Norwegian Meteorological Institute met.no

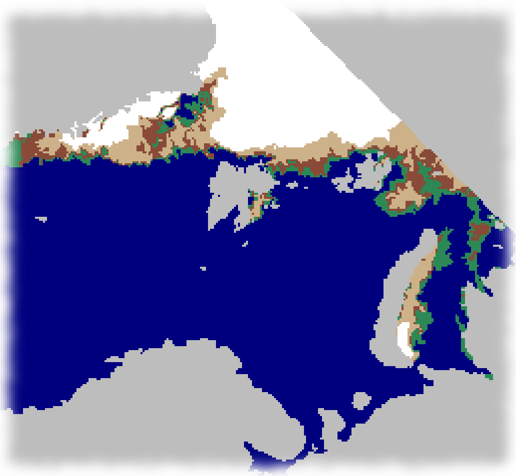
met.no note

Note No. 12/2010
oceanography, remote sensing
Oslo, August 9, 2010

Validation of sea ice concentration in the myOcean Arctic Monitoring and Forecasting Centre[1]

Arne Melsom

myOcean



[1]This document contains hyperlinks that are active when viewed with properly enabled software.

1 Introduction

myOcean is the Marine Core Service project of the Global Monitoring for Environment and Security. The objective of myOcean is to deploy the first concerted and integrated pan-European capacity for Ocean Monitoring and Forecasting. myOcean is co-funded by the 7^{th} Framework Programme (FP7) of the European Union, and by in kind funding from the myOcean partners which include more than 50 institutes.

Targeted areas for use of myOcean products include maritime security, oil spill forecasting, marine resources management, climate change, seasonal forecasting, coastal activities, and water quality and pollution.

Validation of model results my have a number of purposes, but for simplicity we divide these purposes into two categories: (1) information to developers, and (2) information to users. The rational for this division is that

myOcean is a user-oriented project, so our emphasis with the validation activities in the present context, is to provide users with information that can help them assess the quality of the myOcean products. Nevertheless, this constraint does not necessarily exclude the provision of validation products that are also helpful for developers.

Being an activity that relies on comparison of model results and observations, the validation results obviously depend on the platform(s) on which observations have been collected. Frequently observations are incorporated into the model in an analysis whereby the model state is adjusted to reduce its deviation from the observations.

When a subsequent validation is performed, one should have a strong preference to validate model results using other sources of observations than those that are included in the model analysis. Only then can the true performance of the forecast system be assessed, since this system is an integration of a numerical model, observations that are assimilated, and the analysis step.

Another aspect that needs to be taken into account is the quality of the observational data that are used in the validation. A validation which gives results where the root-mean-square differences between model results are of the order of the errors in the observations either point to a very good model or, perhaps more likely, to large uncertainties in the instruments from which the observations are collected.

2 myOcean catalogue products

The Service Management (Work Package 16) of the myOcean project maintains a list of products from all of the project's 7 Monitoring and Forecasting Centres and from its 5 Thematic Assembly Centres. The list of these products is available on-line from `http://catalogue.myocean.eu.org/`.

For the Arctic Ocean region, two relevant observational products for sea ice are available from the Sea Ice and Wind Thematic Assembly Centre. In the myOcean nomenclature, these products are referred to as

SEAICE_ARC_SEAICE_L4_NRT_OBSERVATIONS_011_002 and
SEAICE_ARC_SEAICE_L4_NRT_OBSERVATIONS_011_003.

The former product covers the waters from east of Greenland to Novaya Zemlya, and the western Kara Sea. This product is updated daily on week days. The latter product, which presently is provided irregularly in time, covers the waters off Greenland. Both products provide results for sea ice concentration intervals as defined by the World Meteorological Organization.

Due to its larger spatial domain and more regular updates, we have chosen to use the

SEAICE_ARC_SEAICE_L4_NRT_OBSERVATIONS_011_002

sea ice chart product for the validation of model results. The satellite data used for these ice charts are Synthetic Aperture Radar data from Radarsat and Envisat and visual and infrared data from MODIS and NOAA. The product has a spatial resolution of 1 km.

The model results that are validated by the algorithm that is described in this document, is the nominal sea ice concentration product

ARCTIC_ANALYSIS_FORECAST_PHYS_002_001

from the Arctic Monitoring and Forecasting Centre. This is a product which is available as daily mean fields, updated by a weekly production cycle. The model domain covers the North Atlantic Ocean, the Mediterranean Sea, the Nordic Seas, the Arctic Sea and adjacent shelf seas. The results are available on a spatial resolution of 12.5 km. The weekly production cycle is built from the following components:

1. an analysis step, based on data centred one week prior to the production, and using results from a week-old ensemble simulation (next item)

2. a 7 day ensemble simulation with 100 members, bridging the one-week interval from the analysis to the production date

3. a 10 day deterministic forecast, produced by a continuation of the results from ensemble member no. 1

The analysis step in item 1 includes assimilation of observed sea ice concentrations based on brightness temperatures that are recorded by satellite measurements in the microwave band (either the Special Sensor Microwave/Imager or the Advanced Microwave Scanning Radiometer-EOS).

The model results from the Arctic Monitoring and Forecasting Centre are available as two products, a "best estimate" which is made up of merged results of the ensemble average from all available ensemble simulations from item 2, and a "forecast" product which the set of forecasts from item 3 in the production cycle list above.

Note that the observational product is based on radar data, while the observations that are assimilated in the model system is based on measurements in the microwave band. These sources provide us with independent observations of the sea ice concentration in the analysis step and in the model validation.

3 Validation algorithm

3.1 Background

Assume that we start with the following gridded results for sea ice concentration:

Table 1: Values on a fine spatial resolution.

1.0	1.0	1.0	1.0	0.8	0.2
1.0	1.0	1.0	0.8	0.2	0.0
1.0	1.0	0.8	0.2	0.0	0.0
1.0	0.8	0.2	0.0	0.0	0.0
0.8	0.2	0.0	0.0	0.0	0.0
0.2	0.0	0.0	0.0	0.0	0.0

Here, the values represent the areal average over each cell. The corresponding set of results on a resolution which is coarser by a factor of two along both axes, is well-defined:

Table 2: Values on a coarse spatial resolution.

1.00	0.95	0.30
0.95	0.30	0.00
0.30	0.00	0.00

However, interpolating (and extrapolating) from the coarser grid in Table 2 to the finer grid displayed in Table 1 is not well-defined. One may e.g. use interpolation factors for bi-linear interpolation, or factors representing cubic splines. The only grid cells for which the interpolation may be conducted in a well-defined manner, is those where the average concentration values are 0 or 1 (assuming that the range of valid values are [0,1]).

Next, assume that the results in one of the tables are from a model simulation, and the other table contains observations which we will use in order to validate the model product. The two sets of results are completely in concord with each other. Thus, if e.g. the model is on the coarser grid, the models representation of the observed values agree completely with the actual observations.

All validation metrics should be defined so that model results which are completely consistent with observations, attain a value that can be interpreted accordingly. Based on the examination above, we conclude that all model-data comparison which is performed in validation, should be conducted on the coarser grid on which the two sets of numbers are available.

A standard validation metric is the root-mean-square (rms) difference. A validation should produce a value indicative of a “perfect” model if there are no inconsistencies between model results and observations. In the case of the rms difference, we should then define this quantity so that its value is 0 for a “perfect” model. Note that in this discussion, we have chosen to disregard any observational errors. There will be a brief discussion on this topic towards the end of this section.

Let’s assume that the values in Tables 1 and 2 represent observations and model results, respectively. In the example above, The choice of performing the validation on the coarse grid will provide us with a proper “perfect model” rms difference.

Note that in the case of model results being available on a finer grid than observations, which is frequently the case, it is not obvious that the model is flawless even if the validation produces a rms difference equal to 0. This is so because we have no means of determining whether or not the gradient in model results inside an observational grid cell is realistic.

There are other issues which will make the interpretation of validation metrics less obvious than in the example above. To illustrate such issues, we will discuss the sets at hand for validation of sea ice concentration.

The first issue is that observations are usually not available as gridded averages. Observations are frequently discrete rather than representing an average. Defining a griding procedure which produces areal averages based on observations is anything but straight-forward. In the present case, the observational product is based on a series of snap-shots from radar instruments flying on a number of polar-orbit satellites. Merging this data into a daily product for sea ice concentration is performed subjectively, and it is known that the end product is not a true representation of the daily mean state. Even so, this is the data set we have chosen for validation of the model results for daily averages of the sea ice concentration which are provided to the myOcean users.

Further, the model results, as they are available for the myOcean users, have been interpolated from the actual model grid to a regular grid which is meant to be easier to work with from a user perspective.

The second issue is that the coarser grid rarely overlaps the finer grid as neatly as described here. In the present case, there is not a true overlap since the orientation of the observational grid and model grid are rotated relative to each other. Secondly, a true overlap is not possible since the observations are available on a 1 km $\times$ 1 km grid, whereas the model results are stored on 12.5 km $\times$ 12.5 km, i.e., not a multiple of the observational grid size.

The next issue is the degree of representativeness of the observations with respect to the model variable that is validated. In the present example, we use ice charts that are produced from synthetic aperture radar (SAR) scenes. The data needs to be converted from microwave echo fields to sea ice concentrations, and no algorithm can perform this task exactly.

Finally, all instrument data come with uncertainties (errors), be they small or large. Hence, the observational data are not strictly true, although they are frequently referred to as “truth” in model validation activities. If a standard deviation value for the instrument error can be provided, the validation may be conducted in measures of no. of instrument standard deviations between observations and the corresponding model results. One may then e.g. accept model results within ± 1 standard deviation of the observations as “perfect”. However, such an approach is not implemented in the algorithm for validation of sea ice concentration which is documented in this report.

Moreover, the model results themselves have uncertainties that are not necessarily associated with errors in the numerical representation of the governing equations. A good example is the fact that uncertainties in the initial conditions give rise to stochastic errors which usually grow with increasing forecast lead times. The level of uncertainty in the forecast may be described by a model ensemble approach, but validation of ensemble results are beyond the scope of this report.

3.2 Implementation

The implementation of the validation algorithm for sea ice concentration has in part been described in subsection 3.1 above. A description with some additional details follows.

We handle both sea ice concentration observations as well as model results as representing the daily mean, so no interpolation or averaging is required in time. The differences in spatial resolution is handled by mapping each cell in the observational grid onto the coarser model grid. The average of the observational values that are mapped onto the same model grid cell is subsequently entered into the validation. Only the cells in the coarse (model) grid where both observations and model results are available, can be included in the validation algorithm.

While the model sea ice concentration values are made up of a continuous set of values in the range [0,1], the ice charts give concentration as ice belonging to one of six categories. Each of these categories represent a range of sea ice concentrations, and the present algorithm represent the classes by the ranges’ mean values.

Table 3: Representation of sea ice concentration in the validation algorithm. Note that the category *fast ice*, which is a continuous ice cover attached to a coast line, is not represented by the model.

ice class	ice chart values	model range	representation value
open water	<0.1	[0.,0.1>	0
very open drift ice	0.1, 0.2, 0.3	[0.1,0.35>	0.2
open drift ice	0.4, 0.5, 0.6	[0.35,0.625>	0.5
close drift ice	0.7, 0.8	[0.625,0.85>	0.75
very close drift ice	0.9, 1.0	[0.85,1.]	0.95
fast ice	1.0	N/A	N/A

Note that the algorithm used in the production of the ice charts assumes that the ice concentration can be represented by multiples of $1/10^{th}$ fractions. Note also that the averaging of several observational grid cells in each model cell leads to values with additional decimals. To be consistent, this average sea ice concentration observation is also mapped into the corresponding category’s representation value, as listed in Table 3. For simplicity, consider the case where only three observational grid cells is mapped into the same model cell, and assume that the three observations are represented by values 0.5, 0.5 and 0.95. The average observation then becomes

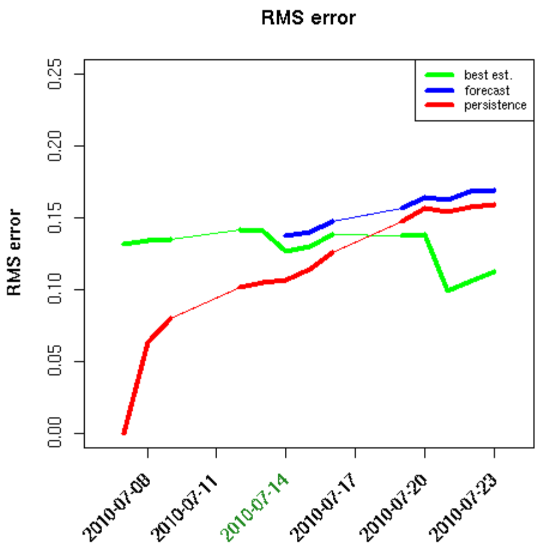


Figure 1: The rms difference between sea ice chart data and three forecasts for sea ice concentration, as a function of time. The green, blue and red lines show the rms value for the "best estimate" from the model, the standard model "forecast", and persistence of observations from the analysis date, respectively. The analysis date is the date from which observations are assimilated in the model. The date displayed in green along the bottom axis is the date at which the forecast was issued (the "bulletin date"). Dates with no observational product are bridged by thin lines.

0.65, belonging to the class *close drift ice* (since it falls within the range [0.625,0.85>). Consequently, the observed sea ice concentration is reset to 0.75 in this cell, prior to the computation of the validation metrics.

4 Validation products for sea ice concentration

The validation results are updated weekly, and validation bulletins are available from `http://myocean.met.no/ARC-MFC/Validation/SeaIceConcentration/`.

4.1 RMS

We start the validation with computing the root of the mean squared (rms) difference between the observations and the model results. This quantity is included as one of the metrics for validation of sea ice concentration due to its frequent use in this context. Hence, the rms difference is also a quantity with which most users are already familiar.

An example of results for rms difference is displayed in Figure 1. The model "forecast" and "best estimate" were defined in Section 2. Note that in the myOcean Arctic Monitoring and Forecasting Centre production at the time of writing, the assimilation is performed at a date that precedes the model run by one week.

We have chosen to display the validation of persistence of observations using the observations that entered into the assimilation initially (i.e. at the analysis time) in the present model cycle. This is a natural choice from a modeler perspective, since the initial conditions pose restrictions on any model's capability of providing an accurate forecast. However, a user is likely to be more interested in a validation of the persistence of observations available on the "bulletin date" (when the forecast is issued). Although this quantity is not displayed in Figure 1, these rms differences will be available once the validation results that are produced one week later are available.

During long periods of the year, sea ice concentration is a near-binary field with a pronounced domination of ice classes *open water* and *very close drift ice* (which includes *fast ice* in this validation). When such a distribution occurs, the rms difference takes on fairly small values as long as the model performs well in the large regions where these two ice classes are found. But a user who operates in the vicinity of the ice edge is likely to be more interested in the model's performance for the remaining ice classes. For this user, the rms difference is not an ideal metric in order to evaluate the applicability of the model results.

4.2 Ice class area

Due to the near-binary character of sea ice concentration, we need to supplement the validation with a product that is tailored to provide information about the accuracy of the model results in the marginal ice zone, which we here define as the region occupied by ice classes *very open drift ice*, *open drift ice* and *close drift ice*.

Since the ice chart data comes with a resolution defined by a set of ice classes as given in Table 3, we adopt this division as a basis for a validation metric which has a focus on the results

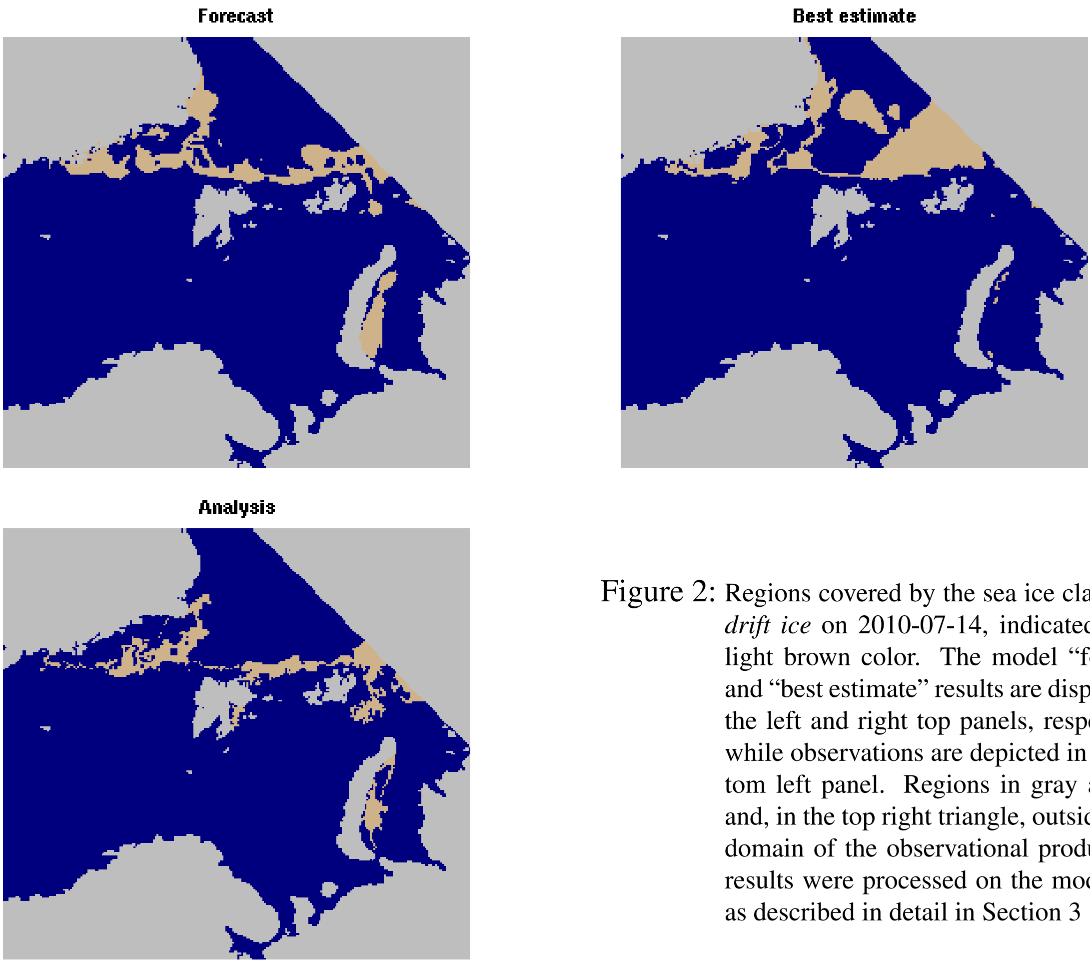


Figure 2: Regions covered by the sea ice class *close drift ice* on 2010-07-14, indicated by the light brown color. The model "forecast" and "best estimate" results are displayed in the left and right top panels, respectively, while observations are depicted in the bottom left panel. Regions in gray are land and, in the top right triangle, outside of the domain of the observational product. All results were processed on the model grid, as described in detail in Section 3

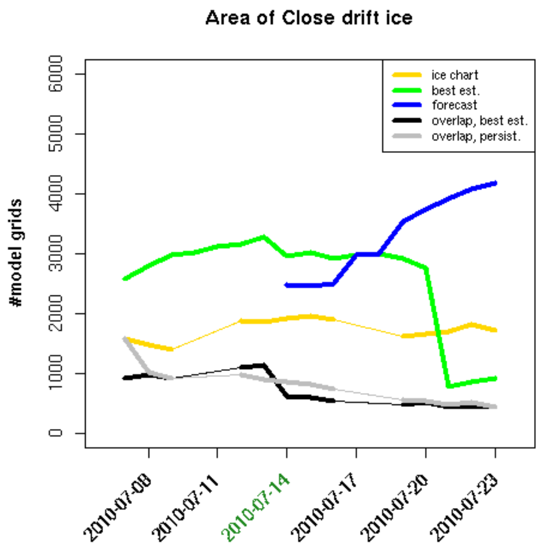


Figure 3: Area covered by the ice class *close drift ice*, as a function of time. The yellow, green and blue lines show the areal coverage in the observations, the "best estimate" from the model, and the standard model "forecast", respectively. The full black line shows the area where the ice class observations overlap with the "best estimate" product. The region where the initial observations of this ice class retains its class in the subsequent observations is indicated by the grey line. Dates with no observational product are bridged by thin lines. Note that the results for 2010-07-14 correspond to the areas displayed in the panels of Figure 2.

in the marginal ice zone: The area covered by the various ice classes.

While a validation of the area covered by an ice class provides information about the model's capability of reproducing this ice class for a given ocean – atmosphere heat flux, there is no information about the accuracy of the positioning of the ice class at hand. In order to gain information to make up for this short-coming, we also include the area in which the ice class overlaps in the observations and in the model results as a metric. Hence, in a representation of this metric like the one displayed in Figure 3, results from a "perfect model" with the "true" area in the "correct" position has all lines overlapping.

Eventually, we will add a separate validation for the sub-domain of the Barents Sea. We will also consider implementing validation of sea ice concentration based on other observational products, notably those that cover other regions of the Arctic Ocean.

Finally, we must add that another metric which provides valuable information on the model performance in the marginal ice zone, is the rms of the distance from the observed ice edge to the ice edge in the model results. (The ice edge is usually defined as the iso-line for a sea ice concentration value of 0.15.) This metric is presently (summer of 2010) not included in the validation of sea ice products from the myOcean Arctic Monitoring and Forecasting Centre. However, present plans are to implement validation of the sea ice edge position early in 2011.

Acknowledgement

myOcean is co-funded by the 7^{th} Framework Programme (FP7) of the European Union with Project Reference no. 218812 . Their support is gratefully acknowledged.