Master Thesis: Assessment of Model Performance in Simulating Arctic Sea Ice Using Taylor Diagrams

Rusul Al-Janabi

9th August 2012

Contents

1 Introduction

Numerical modeling has witnessed great leaps in the last decades. Climate models are tools for understanding past climate change and predicting future possible changes. Thanks to advances in computational capacity, researchers are able now to develop more complex models, by including more processes and more components. Finer resolution and smaller time steps are also now available, leading to more realistic simulations.

Has our global climate system changed? For the last 20000 years, until the start of the industrial revolution CO_2 concentrations were relatively stable. Then they were increased by almost 27 % which led to a rise in the global average temperature by almost 1°. The climate research community is interested in predicting to what extent the global climate system would be affected by this change, and in assisting countries and communities in adapting to these changes.

One of the keys in understanding future changes lies in understanding the dynamics of the polar regions. The significant impact of the polar regions on the global climate system is due largely to the physical properties of sea ice. Sea ice covers 7% of the oceans in the annual mean. If sea ice cover is changed, the surface albedo changes accordingly, which directly affects the terrestrial radiation budget. If sea ice melting rates increase, a substantial amount of fresh water will be introduced to the ocean, which in turn will effect mixed layer processes, deep water formation and the meridional overturning circulation.

According to the IPCC (Intergovernmental Panel on future Climate Change) fourth assessment report, there is already significant and abundant evidence that most of the cryospheric components and especially sea ice in polar regions are undergoing generalized shrinkage, and that there effect in the environmental and in human activities are already detectable*[Lemke et al., 2007]*.

According to satellite and ship observations, sea ice had its lowest record in September 2007. Some models have suggested that sea ice may disappear in summer by the end of the 21st century. The complex dynamics of the polar regions and the role of sea ice have yet to be fully understood in order to make better predictions about the future.

The polar regions may be one of the best examples to illustrate the sensitivity of our climate system to CO_2 increases (Arctic amplification).

Sea ice characteristics differ between the Arctic and the Antarctic. That is due to the geographical differences between both regions. The Arctic is a basin surrounded by land, whereas the Antarctic is a continent surrounded by water. In addition the Arctic is considerably fresher than the Antarctic due to river input from the continents. This leads to

the Arctic ocean being more stable than the Southern Ocean which has implications for deep water formation, MOC, and mixed layer processes

The main characteristics that have direct influence over sea ice-ocean interaction are: brine concentrations, thickness, depth of the snow layer covering the ice and percentages of multi-year and first-year ice.

Oceanic heat fluxes vary considerably between the two regions. This effects the maximum thickness the ice may reach. Oceanic heat fluxes vary around $4Wm^2$ in the Arctic, and it may mount up to about $40Wm^2$ in the Antarctic, causing the Arctic mean thickness to be one meter larger than the Antarctic.*[Thomas and Dieckmann, 2010]*

The amount of snow covering the ice is very important and has to be taken into consideration even though it accounts for less than 10% of sea ice mass. Snow plays a major role in the heat budget of sea ice because of it's higher albedo and it's low thermal conductivity. The thermal conductivity of snow is lower than that of ice by 1 order of magnitude, acting as a insulator between sea ice and the atmosphere. Precipitation rates in the Antarctic are higher than the Arctic, because of the location of the continent, thus, snow layers covering sea ice are deeper in the Antarctic. *[Thomas and Dieckmann, 2010].* In terms of ice age, more than half of the sea ice in the Arctic is multi-year ice. In the Antarctic only a small fraction is multi-year ice.

Formation of sea ice mainly occurs in the Laptev Sea in the Arctic. A Polynya is one location of sea ice production in the southern ocean. It is open water exposed to katabatic cold winds from the Antarctic continent, which travel to the coast from high elevations.

Simulating polar sea ice and it's seasonal cycle is by no means a simple task. It is sensitive to the choice of ocean boundary conditions and model formulation *[Proshutinsky et al., 2001].* Errors in modeling sea ice may lead to substantial errors in the model's climate as a whole. The consequences involve the global radiation balance, the meridional overturning circulation and heat and salt fluxes. Errors in simulating sea ice coverage directly effects the planetary albedo, which in turn affects the radiation balance equation. Errors involving fresh water fluxes affects ocean statification, which lead to incorrect simulation of the MOC. Errors in modeling melting/ freezing rates of sea ice affects heat and salt intake, which effects the globally averaged ocean temperature and salinity.

Model development requires constant validation of model performance. One of the methods of validation is visual comparison between the simulated and observed fields. For models with a multitude of variables and multiple dimensions, visual comparison of the simulated and the observed fields becomes impractical *[Taylor, 2001].* "Metrics" have been developed

in recent years mainly by P. J. Gleckler as a way to rank different models. It has not been widely applied to sea ice-ocean models *[Gleckler and Taylor, 2008]*. Using metrics, it has been found that the relative ranking of models varies considerably from one variable to the next. In addition accurate simulation of one aspect of the climate does not guarantee accurate representation of other aspects *[Gleckler and Taylor, 2008]*. Hence, we must use a practical method of validation that can overcome such problems. The so-called Taylor diagram has been devised, by Karl E. Taylor to provide an objective overview or summary of model performance. It provides evidence of fidelity of climate models, in which plots can show that some pattern of observed variation is reasonably well reproduced by the model *[Taylor, 2001]*.

The aim of this research is to compare different models and model versions easily using the Taylor diagram. This research attempts to answer the following question: to which extent does model behavior resemble the observations, and which models produced better results?

The climate dynamics group in the Alfered Wegener Institute (AWI) in Bremerhaven, is interested in further developing FESOM, the Finite Element Sea ice-Ocean Model, in order to achieve better future predictions of climate change. This research will compare results from FESOM to MITgcm, the Massachusetts Institute of Technology global circulation model, and to observational data, for sea ice variables.

2 Methods

2.1 Numerical models

Numerical models are very powerful tools that help us understand geophysical phenomenas. Models can simulate almost the entire ocean, and can give insight to some areas where measurements are difficult to obtain . Ships and satellites are always limited by the conditions on the ground. There are many restrictions to where a ship may go and what time measurements can be taken. For example Ships can not go into a storm. In addition, even if the conditions on ground are favorable, the measurements are still far apart. Satellites on the other hand cover almost the entire globe but the measurements are limited to the surface and do not penetrate more than a few meters. Weather stations are far apart.
Numerical models apply numerical methods to find solutions of the governing differential equations. The finite difference method, the finite element method and the finite volume method are some to mention. For ocean models these methods are employed to solve the primitive equations of conservation of mass and momentum. The momentum equation can be stated as:

$$\partial_t \vec{u} + \vec{v}.\nabla_3 \vec{u} + f\vec{k} \times \vec{u} + \frac{1}{\rho_0}\nabla p + g\nabla\eta = \nabla.A_h\nabla\vec{u} + \partial_z A_v \partial_z \vec{u} \tag{1}$$

and the vertically integrated continuity equation as:

$$\partial_t \eta + \nabla. \int_{z=-H}^{z=\eta} \vec{u} dz = 0 \tag{2}$$

and the hydrostatic pressure equation as:

$$\partial_z p = -g\rho \tag{3}$$

Where, $\vec{v}$ is the velocity vector in the three dimensional spherical coordinates, $\vec{u}$ is is the horizontal component of the velocity vector, ∇_3 is the three dimensional gradient operator, ∇ is the two dimensional gradient operator, f is the coriolis parameter, $\vec{k}$ is the vertical unit vector, ρ and ρ_0 are density and it's mean value, p is the hydrostatic pressure, η is the sea surface elevation, A_h is the horizontal viscosity, A_v is the vertical viscosity, H is the ocean depth, and g is the gravitational acceleration.

The set of equations stated earlier, can not be solved analytically because of nonlinearity and terbulance terms. In order to apply numerical methods for obtaining a solution, boundary conditions have to be set. The boundary conditions represent the physical domain. In case of the ocean it is set by the ocean surface and bottom. Obtaining a realistic shape for the ocean floor and coastlines is challenging, and adds to the complexity of the solution *[Stewart, 2008]*.
The continuous primitive equations are then spatially and temporally discretized. The spatial resolution of sea ice-ocean models are tens to hundreds of kilometers in the horizontal and tens to hundreds of meters apart in the vertical. With such a high resolution, processes such as turbulence is not represented because turbulence occurs in the scale of a few millimeters. Numerical codes have many sources of error but most are small in practice *Stewart [2008]*
Sea ice motion is derived by ice-atmosphere, ice-ocean and internal stresses, and by the horizontal surface elevation gradient of the ocean *Losch et al. [2010]*. The basic equations that derive sea ice motion are the continuity equation and the velocity equation. The continuity equation states that:

$$\frac{\partial h}{\partial t} + \nabla.(h\vec{u}) = Q_h \tag{4}$$

where h is the mean sea ice thickness, Q_h are the thermodynamic sources and sinks and $\vec{u}$ is sea ice velocity in three dimensions.

The velocity equation is written as follows:

$$m(\frac{\partial}{\partial t} + f\vec{k}\times)\vec{u} = A(\vec{\tau} - c_d\rho(\vec{u} - \vec{u}_w)|\vec{u} - \vec{u}_w|) + \vec{F} - mg\nabla\eta. \tag{5}$$

where m is the mass of ice and snow, τ is the wind stress, c_d is the ocean drag coefficient, ρ is the density of sea water, $\vec{u}_w$ is the ocean surface velocity and η is the sea surface elevation.

2.2 FESOM

The Finite-Element Sea ice-Ocean model (FESOM), has been developed by the Climate dynamics group at the Alfered Wagener Institute (AWI) in Bremerhaven. It is characterized by an unstructured triangular surface grid, that forms tetrahedras in 3-D. Vertical discretization for this model version is on z-levels with grid nodes aligned underneath each surface grid node (see figure 1). Each surface triangle defines a prism that is cut

into smaller prisms by z-levels. Each small prism is then split into three tetrahedras *[Timmermann et al., 2008]*. The resolution ranges from 20 km in polar regions and coastlines, to about 150 km in mid latitudes figure 2. For the results obtained in this research a mesh with a resolution of 25 km is used. FESOM is a hydrostatic ocean circulation model which solves the primitive equations in a spherical geometry. The finite element method is utilized to discretize the equations. It uses a continuous linear representation of temperature, salinity, horizontal velocities, and surface elevation *[Timmermann et al., 2008]*. FESOM has an integration time from 1958-2007.

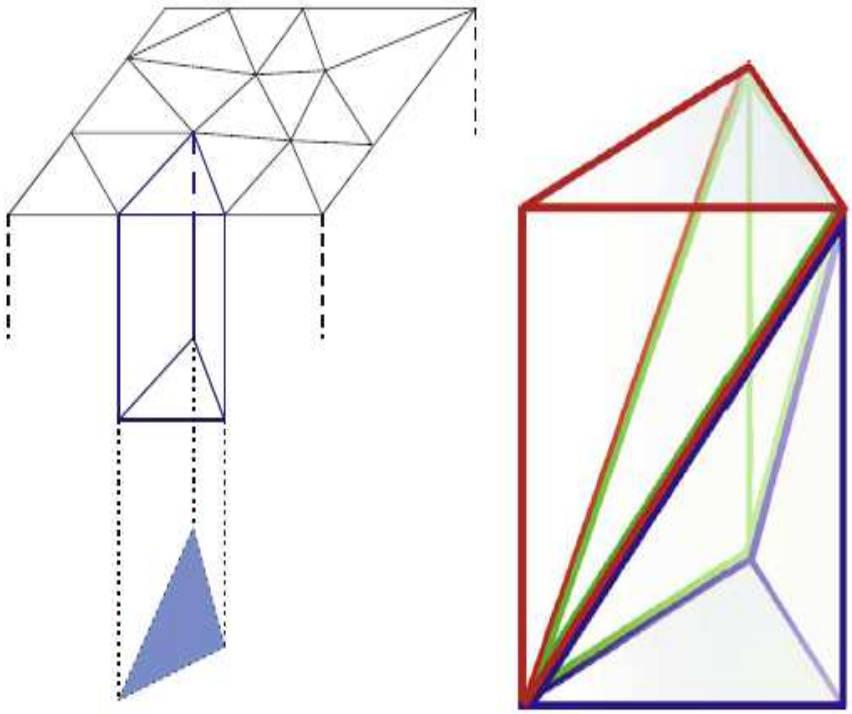

Figure 1: Illustration of the three dimensional structure of the FESOM mesh. Grid nodes in the interior ocean are aligned under the surface triangles (left). Each of the resulting prisms are split into three tetrahedra (right). *[Timmermann et al., 2008]*

Initial temperature and salinity have been derived from the World Ocean Atlas (WOA01) *[Conkright et al., 2002]*. Atmospheric forcing fields for the simulations presented here have been obtained from the NCEP/NCAR atmospheric reanalysis data set *[Large and Yeager, 2008]* for the period 1948-2007. They consist of datasets for 10 m-wind 2 m-temperature,

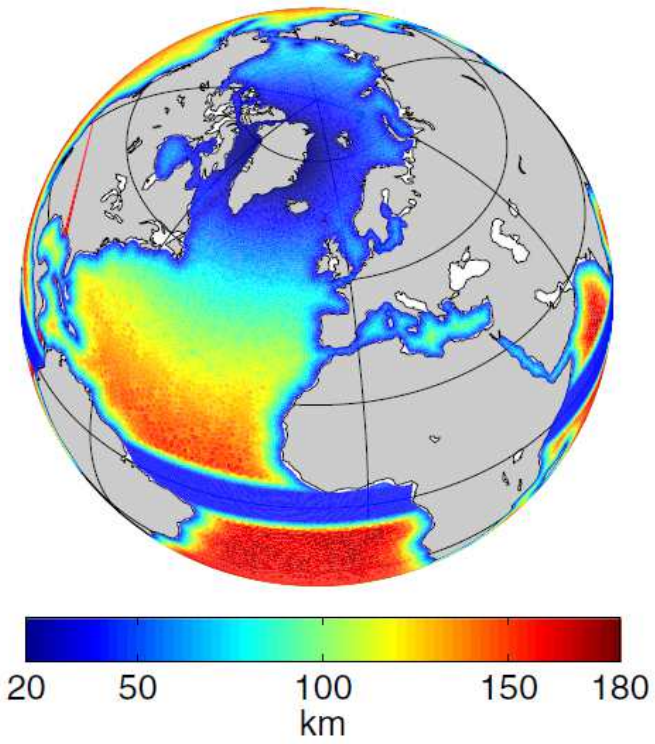

Figure 2: FESOM mesh resolution. About 20 km in the arctic and equator, up to 150 km other areas *[Sidorenko et al., 2011]*

specific humidity, total cloudiness and net precipitation, where the later has been derived from total precipitation and the reanalysis latent heat flux.
In regions with an initial sea surface temperature below -1 °C, an initial sea ice thickness of 1 m, snow thickness of 0.1 m, and ice concentration of 0.9 were prescribed, which yields a sea ice distribution that is close to observed ice concentrations for January *[Timmermann et al., 2008]*.

2.3 MITgcm

The configuration of the MITgcm (Massachusetts Institute of Technology global circulation model) as used in this research is a hydrostatic, finite volume sea ice-ocean model[1]. The model grid is a cube-sphere grid which avoids polar singularities. Each face of the cube has 510*510 structured grid cells. The mean horizontal grid spacing 18 km. The

[1]Data has been provided by Dimitris Menemenlis. For more information visit the website: www.ecco2.jpl.nasa.gov

vertical is composed of 50 z-levels ranging between 10-450m, according to the depth of the ocean.The current integration covers the years from 1992 to 2007. The forcing used is the Japanese 25-year ReAnalysis (JRA-25), which is a 26-year reanalysis project, conducted by the Japan Meteorological Agency (JMA) and the Central Research Institute of Electric Power Industry (CRIEPI).
The sea ice component of MITgcm is based on a variant of the viscous-plastic (VP) dynamic-thermodynamic sea-ice model of Hibler *[Hibler, 1979]*. The model has been written for an Arakawa C grid. Sea ice momentum equations are solved either using line-successive-over-relaxation (LSOR). The finite-volume discretization of the momentum equation on the Arakawa C grid is straightforward *[Losch et al., 2010]*.
Surface fluxes of fresh water and heat are computed from the atmospheric state and bulk formula.

2.4 FESOM/ECHAM

The coupled FESOM/ECHAM model is the coupled ocean-sea ice and atmosphere model developed at the climate dynamic group at AWI. The atmospheric model (ECHAM) was coupled to the ocean-sea ice component (FESOM) using a parallel coupler OASIS4 (Ocean Atmosphere Sea Ice Soil version 4) (see figure 3). The approach followed in climate modeling is to develop each component of the climate system independently. Later all these components are coupled together to build a model that describes the whole climate system.
ECHAM is an atmospheric general circulation model. It has been developed at Max Plank Institute for Meteorology (MPIM) in Hamburg. It evolved from the spectral weather prediction model of the European Centre for Medium Range Weather Forecasts (ECMWF). The first two letters EC are from ECMWF, the other three HAM is from HAMburg *[Roeckner et al., 2003]*.
In the coupled context ECHAM is currently used with the T63L31 grid (ca. 1.87 degrees or 200km in the horizontal). CO_2 and aerosol concentrations were set as from the year 2000. The different models have usually different time steps *[Redler et al., 2010]*. FESOM's time step is 30 min ECHAM's is 12 min. This is why the OASIS4 coupler is used to carry out accumulation in time and interpolation between the two models. Coupling step is carried out every 6 hours. The World Ocean Atlas (WOA2005) *[Conkright et al., 2002]* climatologies were used for initialization of temperature and salinity in FESOM.
The overall coupling technique adopts the following strategy: 1.ECHAM receives sea surface fields from FESOM, which include surface temperature, sea ice concentration, sea ice

and snow thickness. 2.ECHAM uses these fields to compute moisture, heat and momentum fluxes and passes these fluxes to FESOM. 3.FESOM uses these fluxes to drive the ocean.

Alltogether 4 fields and 12 fluxes are exchanged between FESOM and ECHAM models.

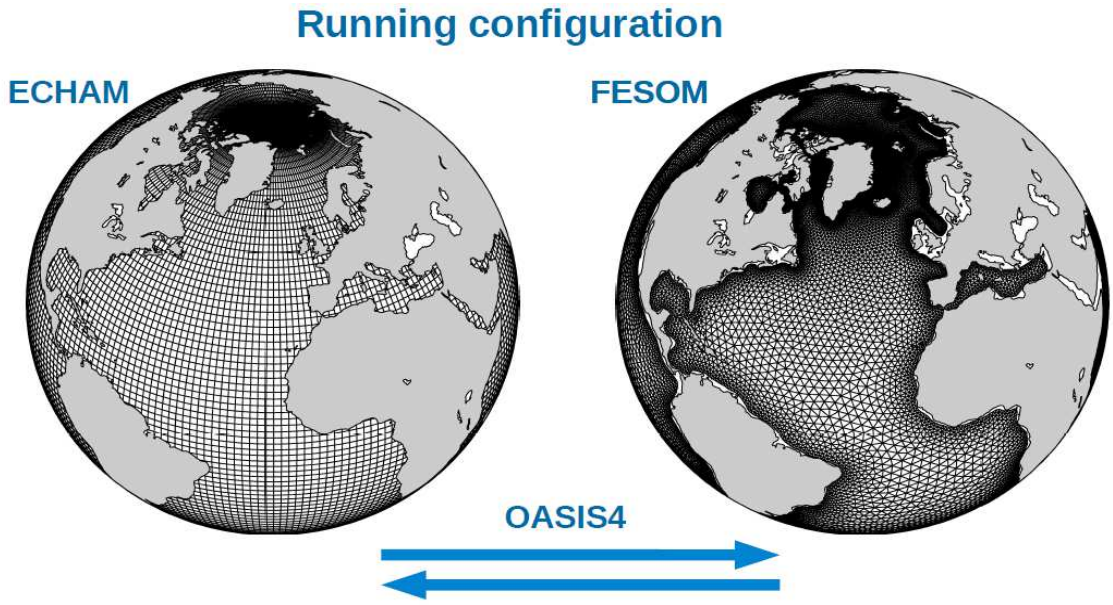

Figure 3: Coupling strategy between FESOM and ECHAM using the parallel OASIS 4 coupler *Sidorenko et al. [2012]*

2.5 Validation of climate models

The first climate model was developed by Kirk Bryan and Michael Cox in 1969 in the Geophysical Fluid Dynamics Laboratory in Princeton. Theirs was a simulation, primitive equation model to calculate the 3-dimensional flow in the ocean. Since then, climate models have come a far way. Many geophysical institutes in the world have developed their own sea ice-ocean model. Therefore comparison between the different models is necessary.
One way to rank different models is to use "Metrics". Metrics is an objective method for ranking, according to the model's efficiency in simulating a certain climate variable. It is based on the relative errors of each model independently for each variable. The first attempt to achieve a thorough model intercomparison using metrics was carried out by Gleckler. Current interest is focused on applying metrics to sea ice-ocean and coupled models. However, such a task of validation is not an easy one due to the sparse observations in the ocean.

2.6 The Taylor diagram

A Taylor diagram is a 2 dimensional plot showing three statistical quantities; the ratio of variances of both simulated and observed fields, the centered root-mean-square error and the correlation coefficient between the two fields for the model variable under consideration in one point. It can to some extent summarize the agreement between the observed and the simulated. It is useful in assessing the relative merits of competing models and in monitoring overall performance as the model evolves *[Taylor, 2001]*.

In order to determine the degree of agreement between the observed and simulated fields, we have to determine the pattern of each. The term "pattern" is used in the generic sense, not restricted to spatial dimensions *[Taylor, 2001]*. A pattern can be characterized by a phase and an amplitude. We need to determine how the phasing and amplitude for the two patterns differ from each other at each grid point.
The correlation coefficient is the quantity that measures the degree of agreement of the phases of both fields. It can tell us if the two fields are simultaneously increasing (r=1), or simultaneously decreasing (r=-1), or there is no relation (r=0). The centered root-mean-square error difference is the quantity that measures the degree of agreement in amplitude. It is simply the sum of the anomaly differences of the two fields. The correlation coefficient and the centered pattern root-mean-square difference provide complementary statistical information quantifying the correspondence between the two patterns *[Taylor, 2001]*. The variances are the other two quantities needed to give a full description of the two patterns.

Consider two fields the observed o and the simulated s. Assume we have N values for the two fields, calculated on the same N point grid. We calculate the correlation coefficient by using the following formula :

$$R = \frac{1}{N}\sum_{n=1}^{N}\frac{(s_n - \bar{s})(o_n - \bar{o})}{\sigma_o \sigma_s} \tag{6}$$

where $\bar{s}$ is the mean of the simulated pattern, and $\bar{o}$ is the mean of the observed pattern. To quantify the difference in the two fields we need to find the centered root-mean-square difference by using the following formula :

$$E = \sqrt{\frac{1}{N}\sum_{n=1}^{N}[(s_n - \bar{s}) - (o_n - \bar{o})]^2} \tag{7}$$

The previous two quantities along with the standard deviation of the simulated and observed fields, form the sides of a triangle as shown in figure 4. We can write the triangle equation for the desired statistical quantities as :

$$(\frac{E}{\sigma_o})^2 = (\frac{\sigma_s}{\sigma_o})^2 + 1 - 2R(\frac{\sigma_s}{\sigma_o}) \tag{8}$$

Note that the equation above is normalized. This is useful for plotting different variables with different units on one diagram.

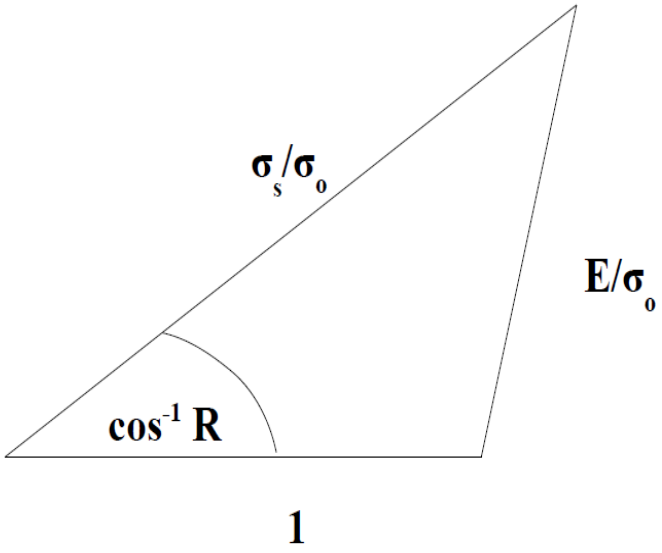

Figure 4: Geometric relationship between the correlation coefficient R, the normalized centered pattern NRMS error E/σ_o, and the ratio of standard deviations σ_s and σ_o of the simulated and observed fields respectively *[Taylor, 2001]*.

The previous equation is the basis of the Taylor diagram. The Taylor diagram as shown in figure 5 is a polar style graph. The dashed green circles correspond to circumferences of equal centered normalized root-mean-square (NRMS) difference, the dotted-dashed blue lines correspond to lines of equal correlation coefficients, and doted black circles correspond to circumferences of equal ratio of standard deviations. The point denoted by "ref" is the point of perfect simulation. At this point the centered RMS error is equal to 0 and both the ratio of standard deviations and the correlation coefficient are equal to 1. Each model can be represented by a point. Figure 5 illustrates a Taylor diagram constructed for two models. The examiner can determine the ranking of the each model in relation to the other, by comparing the distance to the "ref" point. "model2" appears to be closer

to "ref" than "model1", which means that it has a higher correlation coefficient, a lower NRMS difference and a ratio of standard deviations closer to 1 than the other point. Thus, we conclude that "model2" performs a better simulation.

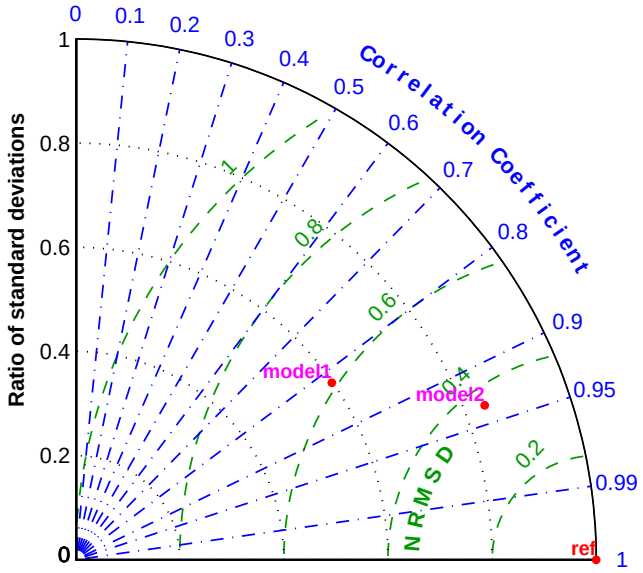

Figure 5: The Taylor Diagram

In obtaining the Taylor diagram for sea ice variables two processes have to be carried out. The first is interpolating the simulated and observed fields to the same grid, in order for all the grid elements to have the same weight. The second is to locate grid points where no values have been produced in both fields, and exclude these points from the calculations.

2.7 Observations

2.7.1 ICESat

Sea ice thickness observational data were obtained from the National Snow and Ice Data Center (NSIDC). NSIDC archieves and distributes data products from the Geoscience Laser Altimeter System (GLAS) instrument on board the NASA Ice, Cloud and Land Elevation satellite (ICESat), launched on 12 January 2003.

ICESat has completed ten data acquisition campaigns from 2003 to 2007. Each campaign spanned approximately 35 days, with a separation between spring and autumn campaigns lasting for about 3 months. This sampling strategy is followed in order to detect seasonal and interannual cahnges in the global ice cover *[Kwok and Cunningham, 2008]*. ICESat campaign periods are listed in tables 1 and 2

Campaign (mmyy)	Period	Days of Operation
ON03	Sep 24-Nov 18	55
ON04	Oct 03- Nov 08	37
ON05	Oct 21-Nov 24	35
ON06	Oct 25-Nov27	34
ON07	Oct 02-Nov 05	37

Table 1: Table of ICSat autumn campaign periods

The Geoscience Laser Altimeter Sysytem (GLAS) is used in obtaining sea ice thickness estimates from measuring "freeboards", in a process known as "freeboard retrieval". ICESat freeboard h_f is defined as the vertical distance between the air-snow interface and the local sea surface, as illustrated in figure 6 *Kwok and Cunningham [2008]*. h_f consists of a snow layer, h_{fs} and an ice layer, h_{fi}. In order to estimate h_{fs}, meteorological products have to be used. In the case of ICESat the ECMWF (European Centre for Medium-Range Weather Forcast) snowfall estimates were used. After obtaining measurements for h_f and h_{fs}, the total ice thickness h_i can be calculated as follows:

$$h_i = (\frac{\rho_w}{\rho_w - \rho_i})h_f - (\frac{\rho_w - \rho_s}{\rho_w - \rho_i})h_{fs} \tag{9}$$

Campaign (mmyy)	Period	Days of Operation
FM04	Feb 17-Mar 21	34
FM05	Feb 17- Mar 24	36
FM06	Feb22-Mar27	34
MA07	Mar 12-Apr 24	34

Table 2: Table of ICSat spring campaign periods

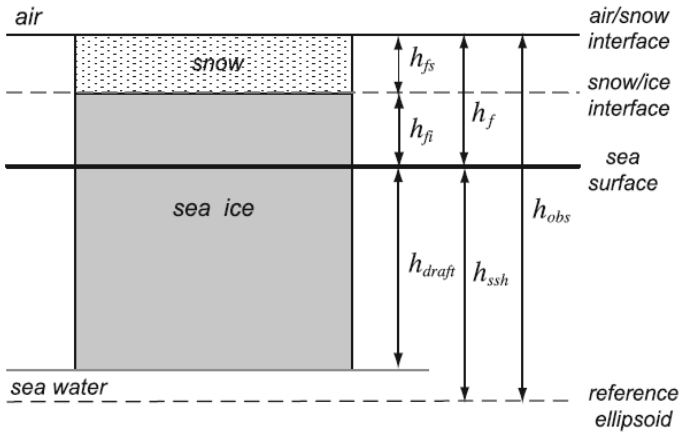

Figure 6: Freeboard retrieval scheme *[Kwok and Cunningham, 2008]*

where ρ_i is the density of ice, ρ_s is the density of snow and ρ_w is the density of water.

We expect ICESat data to have random and bias errors in the measurements themselves, and sampling and analysis errors due to data processing. Estimating the uncertainties in ICESat data requires to take the assumptions made in "freeboard retrieval" into consideration. *Kwok and Cunningham [2008]* estimate the overall uncertainty in the ice thickness estimates within 25-km ICESat segment to be 0.7 m but it varies with the relative thickness of the total freeboard and snow depth. Uncertainities are higher over MY ice due to deformation.

2.7.2 Nimbus 7 satellite

The National Snow and Ice Data Center NSIDC provides an up-to-date source of sea ice concentration values and images. The data are provided in the polar stereographic projection at a grid cell size of 25 x 25 km. NSIDC provides a consistent time series of sea ice concentrations (the percentage of ocean area covered by sea ice)*[Cavalieri et al., 1996]*.
The data are generated using the NASA Team algorithm developed by the Oceans and Ice Branch Laboratory for Hydrospheric Processes at the NASA Goddard Space Flight Center (GSFC) *[Cavalieri et al., 1996]*.
These data include gridded monthly averaged sea ice concentrations for the north polar regions, up to 86 °N. Data is provided from 26 October 1978 and are produced from the Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR). SMMR records bright-

ness temperatures from the surface. SMMR has a dual-polarization and many spectral channels. This is useful in distinguishing between FY and MY ice, since each type of ice has its own microwave characteristics. In general MY ice is more difficult to distinguish because of deformation.

The sensor receives brightness temperatures from the polar surface according to the equation :

$$T_B = T_W(1 - C) + T_{FY}C(1 - F) + T_{MY}CF \tag{10}$$

where T_W is the radiance of open water, T_{FY} and T_{MY} are the radiances of first-year ice and multi-year ice, which are polarization and wavelength dependent. The sea ice concentration C, is the percentage of ocean covered by sea ice within the instruments field of view. F is the percentage of muti-year sea ice. The principle channels used for sea ice observations are the 0.81 and 1.7 cm horizontal and vertical *[Cavalieri and Gloersen, 1984].* Usually MY ice is located in the middle of the arctic basin, whereas FY ice is located in the edges. In the Antarctic, most sea ice is FY ice.
Errors are usually difficult to estimate because of the many assumptions made in the process of estimating ice concentrations. *Cavalieri and Gloersen [1984]* estimates the relative accuracy to be 13-25 %.

3 Results

3.1 Sea ice thickness

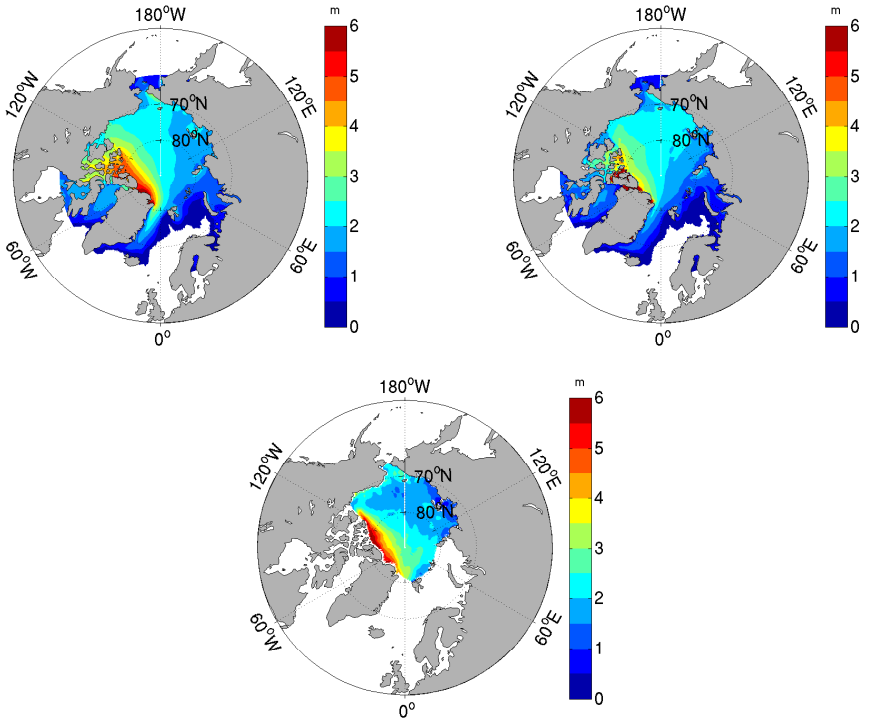

Figure 7: Sea ice thickness spring mean 2003-2007, FESOM (upper left), MITgcm (upper right) and ICESat (below). Values of zero are not shown. We notice that the two models simulate correctly the distribution of ice, with FESOM having the apparent advantage.

Spring and autumn sea ice thickness temporal mean maps were constructed for the ICESat periods for FESOM, MITgcm and ICESat, in figures 7 and 8. ICESat periods (tables 1 and 2) span approximately 35 days in February-March and October-November from 2003 to 2007. The means were constructed using an intermediate grid with a resolution of 0.1°.

Considering the spring mean map for ICESat (figure 7 bottom), we notice that for most of the Arctic basin the ice thickness is about 2 m. In the Laptev Sea where the ice is produced, thickness is around 1 m, the thinnest to be found in this region. North of Greenland and the Canadian Arctic Archipelago, we find the thickest ice, with values ranging between 4-6 m.

Next consider both FESOM and MITgcm simulations (figure 7, upper left and upper right).

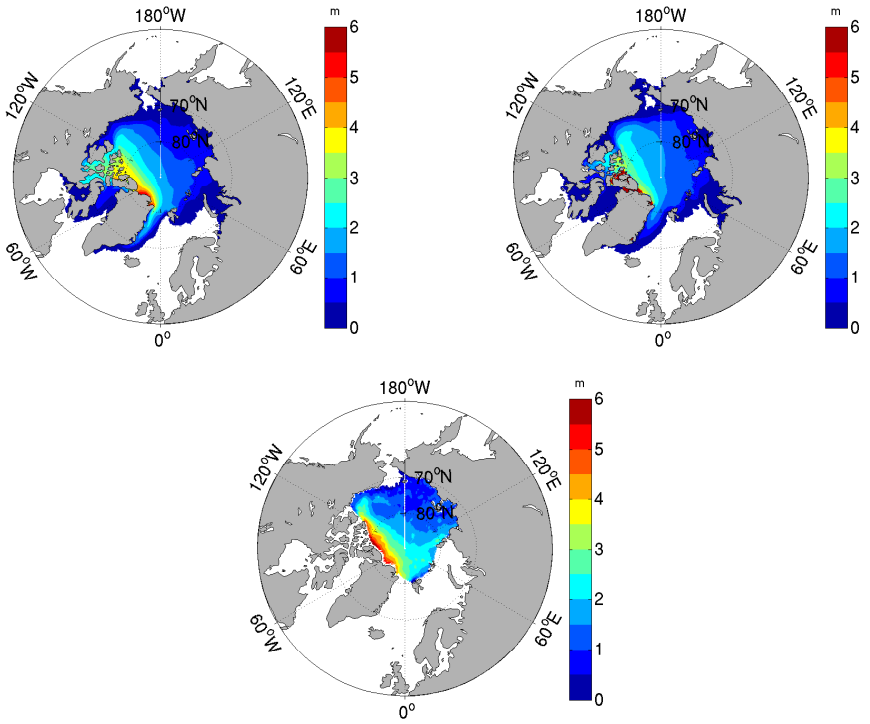

Figure 8: Sea ice thickness autumn mean, 2003-2007, FESOM (upper left), MITgcm (upper right) and ICESat (below). Values of zero are not shown. Notice that both models shows lower thickness values north of Greenland than the satellite shows.

Notice that the models produce sea ice in parts where no observations are available, in the Kara sea, parts of the Barents sea, the Greenland sea and in the Baffin bay until the Davis strait.

FESOM simulated the overall ice thickness distribution quite successfully. FESOM produces the 2 m mean thickness over most of the Arctic ocean, the thickest ice north of Greenland and the thinnest ice in the Laptev sea. MITgcm shows thinner ice overall, but has correctly simulated the location of the thicker ice north of Greenland (however with low values), and the very thin ice in the Laptev sea.

For the autumn mean map, figure 8, notice the overall shrinkage of sea ice area, and the lower overall mean (around 1 m). On the ICESat September map (figure 8, bottom), the thickest ice north of Greenland covers a smaller area than it did in spring and has sightly lower values ranging from 5-6 m. FESOM simulates the thicker ice in the same region but with lower values. MITgcm produces thinner ice overall. However, it produces thickness values ranging between 5-6 m between Greenland and the Canadian Arctic Archipelago, and around Ellesmere Island. Overall FESOM appears to simulate ice thickness better

than MITgcm.

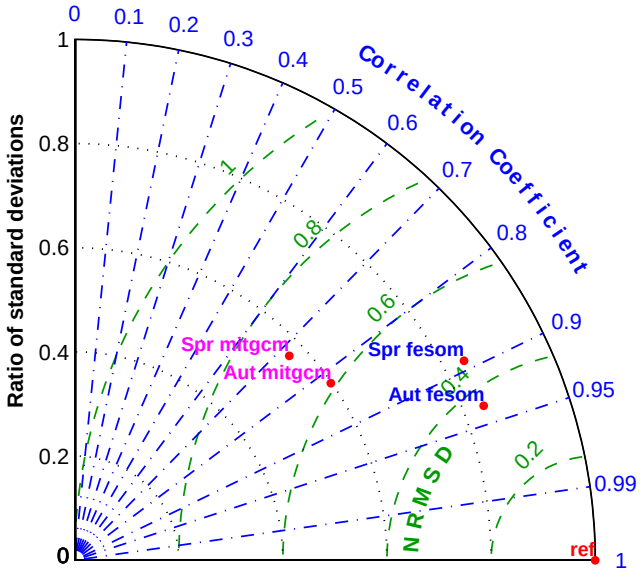

Figure 9: Taylor Diagram for sea ice thickness

The Taylor diagram in figure 9 is constructed for sea ice thickness spring and autumn means 2003-2007. The satellite and model thickness values were interpolated on the same intermediate grid with a 0.1 ° resolution. At some grid points the models produced values where no observations were available. These grid points cannot be used in the construction of the Taylor diagram, therefore, they were excluded from the calculations of the necessary statistical quantities. Also, grid points where all FESOM, MITgcm and ICESat were equal to zero have been excluded as well.

Figure 9 contains 4 points: "Spr fesom", "Aut fesom", "Spr mitgcm" and "Aut mitgcm". In order to evaluate model performance, we draw a line from the reference point "ref" to each point and measure the distance. The model point with the smallest distance is the model with the best simulation. On the diagram we notice two clusters, one for each model. The FESOM cluster is closer to the "ref" point than the MITgcm cluster. Hence, we are able to conclude that FESOM produced a closer simulation to the ICESat measurements than MITgcm. We also notice that the point "Aut fesom" is the nearest of

the two FESOM points, indicating that the autumn simulation is better. Going back to the visual comparison in (7) and (8) we notice that for FESOM, the simulation for spring appeared better than it did in autumn. In autumn FESOM produced lower values north of Greenland than is seen the ICESat map. This is one of the examples that illustrate the helpfulness of the Taylor diagram. It is an objective measure of model performance.
By considering the two MITgcm points, we notice that "mitgcm aut" is closer to "ref" than "mitgcm spr".

3.2 Sea ice concentration

Figures 10 and 11 show the constructed March and September means 1992-2007, for FESOM, MITgcm, FESOM/ECHAM and NSIDC. For March (figure 10), FESOM and MITgcm simulations are quite similar. Sea ice concentration is above 95 % over the entire Arctic ocean, the continental shelves, the Canadian Arctic Archipelago, west of Greenland and down to the Hudson Bay and the Okhotsk sea. Both model simulations are in accordance with the NSIDC data, except for some parts were the NSIDC concentration goes down to 90 % (around the coasts in the Hudson bay and the Nansen Basin). The FESOM/ECHAM simulation is close to the observations. However, the coupled model fails to give results near the coastlines and in the Okhotsk sea, whereas it produces sea ice well into the Labrador sea. For September figure 11, NSIDC (lower right) shows that

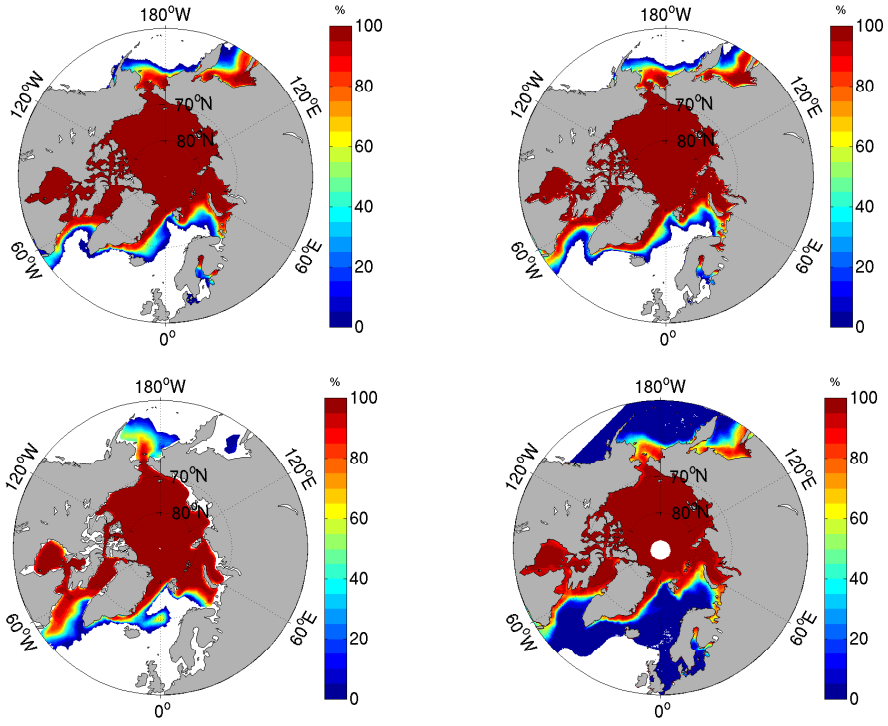

Figure 10: Sea ice concentration March mean 1992-2007. FESOM (upper left), MITgcm (upper right), FESOM/ECHAM (lower left) and NSIDC (lower right). The hole indicates the area above 86 °N were NSIDC measurements where not available. Concentrations of zero are not shown

concentrations start low near the coastlines and gradually rise at the end of the continental shelves, reaching a maximum of 85-95 % north of Greenland and the Canadian Arctic Archipelago. FESOM (upper left) is quite similar to NSIDC (lower right). North of Green-

land FESOM seems to produce similar concentration values as the ones seen in the NSIDC map. However, at the continental shelves FESOM has higher values. MITgcm produces higher concentrations overall. The gradual rise in concentration starts at the coastlines in some regions. The maximum concentration north of Greenland has higher values than both NSIDC and FESOM. FESOM/ECHAM produces over all lower concentrations and it

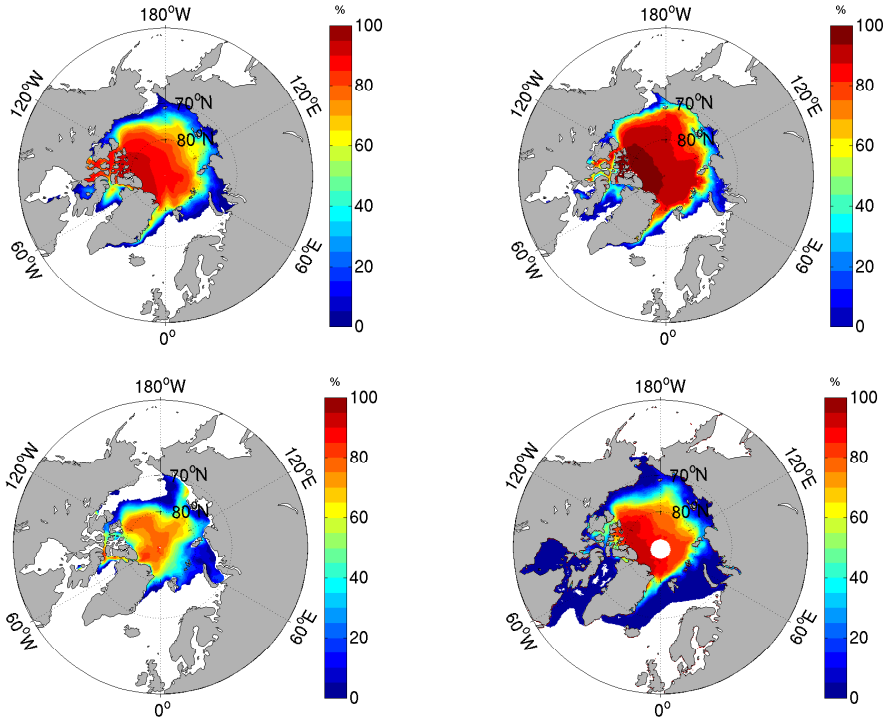

Figure 11: Sea ice concentration September mean 1992-2207. FESOM (upper left), MITgcm (upper right), ECHAM/FESOM (lower left) and NSIDC (lower right). The hole indicates the area above 86 °N were NSIDC measurements were not available. Concentrations of zero are not shown

fails to produce the maximum concentration north of Greenland and the Canadian Arctic Archipelago. FESOM/ECHAM fails to produce ice near the coastlines.

The Taylor diagram shown in figure 12 was obtained from the temporal concentration mean maps for March and September 1992-2007. It was constructed after interpolating all model values to the NSIDC grid. As mentioned earlier some grid points for the models or the observations have no values. Values at these grid points have to be excluded from the construction of the diagram. Moreover, grid points where all FEOSM, MITgcm, FESOM/ECHAM and NSIDC had zero concentration were excluded too. After describing the concentration maps, the reader might expect that the Taylor diagram would indicate

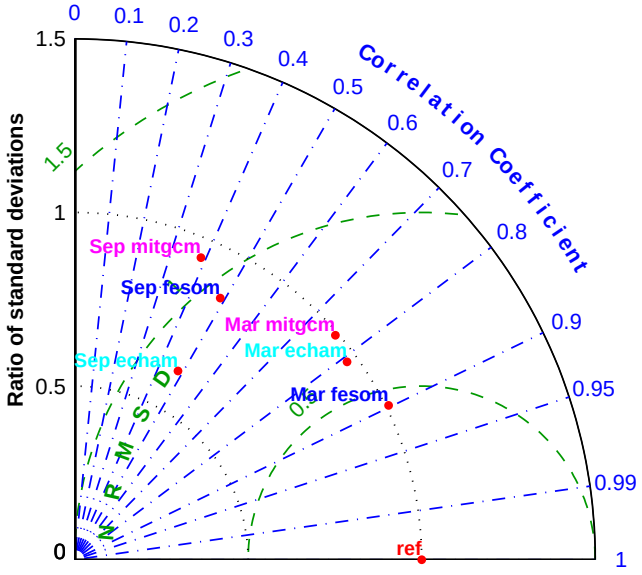

Figure 12: Taylor Diagram for sea ice concentration

that FESOM has the best results, and FESOM/ECHAM has the worse. Is this the case? To answer this question we start to measure the distances of each point from the "ref" point. We find "Mar fesom" to be the nearest point, hence, achieving the best simulation among the six points. Comparing the March points to the September points, we find that the March cluster is nearer to the "ref" point than the Autumn cluster, suggesting that the March cluster achieved the better simulations. The better March simulations could be explained by more melting in Autumn, which leads to higher variability in concentration, whereas in March the concentrations are above 95 % almost everywhere. The standard deviations for "Mar fesom" and for "Mar mitgcm" are nearly equal to 1, indicating almost perfect simulation of sea ice concentration variability. In the September cluster "Sep fesom" has the smallest distance to "ref". "Sep fesom" and "Sep echam" have very close correlation coefficients, but "Sep echam" has the lower ratio of standard deviation, which is to be expected, according to the visual comparison. According to figure 11, we would expect "Sep mitgcm" to be closer to "ref" than "Sep echam". That is clearly not the case. Again the usefulness of the Taylor Diagram is illustrated, by giving an objective result of model comparison.

3.3 Sea ice concentration standard deviation

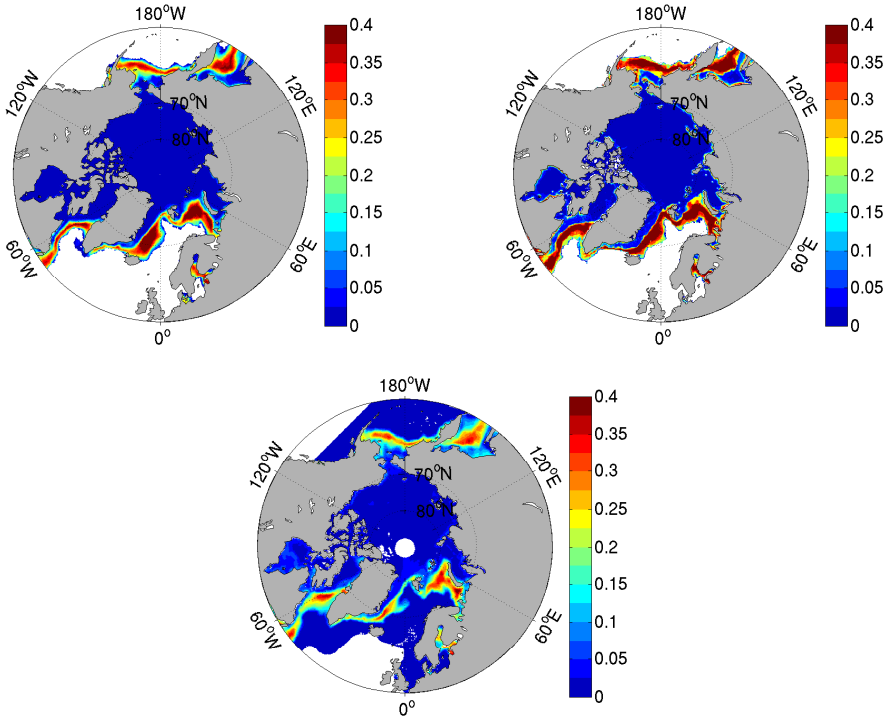

Figure 13: Sea ice concentration standard deviations from the March mean 1992-2007. FESOM (upper left), MITgcm (upper right) and NSIDC (below). The hole indicates the area above 86 °N were NSIDC measurements were not available. Values of zero are not shown

Sea ice concentration standard deviations from the March and September means 1992-2007 are shown in figures 13 and 14. They are constructed for FESOM, MITgcm and NSIDC. For March (figure 13) notice the higher values at the edges, indicating the transition from ice covered areas to open water. September is shown in 14. Notice the shift in the edges to the interior of the Arctic basin. MITgcm shows higher values. This was not expected from the visual comparison of the concentration maps figure 10 and figure 11, where for spring MITgcm produced similar results as FESOM, and for September it produced higher concentrations at the continental shelves.

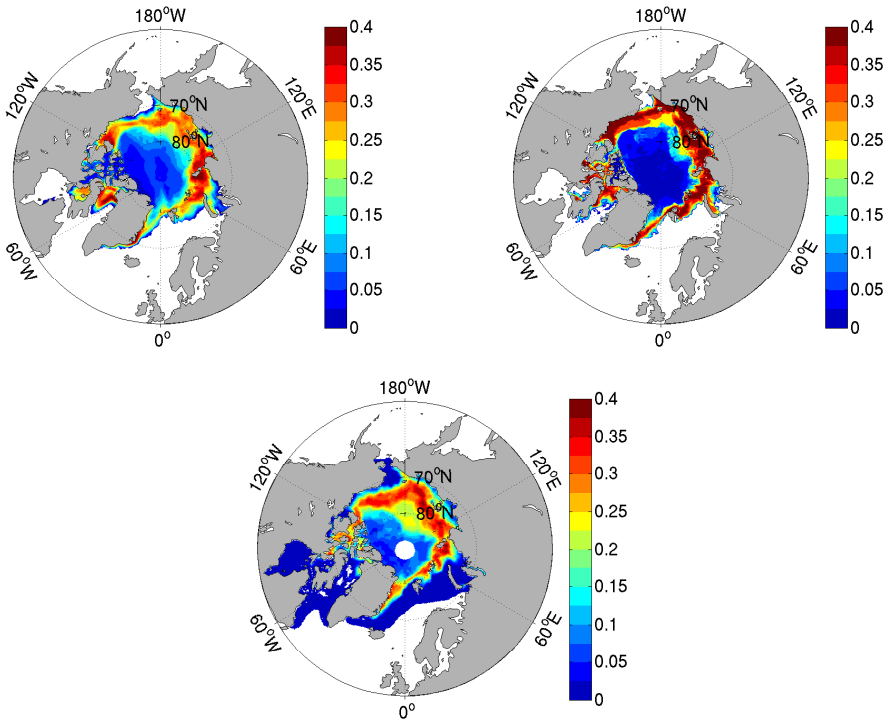

Figure 14: Sea ice concentration standard deviations from the September mean. FESOM (upper left), MITgcm (upper right) and NSIDC (below). The hole indicates the area above 86 °N were NSIDC measurements were not able to be obtained. Values of zero are not shown

The Taylor diagram for sea ice concentration standard deviation is shown in figure 15. Again the same procedure of excluding grid points where no results were obtained for any of FESOM, MITgcm and NSIDC, and grid points where all three were equal to zero, was followed. The two points "Mar fesom" and "Sep fesom" are clearly closer to the "ref" point than "Mar mitgcm" and "Sep mitgcm" suggesting that FESOM simulations are better.
Comparing the two FESOM points in order to determine the point with better simulation, we notice that "Mar fesom" has a higher correlation coefficient by almost 0.1 whereas "Sep fesom" has a better ratio of standard deviations almost equal to one. The NRMS error of both points seem to be equal. We can conclude that "Mar fesom" performed better.
Now comparing the two MITgcm points we notice that both have close NRMS errors and ratio of standard deviations, but "Mar mitgcm" has a higher correlation coefficient. We conclude that "Mar mitgcm" performed better.
The better performance of the March simulations can again be attributed to the high concentrations over the entire Polar region.

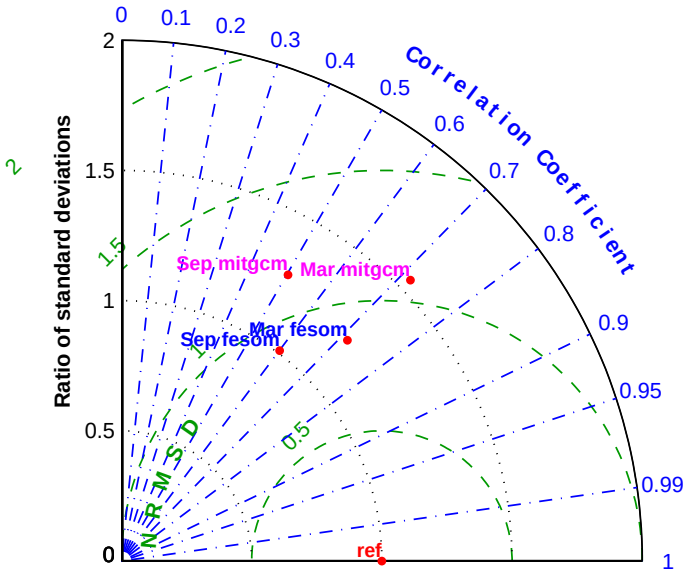

Figure 15: Taylor Diagram for sea ice concentration standard deviation

3.4 Sea ice extent

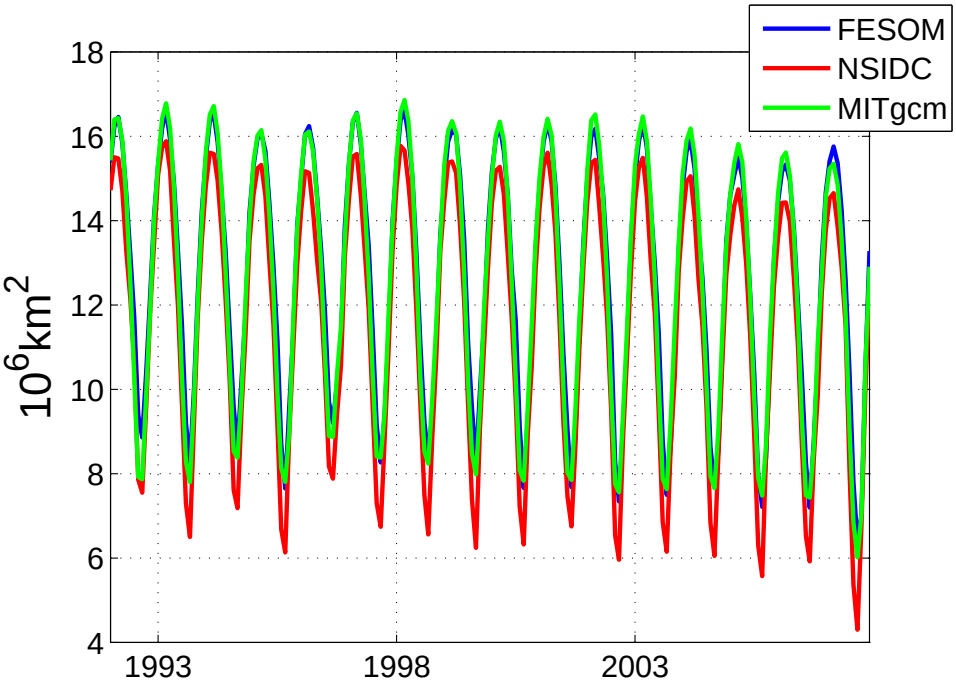

Figure 16: Arctic sea ice extent FESOM, MITgcm and NSIDC for the period 1992-2007

Sea ice extent is the total area of sea covered with at least 15 % of ice. It is obtained by summing grid area's of all grid cells adhering to the 15 % criteria. In general simulating sea ice extent anomaly is more of a challenge than simulating sea ice extent. Figure 16 shows the time series for Arctic sea ice extent. The models are in agreement with the observations in terms of the phase of the time series. However, the amplitude of extent simulated by both models, is slightly higher, than the NSIDC amplitude.

Figure 17 shows the Arctic sea ice extent anomaly. Both models start in phase with NSIDC. Shortly after that MITgcm seems to have an opposite phase as both FESOM and NSIDC, but soon returns in phase with the other two (exept for two points near 1995 and 2006), however with smaller amplitudes. Notice that the models correctly produce the high peek around 1997 and the low peak near 2007. FESOM is found out of phase with NSIDC at a few points (around 1999 and 2006). We conclude that FESOM is in accordance with NSIDC, and to lesser extent MITgcm with NSID.

Arctic sea ice extent annual cycle is shown in figure 18. Model monthly averages are higher than the NSIDC average in accordance with the extent time series (figure 16). MITgcm

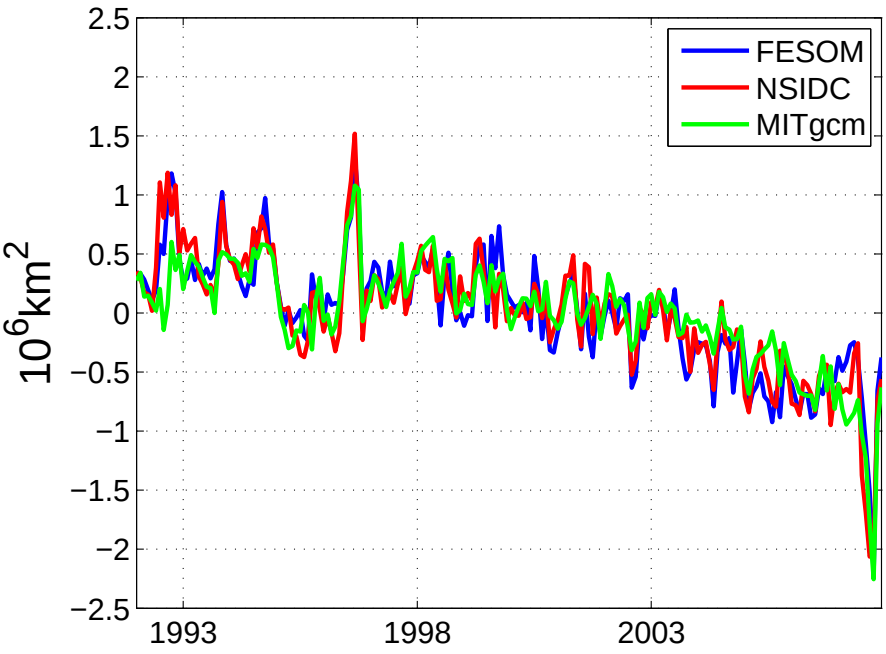

Figure 17: Arctic sea ice extent anomaly, FESOM, MITgcm and NSIDC for the period 1992-2007

has slightly lower values than FESOM from March to September. Seasonal minima and maxima are produced at the correct months.

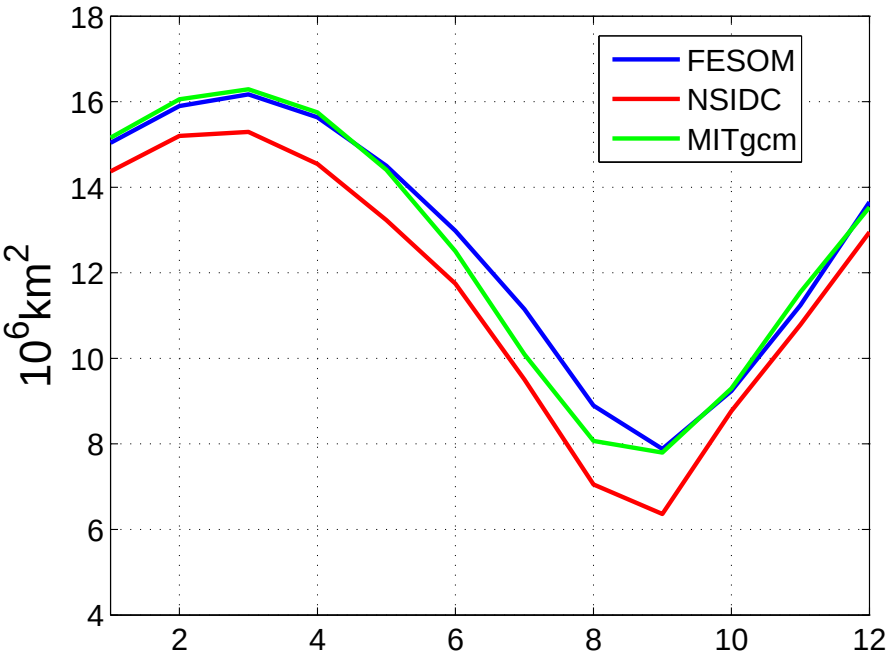

Figure 18: Arctic sea ice extent annual cycle

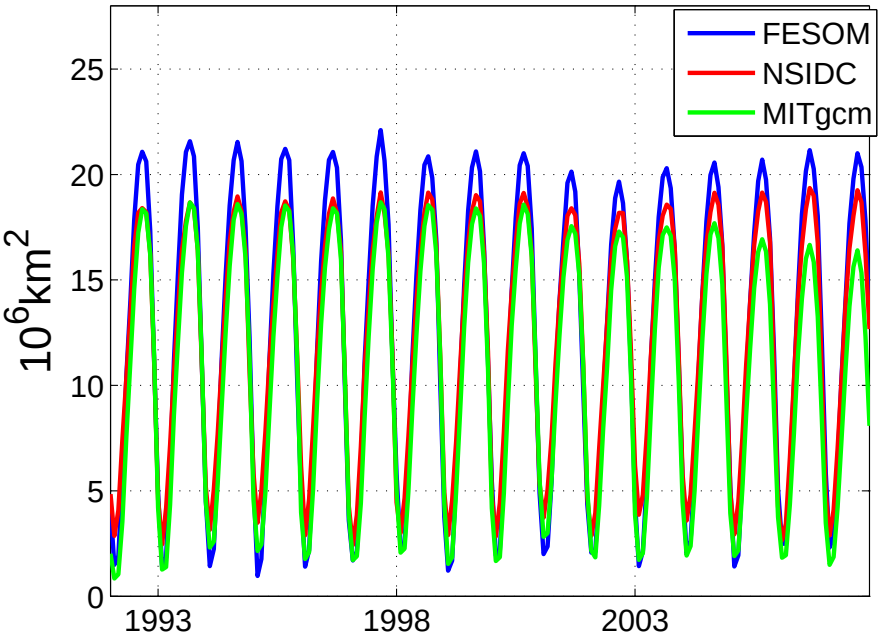

Figure 19: Antarctic sea ice extent FESOM MITgcm and NSIDC for the period 1992-2007

Antarctic sea ice extent time series is shown in figure 19. As in the case of the Arctic, the two models and the observations are in agreement in terms of phase. Comparing the models with NSIDC, we notice that FESOM has higher spring amplitudes. For MITgcm spring amplitudes starts similar to those of NSIDC but after 1993 they become lower.

Antarctic sea ice extent anomaly figure 20 shows disagreement in phase between each of the two models and the observation. The amplitude of the extent anomaly for the models are sometimes similar to the NSIDC amplitude and other times not. For example, notice that MITgcm has twice the amplitude in the opposite direction as NSIDC and FESOM at the very beginning, and it has the right amplitude but in the opposite direction towards the very end. FESOM underestimates the amplitude at various points. MITgcm fails to simulate the amplitude correctly in various places.
The seasonal cycle is shown in figure 21. In the Antarctic both models produce more ice in spring and less ice in autumn. FESOM and MITgcm produce the same amount of ice in autumn, and FESOM produces more ice than MITgcm in spring.
The Taylor diagram for sea ice extent anomaly (figure 22) shows that results for the Arctic

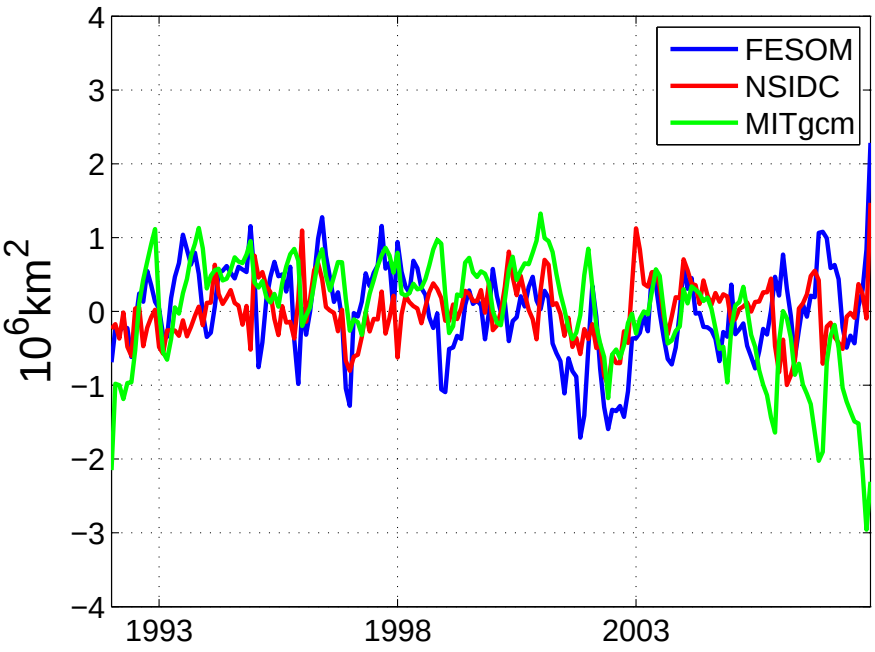

Figure 20: Antarctic sea ice extent anomaly, FESOM, MITgcm and NSIDC for the period 1992-2007

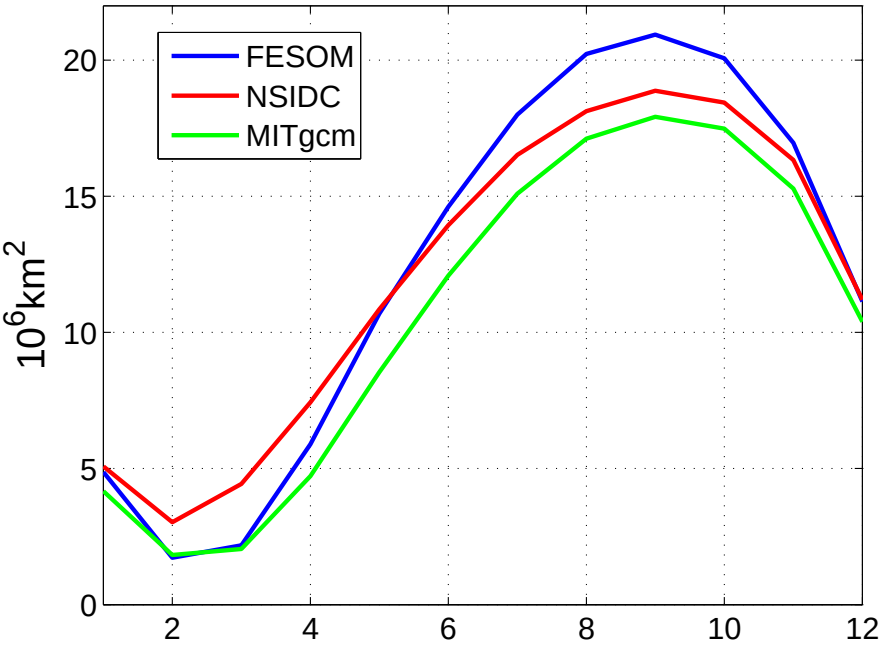

Figure 21: Antarctic sea ice extent cycle

simulation is better than it for the Antarctic. This was expected from the anomaly time series. The reason might be due to the various difficulties in simulating the dynamics of the Antarctic ocean. Small errors in simulating sea ice velocities lead to more or less ice being pushed in or out of the southern ocean, leading to different concentrations.
The points "Arc fesom" and "Arc mitgcm" almost overlap indicating similar model simulations for the Arctic. In the Antarctic FESOM has an obvious advantage.

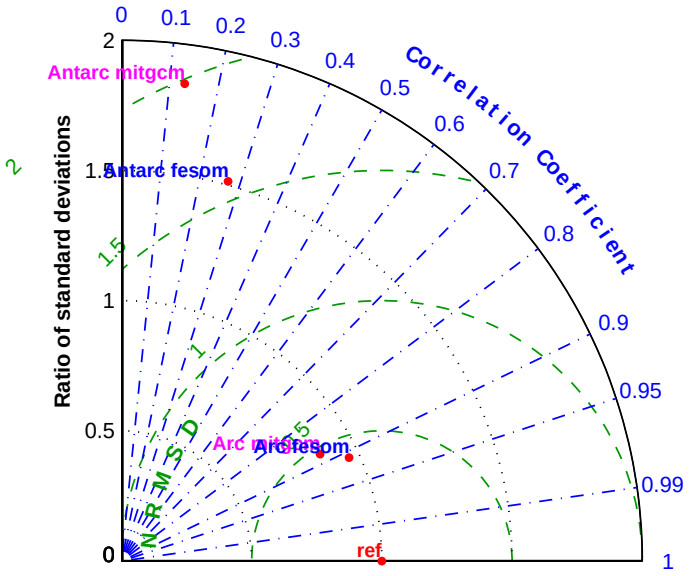

Figure 22: Taylor diagram for sea ice extent anomaly

4 Discussion and Conclusion

In this study the Taylor diagram has been used as a mean for ranking models according to their performance in simulating sea ice variables. It is one of the early attempts to use the Taylor diagram for assessing sea ice-ocean models and coupled sea ice-ocean-atmosphere models. The models used are two ocean-sea ice models, namely FESOM and MITgcm, and one coupled ocean-sea ice-atmosphere model, namely ECHAM/FESOM. The taylor diagram's ranking is based on three statistical quantities that measure the the difference in both amplitude and phase between the observed field and the simulated field. The performance of the model is summarized in one point.

A visual comparison of the constructed temporal means between the simulated and observed fields for sea ice thickness, sea ice concentration and sea ice concentration standard deviation, and of the time series for sea ice extent anomaly was carried out. These maps were later compared to the results of the Taylor diagrams.

The visual comparison and the diagram's rankings were not always consistent. For sea ice thickness the visual comparison led to the assumption that the spring simulation is better than the autumn simulation. However, on the taylor diagram autumn points had a higher correlation coefficient and a lower centered NRMS difference. One explanation for this discrepancy may be attributed to the fact that the statistical quantities are defined for Gaussian distributions, for which sea ice thickness is definitely not.

For sea ice concentration, there was agreement between the Taylor diagram and the visual comparison. Both suggested that results for March were better that it for September. The explanation for the worse results in September can be attributed to melting which leads to more spatial variation in concentration.

For sea ice concentration standard deviation, visual comparison and the Taylor diagram lead to the same conclusion, that FESOM results were in better agreement with NSIDC than MITgcm's results.

Finally for sea ice extent anomaly, the Taylor diagram verified the results of the time series for both the Arctic and the Antarctic. Antarctic results for both models were not in agreement with NSIDC. This is due to the difficulties in simulating sea ice in that region.

There are a number of shortcomings of the Taylor diagram that has to be taken into consideration. Maybe the most obvious one is that there are no error estimates. The points that represent the model's performance has no error bar, giving the impression that the position of the points are "true" *[Taylor, 2001]*. Moreover, errors in the observations have not been taken into account. For climate system variables, it is in general difficult to es-

timate observational errors. Sources of uncertainty include random and systematic errors, and analysis errors when the observational data are processed. *Gleckler and Taylor [2008]* argue that ignoring errors in the reference data is acceptable as long as these errors remain much smaller than the errors in the models. Also, satellite data are usually available for only a few years, whereas model climatologies compared with such data are usually computed from twenty or more simulated years. *Gleckler and Taylor [2008]* further state that for the metrics considered here, simple sampling tests (performed on model simulations) indicate that the impacts of uncertainties associated with a limited observational record are small when compared to the magnitude of current model errors.

Even though the Taylor diagram has some short comings, it can still be used as one of the tools in assessing model performance. The main advantage of the diagram is that it provides a summary of performance. One useful application of the Taylor diagram is assessing performance in the process of model evolution. It can be used at each step of the model development to get some kind of indication of the results of the modification. Another useful application is assessment in cases where visual comparison becomes difficult or impractical, as in the case when the variable under consideration is both a function of space and time.

From the results presented in this work we conclude that for the configurations used, FESOM appears to simulate sea ice variables better than MITgcm.

References

D. Cavalieri and P. Gloersen. Determination of sea ice parameters with the nimbus 7 smmr. *J. Geophys. Res.*, 1984.

D. Cavalieri, C. Parkinson, P. Gloersen, and H. J. Zwally. Sea ice concentrations from nimbus-7 smmr and dmsp ssm/i-ssmis passive microwave data. *Boulder, Colorado USA: National Snow and Ice Data Center. Digital media*, 1996.

M. Conkright, A. Locarnini, H. Garcia, T.O'Brien, T. Boyer, C. Stephens, and J. Antonov. World ocean atlas 2001: Objective analysis, data statistics, and figures, cd-rom documentation. *National Oceanographic Data Center, Silver Spring*, 2002.

P. J. Gleckler and K. E. Taylor. Performance metrics for climate models. *JOURNAL OF GEOPHYSICAL RESEARCH*, 113, 2008.

W.D. Hibler. A dynamic thermodynamic sea ice model. *Phys. Oceanogr*, 9:815–846, 1979.

R. Kwok and G. F. Cunningham. Icesat over arctic sea ice: Estimation of snow depth and ice thickness. *J. Geophys. Res.*, 113, 2008.

W.G. Large and S.G. Yeager. The global of an interannually varying air-sea flux data set. *Clim. Dyn*, 33, 2008.

P. Lemke, J. Ren, R.B. Alley, I. Allison, J. Carrasco, G. Flato, Y. Fujii, G. Kaser, P. Mote, R.H. Thomas, and T. Zhang. Climate change 2007: The physical science basis. contribution of working group i to the fourth assessment report of the intergovernmental panel on climate change. In S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L. Miller, editors, *Observations: Changes in Snow, Ice and Frozen Ground*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

M. Losch, D. Menemenlis, J.Campin, P. Heimbach, and C. Hill. On the formulation of sea-ice models. part 1: Effects of different solver implementations and parameterizations. *Ocean Modelling*, 2010.

A. Proshutinsky, M. Steele, J. Zhang, G. Holloway, N. Steiner, S. Haekkinen, D. Holland, and C. Koeberle R. Gerdes, M. Karcher, W. Maslowski, Y. Zhang, W. Hibler, and J. Wang. The arctic ocean model intercomparison project. *AOMIP*, 2001.

R. Redler, S. Valcke, and H. Ritzdorf. Oasis4-a coupling software for next generation earth system modelling. *Geoscientific Model Development*, 2010.

E. Roeckner, G. Baeuml, L. Bonaventura, R. Brokopf, M. Esch, M. Giorgetta, S. Hagemann, I. Kirchner, L. Kornblueh, E. Manzini, A. Rhodin, U. Schlese, U. Schulzweida, and A. Tompkins. The atmospheric general circulation model echam5 part1: Model description. *Max-Planck-Institut fuer Meteorologie*, 2003.

D. Sidorenko, Q. Wang, S. Danilov, and J. Schroeter. Fesom under coordinated ocean-ice reference experiment forcing. *Ocean Dynamics*, 61:881–890, 2011.

D. Sidorenko, T. Rackow, T. Jung, Q. Wang, S. Danilov, J. Schroeter, and D. Barbi. Coupled fesom/echam5 setup, strategies and solutions. *EGU General Assembly, Vienna, Austria, 22 April 2012-27 April 2012*, 2012.

H. Stewart. *Introduction to Physical Oceanography*. Orange Groove Books, 2008.

K. Taylor. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, 106:7183–7192, 2001.

D. N. Thomas and G. S. Dieckmann. *Sea Ice*. Wiley-Blackwell, 2010.

R. Timmermann, S. Danilov, J. Schroeter, C. Boening, D. Sidorenko, and K. Rollenhagen. Ocean circulation and sea ice distribution in a finite element global sea ice-ocean model. *Ocean Modelling*, 27:114–129, 2008.