doing so, we find that apart from July, differences in estimated sea-ice area are less than 1 million km<sup>2</sup> (blue curve in Fig. 13e). The same is found for an estimate of twice the standard deviation (purple curve in Fig. 13e). Hence, a value of 1 million km<sup>2</sup> can be taken as a conservative approximation of the uncertainty of retrieved sea-ice area throughout the year.

Repeating such analysis for sea-ice trends, we find that uncertainties are less than 0.4 million km<sup>2</sup> decade<sup>-1</sup> throughout the year, with smaller values in winter time (Fig. 13f). Hence, this value can be taken as a conservative approximation of the uncertainty of retrieved trends in sea-ice area. The bulk uncertainty of regional estimates of sea-ice concentration ranges from less than 5% throughout winter and spring to around 10% in summer and autumn.

A number of models have smaller biases in sea-ice area than 1 million km<sup>2</sup> relative to satellite retrievals. For these models, biases in this integrative measure could therefore simply be explained by the uncertainty range of the satellite retrievals. For the absolute biases in mean concentration, however, all models show larger biases towards satellite retrievals than the retrievals do among each other. For this most meaningful measure, the biases in the models are hence not explicable by measurement uncertainty. For a quantitative assessment of model quality, this measure is therefore currently most meaningful to quantify overall, integrative model performance in the light of uncertainties in satellite retrievals.

For a more detailed analysis of modeled sea-ice coverage, regional biases must be analysed. Therefore, the mapping of differences in modeled mean sea-ice concentra-

**Evaluation of modeled sea-ice concentration**

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





## 5 Conclusions

By analyzing the differences between sea-ice extent and the geophysically more relevant sea-ice area for model-evaluation purposes, we have found the following:

1. The summer sea-ice cover in the Arctic is seen as a compact ice cover with a high percentage of high-concentration sea ice by satellite retrievals based on the Bootstrap and the ASI algorithm. In contrast, satellite retrievals based on the NASA-Team algorithm see a loose ice cover with a low percentage of high-concentration sea ice. In winter, all satellite algorithms see a compact ice cover.
2. About half of the CMIP5 models simulate a compact ice cover in summer, while the other half simulates a loose ice cover in summer. This difference can be understood by the different distribution of excess heat for lateral melting of the ice or for thinning of the ice. In winter, all models simulate a compact ice cover.
3. Simulations with a compact ice cover have a similar bias in sea-ice extent and in sea-ice area relative to satellite retrievals based on the Bootstrap algorithm or the ASI algorithm. Simulations with a loose ice cover with a negative bias in sea-ice extent have a smaller bias in sea-ice area, while simulations with a positive bias in sea-ice extent have a larger bias in sea-ice area. This is caused by the fact that the difference between sea-ice extent and sea-ice area is larger in simulations with a loose ice cover than it is for a compact ice cover.
4. Models that simulate too fast a retreat of sea-ice extent have generally a smaller bias in simulated sea-ice-area trends relative to Bootstrap retrievals. Models that simulate too slow a retreat of sea-ice extent have generally a larger bias in sea-ice area trends. This is independent of the compactness of the ice cover.
5. The error that is introduced in the calculation of sea-ice extent and sea-ice area by different grid geometries is negligible.

### Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion





German Climate Computing Centre (DKRZ) whose data portal facilitated data access tremendously. NASA Team and Bootstrap algorithm sea-ice concentration data were obtained from the National Snow and Ice Data Center NSIDC, Boulder, Colorado, US. ASI Algorithm sea-ice concentration were obtained from the Internet on-line information page <http://icdc.zmaw.de/> maintained by the Center of Excellence for Climate System Analysis and Prediction (CliSAP), University of Hamburg, Germany. This work has been funded through a Max-Planck Research-Group Fellowship.

The service charges for this open access publication have been covered by the Max Planck Society.

## References

- Agnew, T. and Howell, S.: The use of operational ice charts for evaluating passive microwave ice concentration data, *Atmos.-Ocean*, 41, 317–331, doi:10.3137/ao.410405, 2003. 3100
- Cavalieri, D. J., Gloersen, P., and Campbell, W. J.: Determination of sea ice parameters with the NIMBUS 7 SMMR, *J. Geophys. Res.*, 89, 5355–5369, doi:10.1029/JD089iD04p05355, 1984. 3100
- Comiso, J. C.: Characteristics of Arctic winter sea ice from satellite multispectral microwave observations, *J. Geophys. Res.*, 91, 975–994, doi:10.1029/JC091iC01p00975, 1986. 3100
- Fetterer, F., Knowles, K., Meier, W., and Savoie, M.: Sea Ice Index, Digital media, National Snow and Ice Data Center, Boulder, Colorado USA, 2002, updated 2012. 3100
- Hibler, W. D.: A dynamic thermodynamic sea ice model, *J. Phys. Oceanogr.*, 9, 815–846, 1979. 3103
- Hunke, E. C. and Lipscomb, W. H.: CICE Documentation and User's Guide, 2010. 3103
- Kaleschke, L., Heygster, G., Lüpkes, C., Bochert, A., Hartmann, J., Haarpaintner, J., and Vihma, T.: SSM/I sea ice remote sensing for mesoscale ocean-atmosphere interaction analysis: ice and icebergs, *Can. J. Remote Sens.*, 27, 526–537, 2001. 3100
- Massonnet, F., Fichet, T., Goosse, H., Bitz, C. M., Philippon-Berthier, G., Holland, M. M., and Barriat, P.-Y.: Constraining projections of summer Arctic sea ice, *The Cryosphere*, 6, 1383–1394, doi:10.5194/tc-6-1383-2012, 2012. 3097

## Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



## Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Meier, W. and Notz, D.: A note on the accuracy and reliability of satellite-derived passive microwave estimates of sea-ice extent, *Clic arctic sea ice working group consensus document*, World Climate Research Program, 2010. 3101

Meier, W., Fetterer, F., Savoie, M., Mallory, S., Duerr, R., and Stroeve, J.: NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Digital media, National Snow and Ice Data Center, Boulder, Colorado USA, 2011. 3100

Notz, D.: Challenges in simulating sea ice in Earth System Models, *WIREs Clim. Change*, 3, 509–526, doi:10.1002/wcc.189, 2012. 3103

Notz, D., Haumann, A., Haak, H., Jungclaus, J., and Marotzke, J.: Arctic sea ice as simulated by MPI-ESM, *J. Adv. Model. Earth Syst.*, 5 doi:10.1002/jame.20016, 2013. 3098, 3103, 3113

Parkinson, C. L., Comiso, J. C., Zwally, H. J., Cavalieri, D. J., Gloersen, P., and Campbell, W. J.: Arctic sea ice, 1973–1976: satellite passive-microwave observations, National Aeronautics and Space Administration, Washington, D.C., 1987. 3097

Partington, K., Flynn, T., Lamb, D., Bertoia, C., and Dedrick, K.: Late twentieth century Northern Hemisphere sea-ice record from US National Ice Center ice charts, *J. Geophys. Res.*, 108, 3343, doi:10.1029/2002JC001623, 2003. 3100

Spreen, G., Kaleschke, L., and Heygster, G.: Sea ice remote sensing using AMSR-E 89-GHz channels, *J. Geophys. Res.*, 113, C02, doi:10.1029/2005JC003384, 2008. 3100

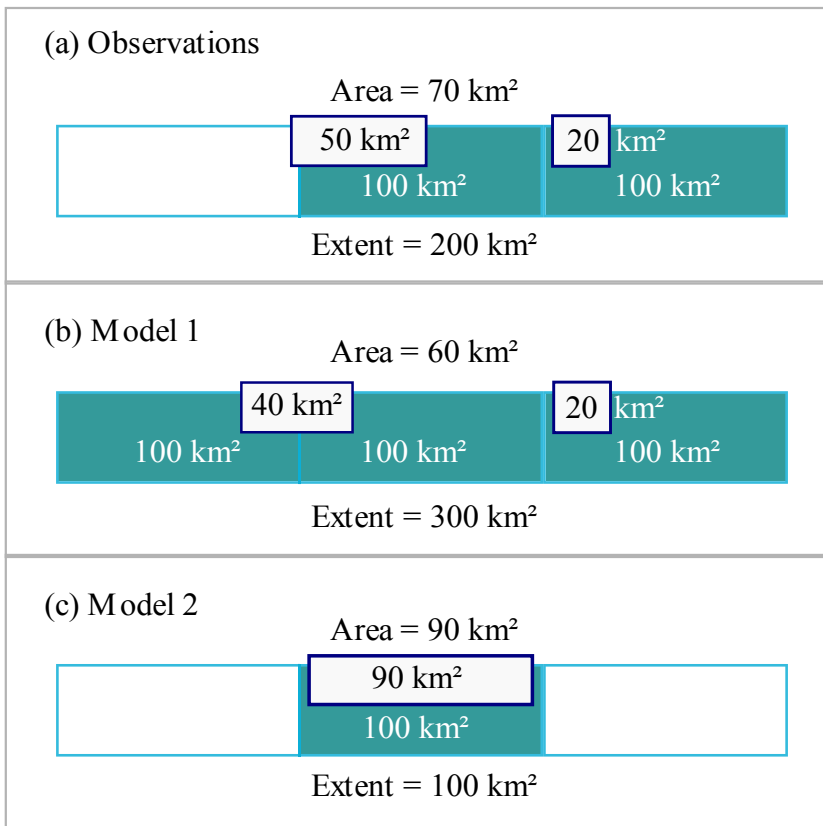
Stroeve, J., Holland, M. M., Meier, W., Scambos, T., and Serreze, M.: Arctic sea ice decline: faster than forecast, *Geophys. Res. Lett.*, 34, L09501, doi:10.1029/2007GL029703, 2007. 3097

Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M., and Meier, W. N.: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations, *Geophys. Res. Lett.*, 39, L16502, doi:10.1029/2012GL052676, 2012. 3097

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, *B. Am. Meteor. Soc.*, 93, 485–498, doi:10.1175/BAMS-D-11-00094.1, 2012. 3099

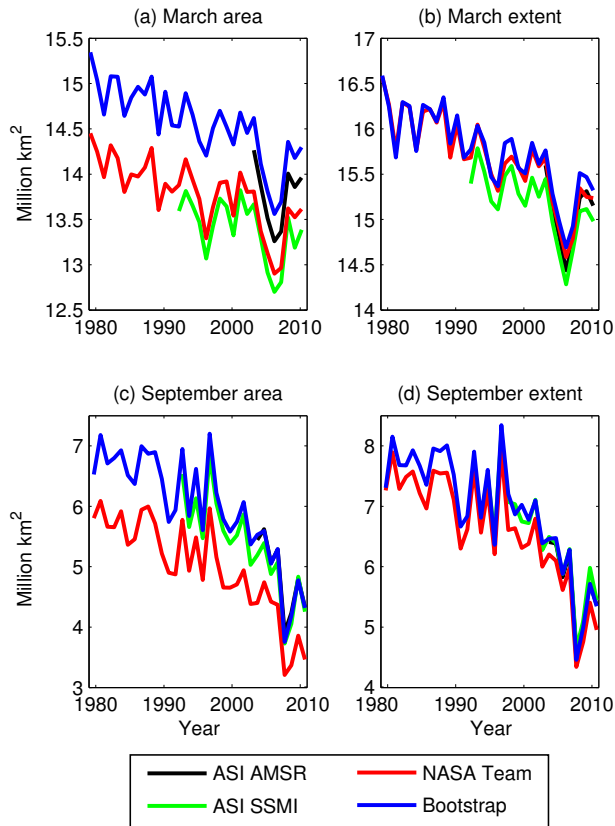
Winton, M.: Do climate models underestimate the sensitivity of Northern Hemisphere sea ice cover?, *J. Climate*, 24, 3924–3934, doi:10.1175/2011JCLI4146.1, 2011. 3109

Zwally, H. J., Parkinson, C. L., and Comiso, J. C.: Variability of Antarctic sea ice and changes in carbon dioxide, *Science*, 220, 1005–1012, doi:10.1126/science.220.4601.1005, 1983. 3097

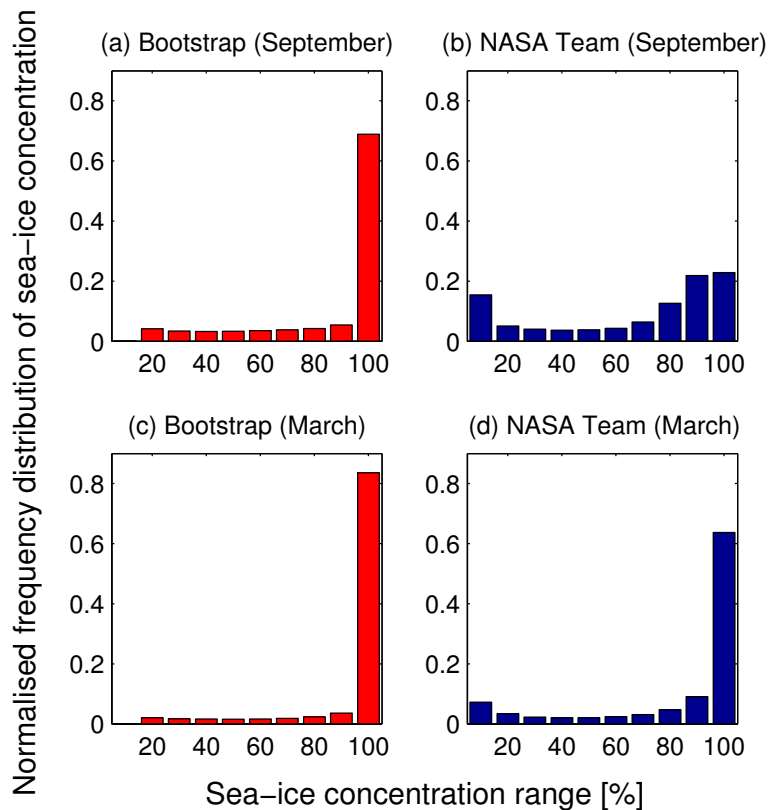


**Fig. 1.** A fictitious example to illustrate the possible non-intuitive relationship between sea-ice area and sea-ice extent **(a)** in the observations, the ice pack is distributed such that two grid cells are covered by more than 15 % ice. **(b)** In a fictitious model simulation, less sea ice than in the observations is distributed such that three grid cells are covered by more than 15 % ice. **(c)** In a fictitious model simulation, more sea ice than in the observations is distributed such that only one grid cell is covered by more than 15 % ice.





**Fig. 2.** September and March sea-ice area and sea-ice extent as retrieved from satellite for the period 1979–2010. Different colors denote different algorithms or satellites. Area and extent were calculated based on sea-ice concentration fields on EASE grids with 25 km resolution (NASA Team and Bootstrap, 1979–2010), 12 km resolution (ASI SSM/I, 1992–2010) and 6 km resolution (ASI AMSR-E, 2002–2010).



**Fig. 3.** Histogram of 1979–2005 sea-ice concentration in all areas with at least 0.1% sea-ice concentration. **(a, c)** Satellite retrievals based on the Bootstrap algorithm and **(b, d)** satellite retrievals based on the NASA-Team algorithm for **(a, b)** September and **(c, d)** March. The numbers on the x-axis denote the upper limit of each bar: 20 e.g. denotes the concentration range 10 to 20%.

Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

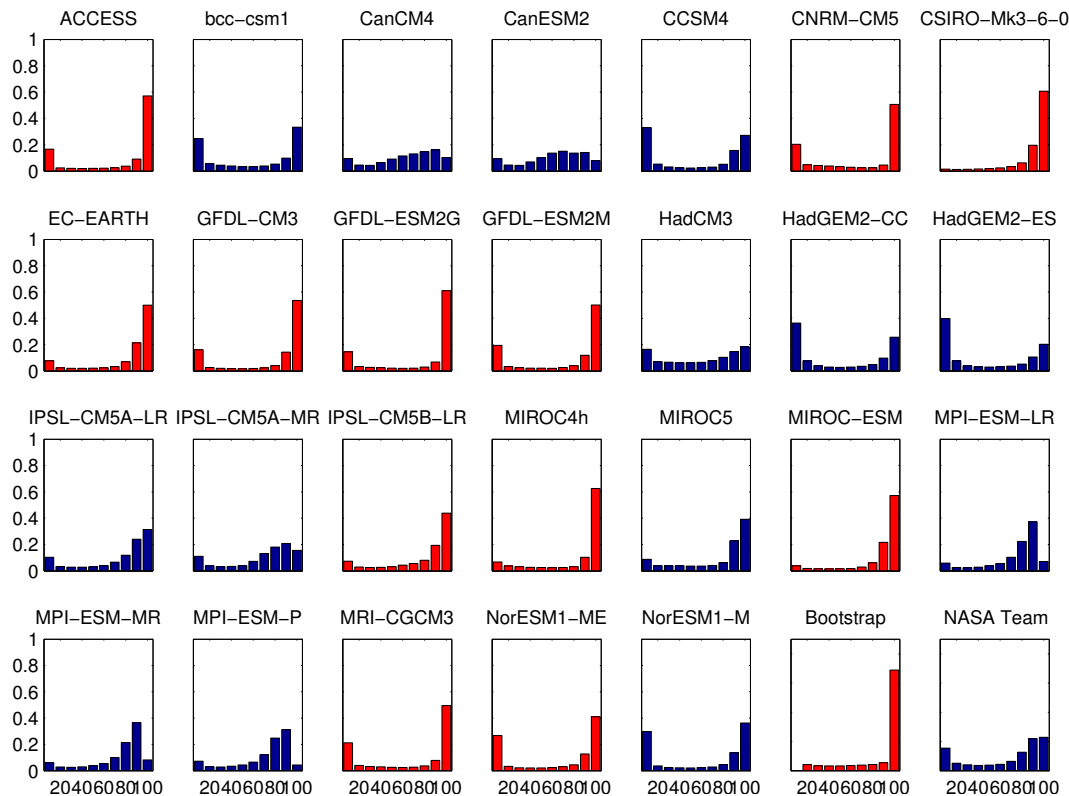
Printer-friendly Version

Interactive Discussion



## Evaluation of modeled sea-ice concentration

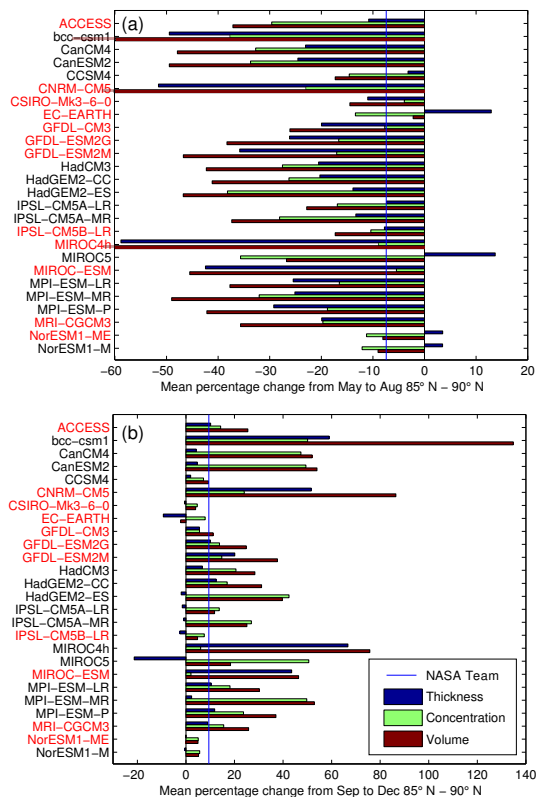
D. Notz



**Fig. 4.** Histogram of 1979–2005 September sea-ice concentration in all grid cells with at least 0.1 % sea-ice concentration in CMIP5 model simulations. The numbers on the x-axis denote the upper limit of each bar: 20 e.g. denotes the concentration range 10 to 20 %. Red panels denote histograms with a compact sea-ice cover, while blue panels denote histograms with a loose sea-ice cover. For models with multiple simulations, the ensemble mean is shown.

Evaluation of modeled sea-ice concentration

D. Notz



**Fig. 5.** Change in thickness, concentration and volume in model simulations for the period (a) May to September and (b) September to December in the central Arctic north of 85° N during the period 1979–2005. The vertical blue line denotes the concentration change as retrieved by the NASA Team algorithm. The concentration change of the Bootstrap algorithm is almost zero for both periods. For models with multiple simulations, the ensemble mean is shown. Models with a compact ice cover are labeled in red.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

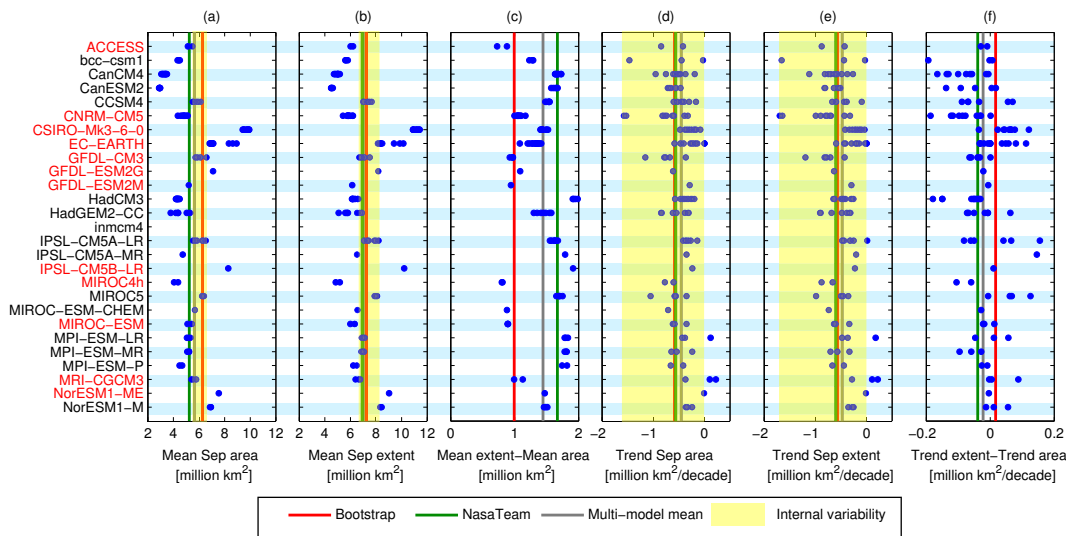
Printer-friendly Version

Interactive Discussion



Evaluation of modeled sea-ice concentration

D. Notz



**Fig. 6.** Overview of the September sea-ice coverage in the 117 “historical” CMIP5 simulations analyzed for this study. Each individual dot corresponds to a single simulation. The vertical lines show the values of the observational record and the mean of all simulations. The yellow shading indicates estimated internal variability. All data refer to the period 1979–2005. **(a)** Mean September sea-ice area, **(b)** mean September sea-ice extent, **(c)** difference between mean September sea-ice extent and mean September sea-ice area, **(d)** linear trend in September sea-ice area, **(e)** linear trend in September sea-ice extent, **(f)** difference between trend in September sea-ice extent and trend in September sea-ice area. Models with a compact ice cover are labeled in red.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

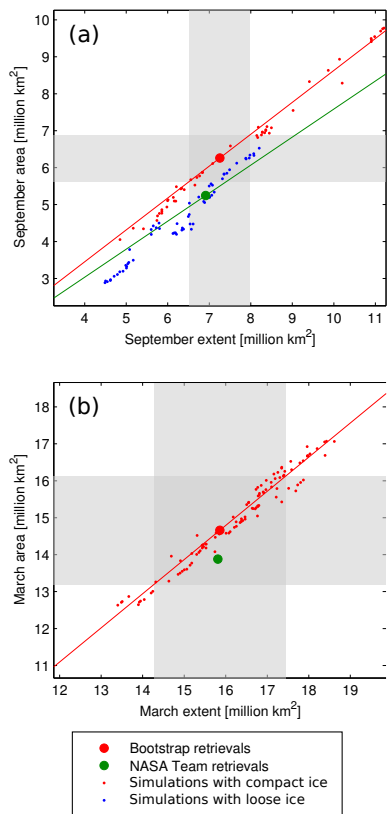
Printer-friendly Version

Interactive Discussion



## Evaluation of modeled sea-ice concentration

D. Notz

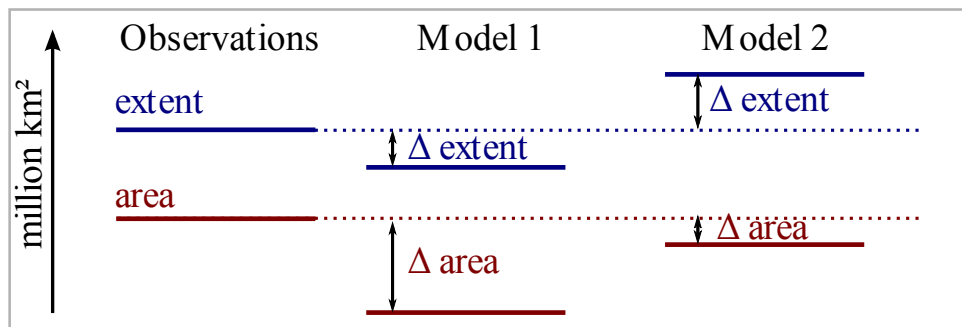


**Fig. 7.** (a) September and (b) March sea-ice area versus sea-ice extent in models and satellite retrievals. The red line connects all value pairs that have the same percentage bias in sea-ice extent and in sea-ice area relative to the Bootstrap retrievals. The gray shading indicates a  $\pm 10\%$  range around the values obtained from the Bootstrap retrievals.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)

## Evaluation of modeled sea-ice concentration

D. Notz



**Fig. 8.** Schematic to explain the findings in Fig. 7: because of the smaller difference between sea-ice extent and sea-ice area in observations with a compact ice cover than in simulations with a loose ice cover, models with a loose ice cover and slightly too large a simulated sea-ice extent result in a comparably small bias in simulated sea-ice area.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

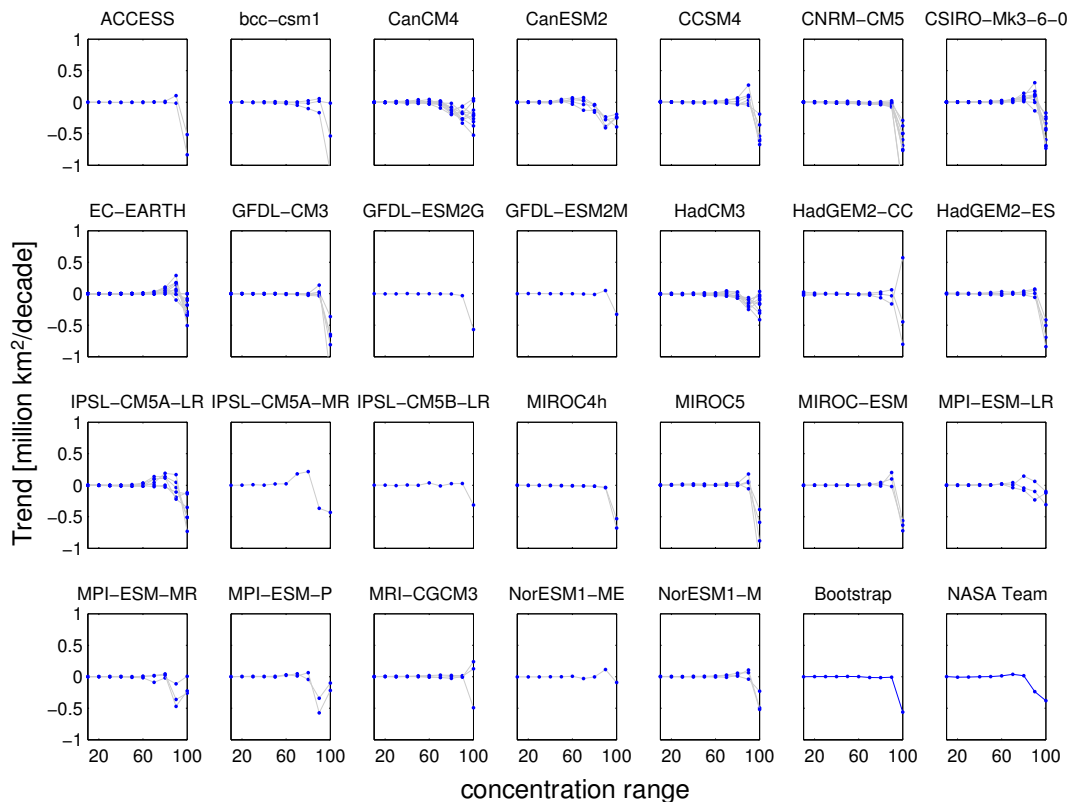
Printer-friendly Version

Interactive Discussion



## Evaluation of modeled sea-ice concentration

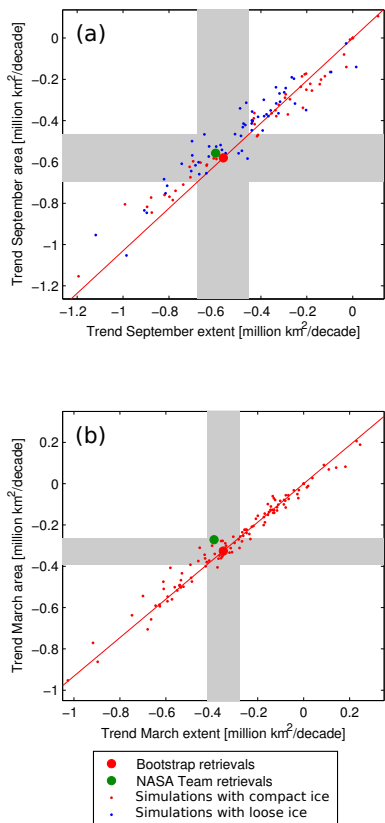
D. Notz



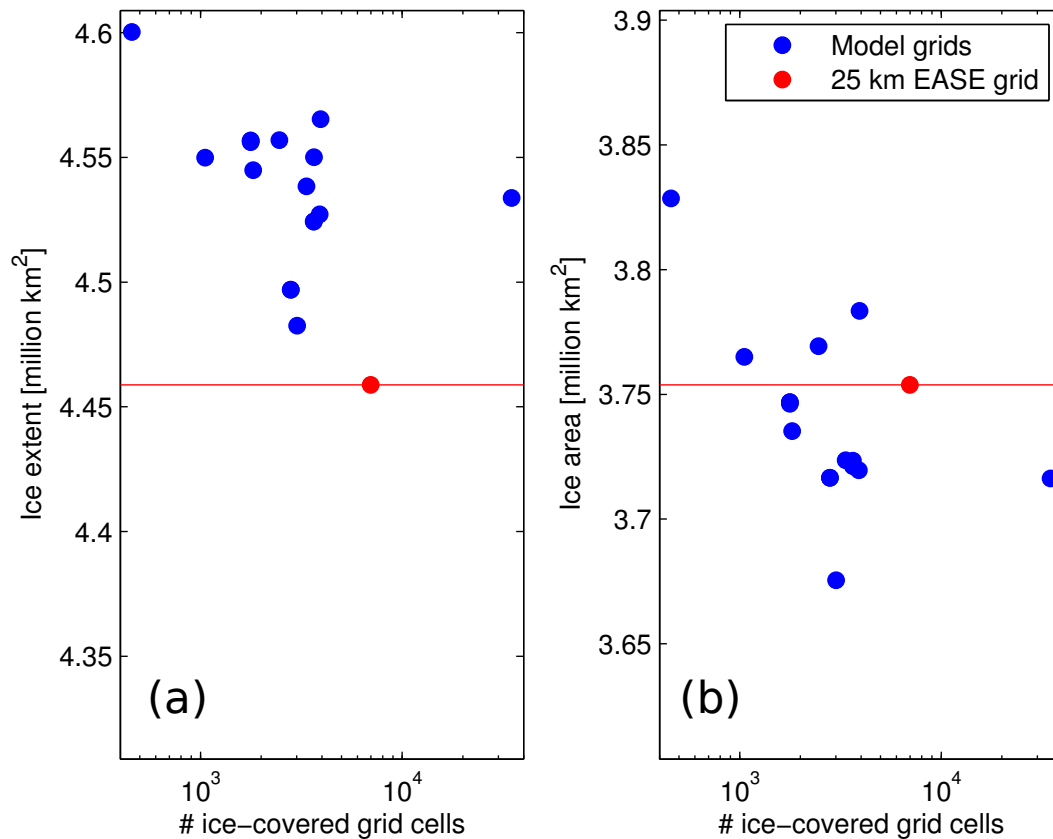
**Fig. 9.** Trends in September sea-ice area per ice-concentration category. The numbers on the x-axis denote the upper limit of each bar: 20 e.g. denotes the concentration range 10 to 20.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[◀](#)
[▶](#)
[◀](#)
[▶](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)



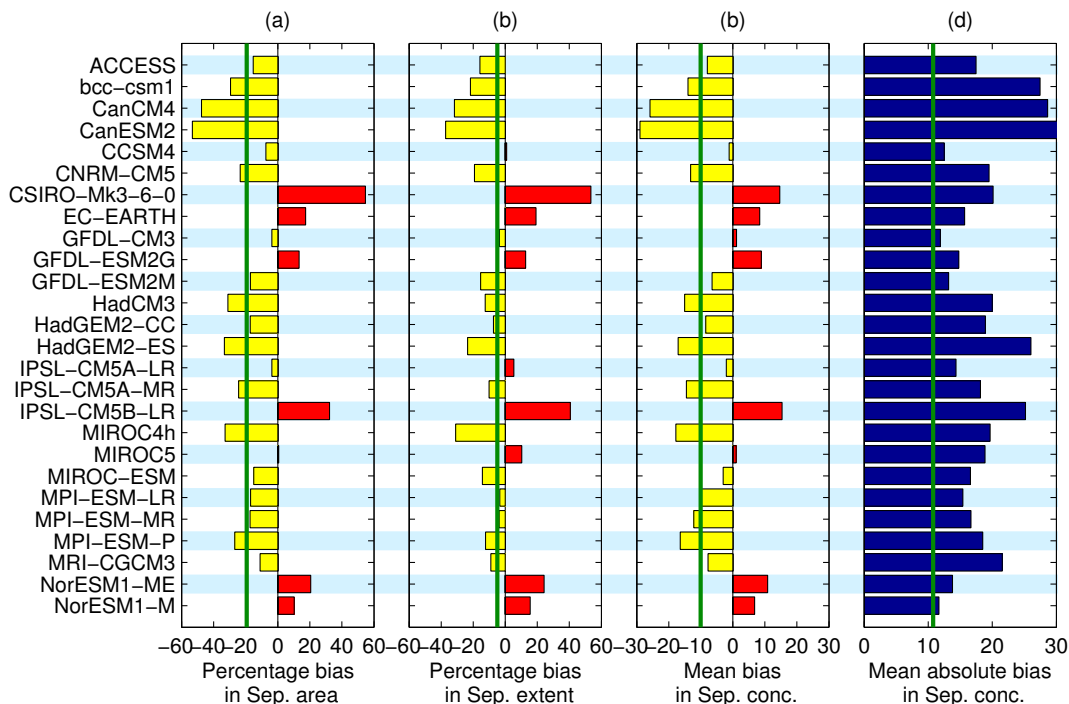
**Fig. 10. (a)** September and **(b)** March trends in sea-ice area versus trends in sea-ice extent for the period 1979–2005. The red line connects all value pairs that have the same percentage bias in sea-ice extent and in sea-ice area relative to the Bootstrap retrievals. The gray shading indicates a  $\pm 20\%$  range around the trends obtained from the Bootstrap retrievals.



**Fig. 11.** (a) Sea-ice extent and (b) sea-ice area versus number of ice-covered grid cells for observational data of September 2007 interpolated onto all model grids examined for this study.

## Evaluation of modeled sea-ice concentration

D. Notz



**Fig. 12.** Overview of biases in September sea-ice coverage in the 117 “historical” CMIP5 simulations analyzed for this study relative to Bootstrap satellite retrievals. **(a)** Mean bias in September sea-ice area. **(b)** Mean bias in September sea-ice extent. **(c)** Mean bias in sea-ice concentration **(d)** mean absolute bias in sea-ice concentration. The vertical green lines denote the respective bias of the NASA Team retrieval relative to the Bootstrap retrieval.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

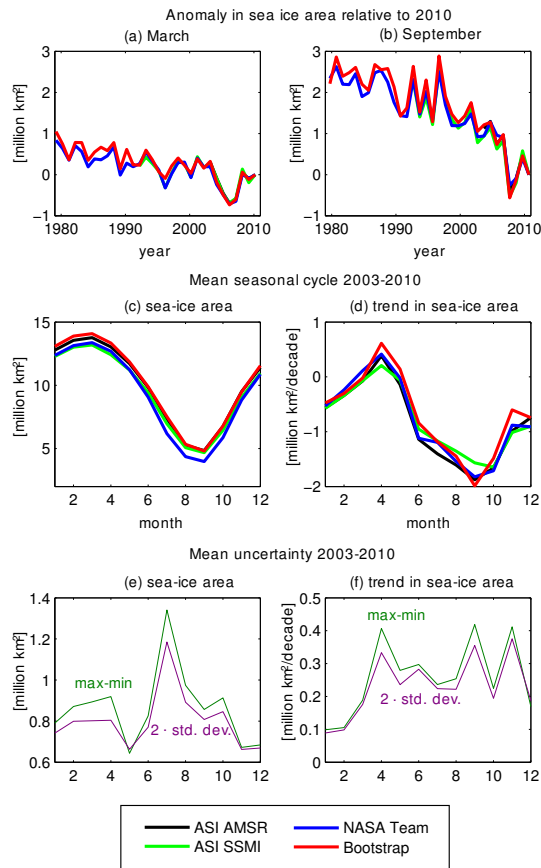
Printer-friendly Version

Interactive Discussion



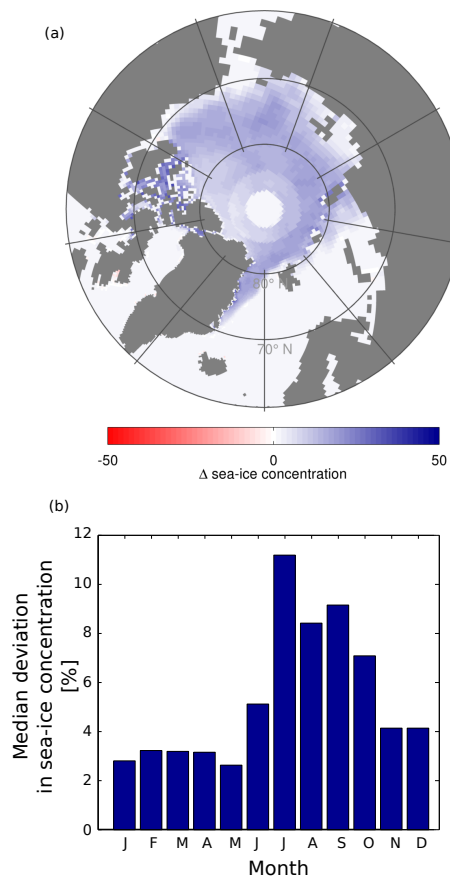
## Evaluation of modeled sea-ice concentration

D. Notz



**Fig. 13.** (a) March and (b) September anomaly in sea-ice area as retrieved from satellites for the period 1979–2010. Different colors denote different algorithms or satellites. (c) Seasonal cycle in sea-ice area and (d) in sea-ice-area trend as retrieved from satellites for the period 2003–2010. (e) Uncertainty in retrieved sea-ice area and (f) in retrieved trend of sea-ice area.

[Title Page](#)
[Abstract](#)
[Introduction](#)
[Conclusions](#)
[References](#)
[Tables](#)
[Figures](#)
[Back](#)
[Close](#)
[Full Screen / Esc](#)
[Printer-friendly Version](#)
[Interactive Discussion](#)



**Fig. 14.** (a) Mean difference in September sea-ice concentration between Bootstrap retrieval and NASA Team retrieval for the period 1979–2005. (b) Monthly median deviation in sea-ice concentration between Bootstrap retrieval and NASA Team retrieval for the period 1979–2005.

Evaluation of modeled sea-ice concentration

D. Notz

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

