

# Improved Sea Ice Forecasting through Spatiotemporal Bias Correction

HANNAH M. DIRECTOR

*Department of Statistics, University of Washington, Seattle, Washington*

ADRIAN E. RAFTERY

*Department of Statistics, and Department of Sociology, University of Washington, Seattle, Washington*

CECILIA M. BITZ

*Department of Atmospheric Sciences, University of Washington, Seattle, Washington*

(Manuscript received 21 March 2017, in final form 18 July 2017)

## ABSTRACT

A new method, called contour shifting, is proposed for correcting the bias in forecasts of contours such as sea ice concentration above certain thresholds. Retrospective comparisons of observations and dynamical model forecasts are used to build a statistical spatiotemporal model of how predicted contours typically differ from observed contours. Forecasted contours from a dynamical model are then adjusted to correct for expected errors in their location. The statistical model changes over time to reflect the changing error patterns that result from reducing sea ice cover in the satellite era in both models and observations. For an evaluation period from 2001 to 2013, these bias-corrected forecasts are on average more accurate than the unadjusted dynamical model forecasts for all forecast months in the year at four different lead times. The total area, which is incorrectly categorized as containing sea ice or not, is reduced by  $3.3 \times 10^5 \text{ km}^2$  (or 21.3%) on average. The root-mean-square error of forecasts of total sea ice area is also reduced for all lead times.


## 1. Introduction

Dynamical models of the climate have considerable predictive capability. However, these predictions can differ systematically from observations. Models drift from their initial conditions and toward their internal mean state as a result of errors in their model physics (Meehl et al. 2014). Imperfections in obtaining or representing initial conditions also factor into these systematic differences (Collins 2002; Hazeleger et al. 2013). Bias correction has emerged as a way to correct for these errors. This class of methods develops statistical representations of the error patterns in climate models using retrospective comparisons of observations and model output. These statistical representations are then used to correct for the expected error in predictions obtained

from dynamical model forecasts (Maraun 2016; Meehl et al. 2014).

Arctic sea ice cover has decreased substantially in recent years, causing increased interest in predicting it (Comiso et al. 2008; Stroeve et al. 2012b). Dynamical models forecast quantities such as sea ice concentration, thickness, and age on a spatial grid. “Perfect model” experiments, which predict results within a climate model rather than on observed variables, have been used to evaluate the predictability of sea ice. These methods, which avoid effects from errors in model physics or initial conditions, suggest that skillful forecasts could be obtained on at least seasonal time scales (Blanchard-Wrigglesworth et al. 2011b; Day et al. 2014; Tietsche et al. 2014) with significant skill for 1–2 years (Guemas et al. 2016). However, current dynamical forecasts only show significant skill for 3–5-month lead times (Guemas et al. 2016). In the Sea Ice Outlook, where predictions for sea ice are annually compiled and compared, Stroeve et al. (2014) found prediction skill to be only marginally better than the skill obtained just from estimating the linear trend. More recently,

---

 Denotes content that is immediately available upon publication as open access.

---

*Corresponding author:* Hannah M. Director, direch@uw.edu

Blanchard-Wrigglesworth et al. (2015) found that dynamical models are also unskilled at predicting each other, suggesting varied errors in model physics and/or initial conditions across prediction systems. Current forecasts are notably affected by which (if any) bias correction technique is applied (Blanchard-Wrigglesworth et al. 2017; Fučkar et al. 2014). This suggests that improved bias correction could reduce the gap between the hypothetical predictability of sea ice and the skill currently being obtained.

A range of techniques have been proposed for bias-correcting forecasts from dynamical models. The simplest approach, typically referred to as mean bias adjustment, corrects forecasts based on the difference in the anomalies of predictions and observations (Meehl et al. 2014), sometimes computing different climatological means for each time period under consideration (García-Serrano and Doblas-Reyes 2012). Additional approaches have been introduced to account for “conditional bias,” or when the bias changes as a function of some other factor (Goddard et al. 2013). These include techniques for correcting bias that vary with forecast time (Kharin et al. 2012), lead time (van Oldenborgh et al. 2012), and initial conditions (Fučkar et al. 2014), among others. While these methods are an improvement over mean bias adjustment, as Goddard et al. (2013) note, nonlinear relationships between the model drift and forecasted value, initialization shocks, and imperfect model responses to forcings all remain challenges for designing appropriate bias correction methods. Extended logistic regression (Wilks 2009) and heteroskedastic extended logistic regression (Messner et al. 2014) have also been used to jointly bias correct and calibrate forecasts of sea ice extent (Krikken et al. 2016). While adjustments are typically applied to a summary measure or to each grid box independently, explicit spatial models of bias have also recently been suggested (Arisido et al. 2017).

For sea ice in particular, bias correction has focused on correcting summary measures, such as the total sea ice area or extent (Fučkar et al. 2014; Krikken et al. 2016). However, many current stakeholders require information about the spatial pattern of sea ice. For example, with the growth of commercial vessel traffic in some regions of the Arctic, there is increased interest in identifying safe routes for ships to follow (Huntington et al. 2015; Smith and Stephenson 2013; Stephenson et al. 2014). Forecasters who produce spatial maps of sea ice currently do not have adequate approaches to directly bias correct the predictions they receive from dynamical models. They must either ignore the known bias in the model output or make ad hoc adjustments to where the sea ice is located based only on whether bias

correction methods for the total sea ice area indicate too little or too much sea ice. This latter approach is problematic, since knowing that the total sea ice area is over or underestimated does not guarantee that any particular location will see more or less sea ice than predicted from the dynamic model. This means that it is unclear where additional sea ice should be added to the model prediction or where existing ice should be removed. Recognizing these limitations of current methods, we propose a new bias correction technique that corrects spatial error in sea ice models. This enables the production of more accurate spatial predictions than uncorrected dynamical model output.

We focus our bias correction efforts on the marginal ice zone. This is where high concentrations of ice change over to open water (Strong 2012) and where the highest amount of error in forecasting occurs (Tietsche et al. 2014). Specifically, we seek to improve the positioning of the contour surrounding contiguous grid boxes with at least 15% ice concentration, a line we refer to as the sea ice contour. The general idea of our bias correction approach is to record how far and in what direction the observed and predicted contours extend from various fixed locations. Using historical data, we can then model the difference between the observation and retrospective prediction at each location and how the difference is changing over time. This provides a bias correction that can be applied to move a new predicted contour to match an observed contour more closely. We refer to this method as contour shifting, since it captures this notion of moving the contour obtained from a dynamical model prediction to a new, more accurate location.

To illustrate how contour shifting can make a prediction more accurate, we compare in Fig. 1 a predicted contour obtained from a dynamical model and its bias-corrected version. Due to some irreducible forecast error, the bias-corrected contour is not expected to lie exactly on the observed region. Here, as in the remainder of the paper, the predictions are from the CM2.5 Forecast-Oriented Low Ocean Resolution (FLOR) model produced by the National Oceanic and Atmospheric Administration’s Geophysical Fluid Dynamics Laboratory (GFDL; Vecchi et al. 2014; Msadek et al. 2014). The observations are from the National Aeronautics and Space Administration (NASA) Bootstrap monthly sea ice concentrations obtained from the *Nimbus-7* SMMR and DMSP SSM/I–SSMIS passive microwave satellites (Comiso 2000, updated 2015).

The potential for increased skill from spatial bias correction becomes apparent when comparing where sea ice was predicted to where sea ice was observed in historical data. In Fig. 2, we map the regions of disagreement between forecasts and observations for



FIG. 1. Predictions for the 15% sea ice concentration contour obtained from the GFDL model with a 4-month lead time (red) and bias corrected using contour shifting (navy blue) for February 2012. The observed sea ice region is plotted in light blue. The black regions are part of the “nonregional ocean” in the ocean mask described in section 2a and are excluded in this analysis. The remaining white area represents grid boxes with sea ice concentration less than 15%.

February, September, and December Arctic sea ice in three successive years. In examining the regions of disagreement between observations and predictions for particular forecast months, we see some regions and months, such as Hudson Bay in December, where sea ice is repeatedly observed but never predicted. Similarly, we observe some regions where sea ice is repeatedly predicted but not observed, such as the Bering Sea in February. This suggests we could obtain more accurate forecasts by correcting persistent errors of this type.

The paper is organized as follows. Section 2 describes the implementation of contour shifting. We discuss how to statistically model the spatial pattern of error using historical comparisons of predictions and observations and how this can be used to bias correct predictions from dynamical models. In section 3 we evaluate the

reduction in error this technique provides for predicting both the spatial pattern of the sea ice and the total sea ice area. Section 4 concludes the paper with discussion.

## 2. Contour-shifting method

### a. Mapping ice sections

Contour shifting corrects the bias in predictions of the sea ice contour by relating where the sea ice edge (i.e., the 15% concentration contour) is predicted in a dynamical model to where the sea ice edge is observed. Thus, to implement contour shifting, we must develop a way to consistently represent the observed and predicted edges of sea ice. To do this, we compare how the predicted and observed sea ice edges relate to fixed reference locations and how their difference changes over time.

Although sea ice is often a broken jumble of floes, most floes accumulate in contiguous sections touching land. This makes it natural to consider how far sea ice extends off land and in what direction. We develop a mapping system accordingly. Figure 3 is an example of how we can map the boundary of predicted or observed regions of sea ice from land. Specifically, we designate fixed lines along many of the borders between land and ocean from which we can record how far sea ice extends. We represent these lines as sequences of spatial points,  $\mathbf{s}_r = (\mathbf{s}_{r,1}, \mathbf{s}_{r,2}, \dots, \mathbf{s}_{r,n_r})$ , where  $r$  denotes the region of the Arctic in which the fixed line is located. Multiple regions and corresponding fixed lines are used in this framework to obtain more detailed representations of the location of sea ice than could be obtained with a single region. Spatial points are placed on the boundary line between the ocean and the land everywhere the line intersects a corner of a grid box. In Fig. 3, these points are plotted as black circles. When connected in order, the points  $\mathbf{s}_r$  exactly trace the boundary between land and ocean. Henceforth, to simplify notation, we suppress the region subscript.

To define how sea ice extends off a fixed line in a particular region, we consider each contiguous section of sea ice that borders the fixed line. For each section, we identify all the points on the boundary of the section that intersect the fixed line. If there are  $k$  such points, we space  $k$  points evenly along the rest of the section boundary, which we denote by  $\mathbf{s}' = (\mathbf{s}'_1, \mathbf{s}'_2, \dots, \mathbf{s}'_k)$ , and refer to this as the mapped line. Specifically, we divide the section boundary line into  $k + 1$  equal-length increments and place a point at the  $k$  divisions between these increments. We then pair up the points on the mapped line with the points on the fixed line sequentially and draw a vector connecting each pair. These two-dimensional vectors, which we denote by  $\mathbf{m} = (\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n)$ , provide a

### Disagreement Between Forecast and Observation (May Initialization)

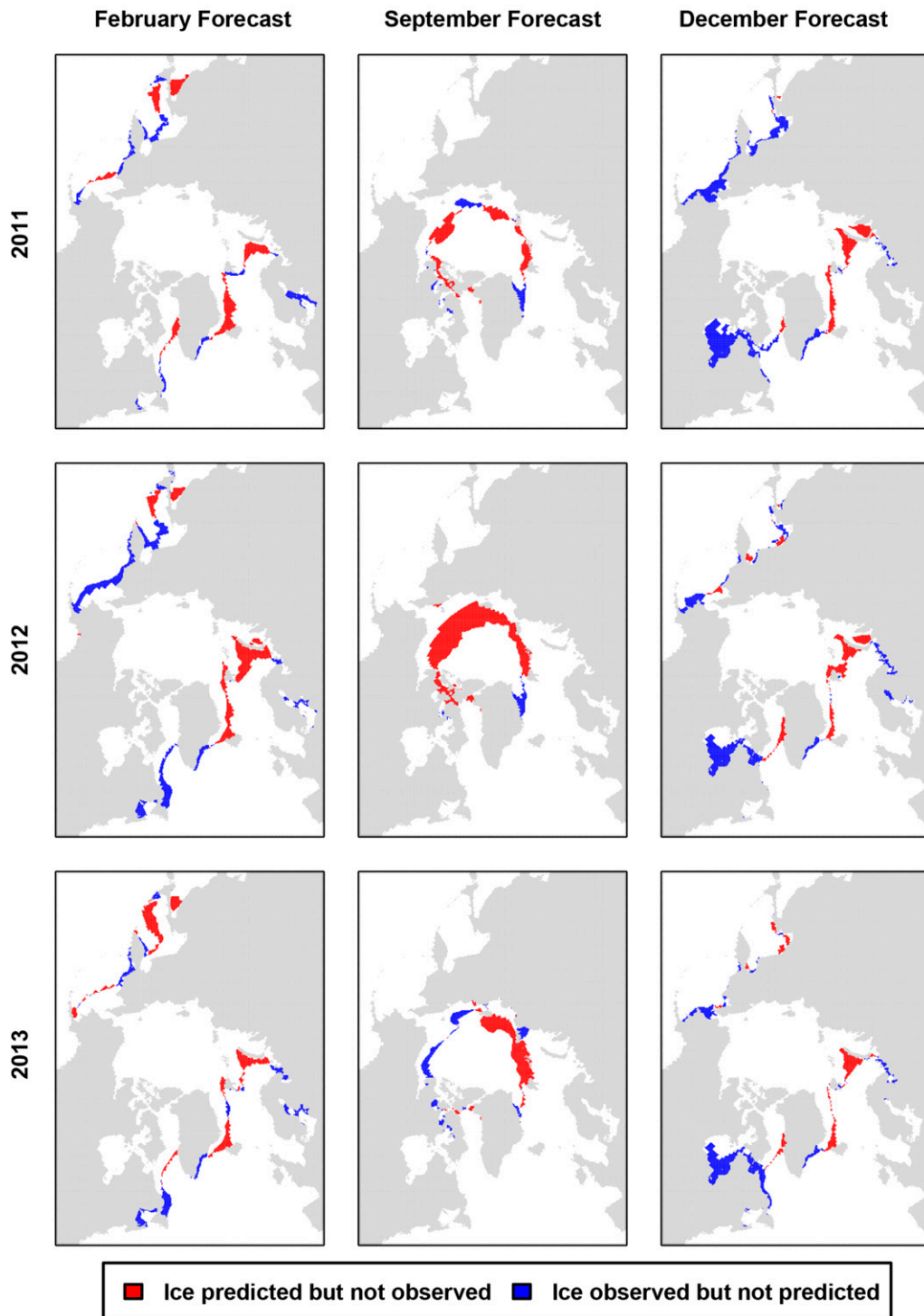


FIG. 2. Maps indicating where the prediction of the presence of sea ice, defined as grid boxes with sea ice concentration of at least 15%, differed from observations for February, September, and December from 2011 to 2013. Regions where sea ice was predicted but not observed are in red, and regions where sea ice was observed but not predicted are in blue. Predictions are from May initializations. For example, the prediction in February 2011 was initialized in May 2010, while the prediction for September and December 2011 was initialized in May 2011.

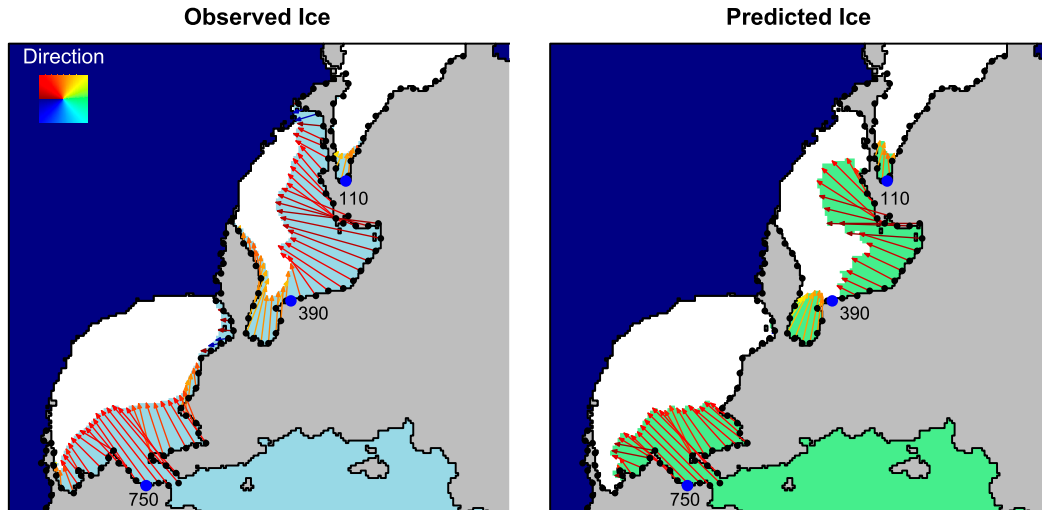


FIG. 3. Examples of mapping vectors for the Bering Sea and the Sea of Okhotsk in (left) an observation and (right) prediction for February 2008. The colored areas denote grid boxes with at least 15% sea ice concentration observed (light blue) and predicted (green). The black points indicate the fixed line on land from which the points are mapped, and the color of the arrows indicate their direction. The arrowheads point to the mapped line. Only every seventh point and line is displayed for ease of visualization. The prediction was initialized in November 2007. The corresponding time series for the three blue numbered points are given in Fig. 6.

“map” of the sea ice extending from the land border, and we refer to them as mapping vectors. They are plotted in Fig. 3 colored by direction. This technique does not constrain vectors on corresponding predictions and observations to go in the same direction.

A given point  $i$  on the mapped line can be represented as

$$\mathbf{s}'_i = \mathbf{s}_i + \mathbf{m}_i. \tag{1}$$

Sequentially connecting the points on  $\mathbf{s}$  and then on  $\mathbf{s}'$  provides a close approximation to the original boundaries of the sea ice sections. Regions and their associated fixed lines are selected to reflect where typical breaks in the sea ice occur and other physical boundaries. In Fig. 4, we show the  $R = 12$  regions outside the center of the Arctic and their fixed lines used in this analysis. These regions were obtained by modifying a region mask described in Parkinson and Cavalieri (2008) and available from the National Snow and Ice Data Center (National Snow and Ice Data Center 2017). In regions that are nearly or completely enclosed by land, we have designated the fixed line to be some portion of the land boundary. If there is a portion of the land boundary where the sea ice clearly more commonly abuts, we use this portion. Otherwise, this selection is arbitrary. In two regions, the fixed line includes a small portion of the region boundary in addition to points on land. This is done to maintain a continuous fixed line where sea ice typically abuts while respecting region boundaries.

Compared to other regions, the sea ice edge in the center of the Arctic can follow more varied paths. We therefore need a modified definition of how to map this region. Our approach is illustrated with a sample map in Fig. 5. We first designate an additional region for the center of the Arctic, indexed as  $R + 1$ . Then, we create an artificial set of reference lines and record where the sea ice intersects them. Specifically, we define an arbitrary point  $c$  near the middle of this central Arctic region and then extend lines from  $c$  to the  $n_{R+1}$  points on the region boundary that intersect the lattice of the spatial grid. In Fig. 5,  $c$  is indicated with a black point. We record the farthest point from  $c$  within the boundary of the sea ice in the region that intersects each line  $i$  as  $\mathbf{s}'_{R+1,i}$ . We let  $\mathbf{s}_{R+1,i} = c$  for all  $i$ . As for other regions, we can then build a mapping vector  $\mathbf{m}_{R+1,i}$  to connect  $\mathbf{s}_{R+1,i}$  and  $\mathbf{s}'_{R+1,i}$  and write

$$\mathbf{s}'_{R+1,i} = \mathbf{s}_{R+1,i} + \mathbf{m}_{R+1,i}. \tag{2}$$

The resulting mapping vectors are shown in Fig. 5 colored by their direction. Unlike in other regions, the angles of the vectors in the central Arctic are fixed and the same for observations and predictions. Connecting the points of  $\mathbf{s}_{R+1,i}$  in order approximately maps out the edge of the sea ice.

Completing this process for all  $R + 1$  regions for a single observation or prediction gives  $R + 1$  sets of coordinates of fixed lines and associated mapping vectors. The vectors will vary for different observations and predictions, but the coordinates will be the same for any observation or



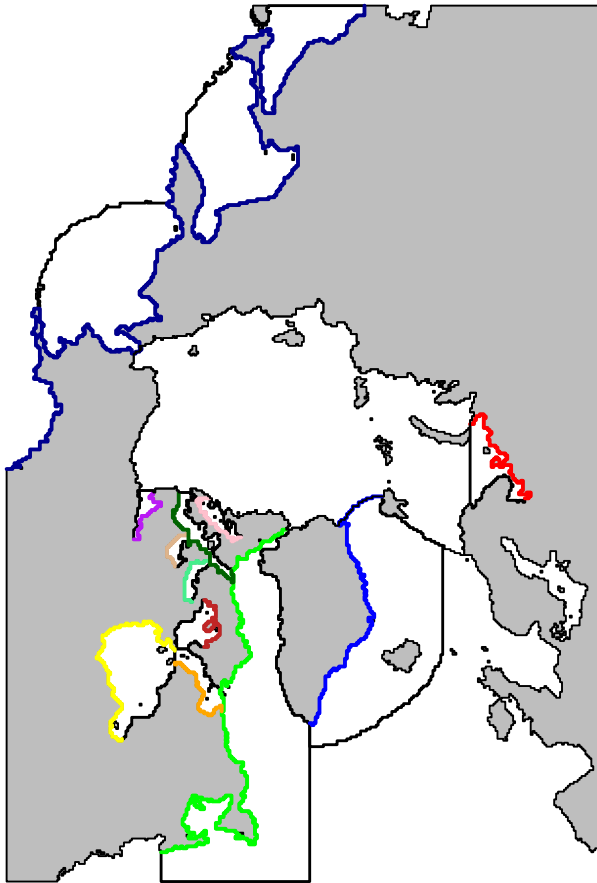


FIG. 4. Map of the fixed lines from which regions will be built. Region boundaries over ocean are in black, and each colored line within a region is the fixed line from which ice will be mapped. The large region in the central Arctic does not have a fixed line as is discussed in Section 2a.

prediction. This means we can develop a detailed spatial picture of how a predicted sea ice edge compares to an observed sea ice edge. We need only evaluate the differences in the mapping vectors along each point on the fixed lines. Repeating this process for a collection of observations and predictions enables relating predictions and observations to one another over time and space. For example, we can conclude that a prediction creates too much sea ice in a particular location if the vectors associated with predictions in that location tend to be longer than those associated with the observations. Similarly, we can assess if sea ice cover is decreasing over time in a region if the vectors associated with coordinates in this region typically get shorter over time.

#### b. Estimating the bias

Now that we have established a way to represent the sea ice edge, we can turn to the issue of estimating bias. With this mapping framework, increasing the accuracy

of the sea ice edge can be achieved by reducing the error in the forecasts of the locations of the observed mapped points  $\mathbf{s}'_{i,\text{obs}}$  for all locations  $i$ . At the time a forecast is issued any  $\mathbf{s}'_{i,\text{obs}}$  is unknown and must be predicted for each  $i$ . We denote a prediction of this quantity by  $\hat{\mathbf{s}}'_{i,\text{obs}}$ . This follows the statistical convention of using the hat symbol above an unknown quantity to indicate a prediction or estimate of that quantity. We denote the mapped point obtained from the prediction from the dynamical model at location  $i$  by  $\mathbf{s}'_{i,\text{pred}}$ . While this quantity is a prediction, the hat notation is not used, since  $\mathbf{s}'_{i,\text{pred}}$  is known at the time the forecast is issued. With this notation, we can now proceed to bias correction. We will use  $\mathbf{s}'_{i,\text{pred}}$  and a model for how  $\mathbf{s}'_{i,\text{pred}}$  is expected to differ systematically from  $\mathbf{s}'_{i,\text{obs}}$  to obtain a bias-corrected forecast  $\hat{\mathbf{s}}'_{i,\text{obs}}$  for  $\mathbf{s}'_{i,\text{obs}}$ .

To do this, we first need to isolate the systematic differences between observations and predictions in the mapped lines. The observed mapped point at location  $i$ ,  $\mathbf{s}'_{i,\text{obs}}$ , can be expressed as a function of both observations and predictions. Beginning with Eq. (1), we can write the observed mapped point at any location  $i$  as

$$\mathbf{s}'_{i,\text{obs}} = \mathbf{s}_i + \mathbf{m}_{i,\text{obs}} = \mathbf{s}_i + \mathbf{m}_{i,\text{pred}} + (\mathbf{m}_{i,\text{obs}} - \mathbf{m}_{i,\text{pred}}). \quad (3)$$

Further,  $\mathbf{s}'_{i,\text{pred}} = \mathbf{s}_i + \mathbf{m}_{i,\text{pred}}$  by definition, so we have

$$\mathbf{s}'_{i,\text{obs}} = \mathbf{s}'_{i,\text{pred}} + (\mathbf{m}_{i,\text{obs}} - \mathbf{m}_{i,\text{pred}}). \quad (4)$$

In other words the observed mapped vector at location  $i$  can be written as a sum of the predicted mapped vector at location  $i$  and the difference between the observed and predicted mapping vectors,  $\mathbf{m}_{i,\text{obs}} - \mathbf{m}_{i,\text{pred}}$ . This difference comes from both systematic discrepancies between the prediction and observation and irreducible error. We can decompose this error as

$$\mathbf{m}_{i,\text{obs}} - \mathbf{m}_{i,\text{pred}} = \mathbb{E}[\mathbf{m}_{i,\text{obs}} - \mathbf{m}_{i,\text{pred}}] + \boldsymbol{\varepsilon}_i, \quad (5)$$

where  $\mathbb{E}[\mathbf{m}_{i,\text{obs}} - \mathbf{m}_{i,\text{pred}}]$  denotes the expected systematic difference between the prediction and observation and  $\boldsymbol{\varepsilon}_i$  represents the irreducible error, which we assume to be random with mean zero.

We now focus on modeling and correcting the systematic component of the error. Our goal is to obtain an unbiased estimator for  $\mathbf{s}_{i,\text{obs}}$ . In general, an unbiased estimator,  $\hat{\theta}$ , is a function of data or other covariates that is used to predict a quantity  $\theta$  and that on average has the same value as  $\theta$ ; that is,  $\mathbb{E}(\hat{\theta}) = \theta$ . Assuming we can obtain an unbiased estimator for  $\mathbb{E}[\mathbf{m}_{i,\text{obs}} - \mathbf{m}_{i,\text{pred}}]$ , denoted by  $\hat{\mathbb{E}}[\mathbf{m}_{i,\text{obs}} - \mathbf{m}_{i,\text{pred}}]$ , we have that

$$\hat{\mathbf{s}}'_{i,\text{obs}} = \mathbf{s}_i + \mathbf{m}_{i,\text{pred}} + \hat{\mathbb{E}}[\mathbf{m}_{i,\text{obs}} - \mathbf{m}_{i,\text{pred}}] \quad (6)$$

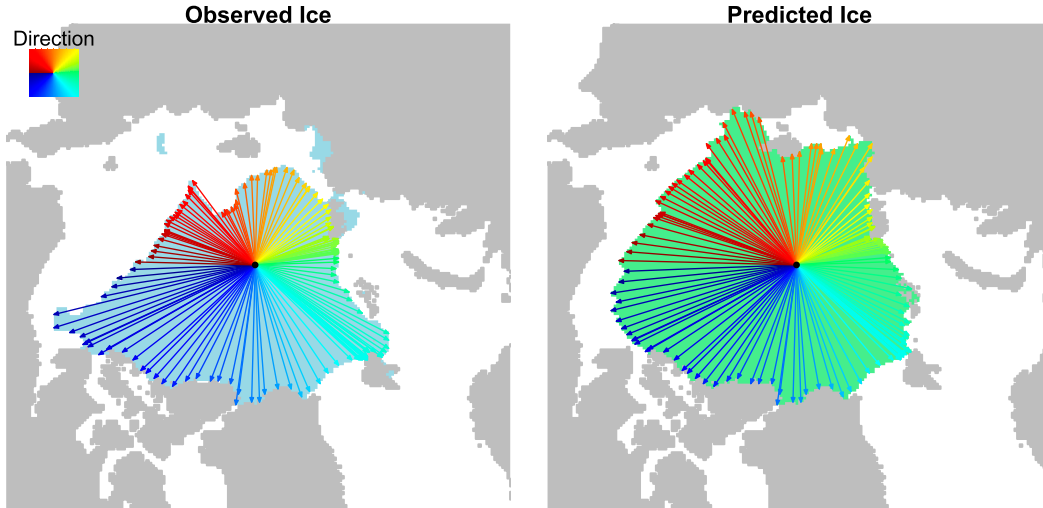


FIG. 5. Sample of mappings in the central Arctic region for September 2008 (left) for the observation and (right) for the prediction with lead times of 4 months. The directions of the colored lines emanating from a central point in the region, indicated with a black circle, are fixed for all predictions and observations. On each line, the farthest point from the center of the region that intersects the sea ice edge is recorded. These points are those that the arrowheads touch. The line connecting these points forms an approximation to the boundary of the sea ice. Only every seventh line is plotted for ease of visualization.

is an unbiased estimator for  $\mathbf{s}'_{i,obs}$ . We can show this by taking the expectation of the estimator:

$$\begin{aligned} \mathbb{E}[\mathbf{s}'_{i,obs}] &= \mathbb{E}[\mathbf{s}_i + \mathbf{m}_{i,pred} + \widehat{\mathbb{E}}[\mathbf{m}_{i,obs} - \mathbf{m}_{i,pred}]] \\ &= \mathbf{s}_i + \mathbf{m}_{i,pred} + \mathbf{m}_{i,obs} - \mathbf{m}_{i,pred} = \mathbf{s}_i + \mathbf{m}_{i,obs} = \mathbf{s}'_{i,obs}. \end{aligned} \tag{7}$$

This means that on average the predicted position of the  $i$ th mapping point will be the same as the  $i$ th mapping point that will be observed. This is in contrast to  $\mathbf{s}'_{i,pred}$ , which on average will be at a position  $\mathbb{E}[\mathbf{m}_{i,obs} - \mathbf{m}_{i,pred}]$  units away from  $\mathbf{s}'_{i,obs}$ . This error decomposition and correction has some similarities to what was proposed by Kharin et al. (2012) but provides results for many spatial locations rather than for a single summary measure.

The derivation in the previous paragraph presupposes that we have an unbiased estimator of  $\mathbb{E}[\mathbf{m}_{i,obs} - \mathbf{m}_{i,pred}]$ . Obtaining an approximately unbiased estimator is feasible with sufficient historical data and retrospective forecasts. We can map the observed and predicted sea ice regions using the technique described in section 2a. If the dynamical model bias was fixed over time, we could then estimate the bias at each location simply as the average difference between the observed and predicted mapping vectors for each year in the dataset. However, this approach fails to account for the rapid change occurring in the Arctic.

Both model predictions and observations show a reduction in sea ice over time, but the rates of decline are not the same, both in total and for individual regions. For

total sea ice extent, Stroeve et al. (2012a) found that the trend in most prediction systems is less than observed. To account for this, we extend the method proposed by Kharin et al. (2012) for bias-correcting total sea ice area to spatial contours. In their approach, the total sea ice area in observations and predictions are each regressed separately on time. Then, the regression models are used to estimate the predicted and observed sea ice area at the forecast time. The difference between these quantities is used as an estimate of the bias at the forecast time.

Extending this approach to our spatial setup, we estimate the bias at each point on the fixed line using regression. Outside the central Arctic, we regress each of the  $x$  and  $y$  components of the mapping vectors for the observation and the prediction separately on time. For each point  $i$  on the mapped line, this gives four regression equations of the form,

$$m_i = \hat{\alpha}_i + \hat{\beta}_i t + \varepsilon_i, \tag{8}$$

for the observed  $x$  component, observed  $y$  component, predicted  $x$  component, and predicted  $y$  component. The  $\varepsilon_i$  are error terms. Consequently, the predicted bias in the mapping line at location  $i$  and time  $t$  is

$$\widehat{\mathbb{E}}[\mathbf{m}_{i,obs} - \mathbf{m}_{i,pred}] = \hat{\mathbf{m}}_{i,obs} - \hat{\mathbf{m}}_{i,pred}, \tag{9}$$

where each  $\mathbf{m}$  and  $\hat{\mathbf{m}}$  are vectors with an  $x$  and  $y$  component. The terms  $\hat{\mathbf{m}}$  are the same as in Eq. (8), except with the error terms removed. To estimate the regression

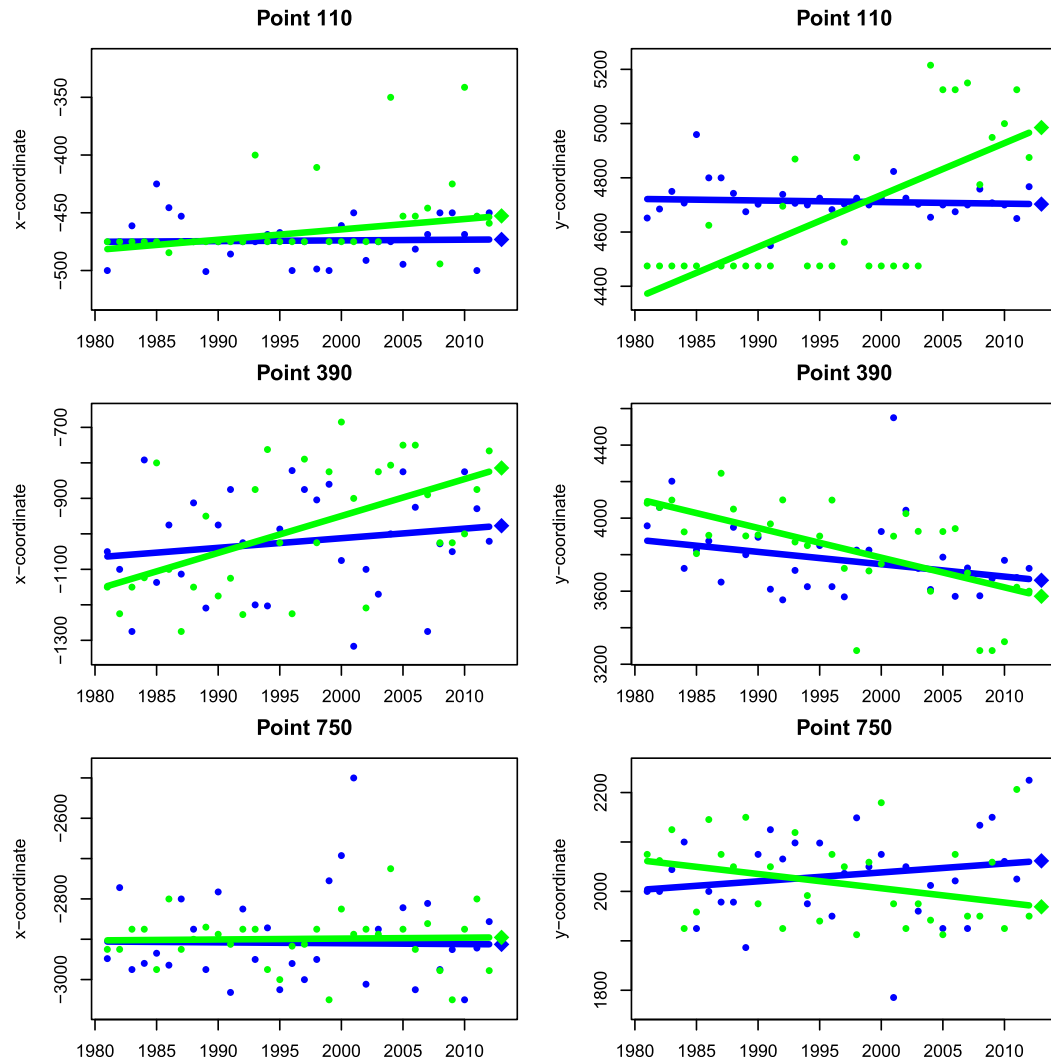


FIG. 6. Time series for the three sample points in Fig. 3 used to predict the bias in February 2012 at a lead time of 4 months: (left)  $x$  coordinate and (right)  $y$  coordinate. Circles are recorded data and diamonds are the prediction for 2012 obtained via Huber M-estimation, a form of robust regression. Blue indicates observation and green indicates prediction. The expected bias is the difference between the two diamonds.

parameters, we use Huber M-estimation, a form of robust regression (Huber and Ronchetti 2011). Implemented in R using the *rlm* function, this technique minimizes a function of the residuals that puts reduced weight on outlying values (Venables and Ripley 2002). This is appropriate in this context, since we do not want a few unusual years to have undue influence on the trend lines. In Fig. 6, we plot the  $x$  and  $y$  components of several observations and predictions along with regression lines for the sample points highlighted in Fig. 3.

In the central Arctic, the mapped points are constrained to stay on specific lines, so the mapping vectors can be represented by just their lengths rather than by separate  $x$  and  $y$  components. Consequently, for region  $R + 1$ , we let

$$\begin{aligned} \widehat{\mathbb{E}}[m_{i,\text{obs}} - m_{i,\text{pred}}] &= \hat{m}_{i,\text{obs}} - \hat{m}_{i,\text{pred}} \\ &= (\hat{\alpha}_{i,\text{obs}} + \hat{\beta}_{i,\text{obs}}t) - (\hat{\alpha}_{i,\text{pred}} + \hat{\beta}_{i,\text{pred}}t), \end{aligned}$$

where each  $m$  is now a single length rather than a vector. This reduces the number of parameters that need to be estimated but otherwise does not affect the bias correction procedure.

### c. Finalizing a prediction

To finalize a prediction, we replace the mapped points with the bias-corrected mapped points and connect them in the same way as in section 2a. Any sections of sea ice predicted that do not touch land outside the



## Getting to the Final Contour

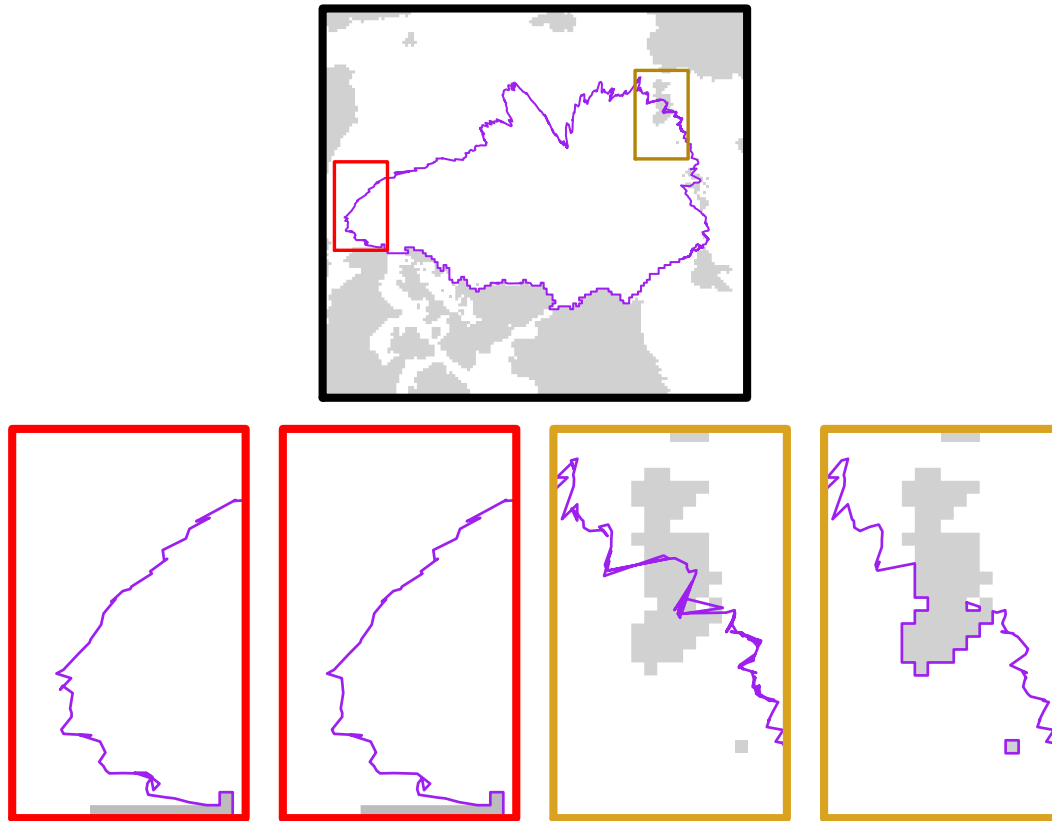


FIG. 7. Example of modifying a bias-corrected contour when the contour line crosses over itself and/or passes through land. (top) A sample contour. The red box highlights a region where the contour (purple) crosses over itself and the yellow box highlights a region where the contour goes through land. (bottom) These sections are zoomed in on, showing (left) the original contour and (right) the contour after correction. All issues have been corrected by removing land regions and applying the algorithm described in [section 2c](#).

central Arctic region are not bias corrected; that is, they are left as predicted by the dynamical model. In the central Arctic region, sea ice sections that are not part of the largest contiguous region are also left unadjusted.

Some issues with the bias-corrected contour remain. In some cases the resulting contour line will intersect itself, making it an invalid boundary for a contiguous region. To remove these self-intersections, we simplify the sections of the contour surrounding self-intersections with the Douglas–Peucker algorithm ([Douglas and Peucker 1973](#)). This algorithm takes in a line, represented as a sequence of connected points, and outputs a new line that resembles the original line. The new line uses as few points as possible while ensuring that the new line differs by no more than a user-specified tolerance  $\varepsilon$  from the original line. For each section of the contour that contains a self-intersection, we replace it with a similar line obtained with the Douglas–Peucker algorithm.

The tolerance  $\varepsilon$  is determined by an automated process. The Douglas–Peucker algorithm with a low  $\varepsilon$  is initially applied to the contour and a check is performed to determine if self-intersections remain. If there are self-intersections,  $\varepsilon$  is increased and the Douglas–Peucker algorithm is run again. This process is repeated until an  $\varepsilon$  is found that is large enough to give a contour with no self-intersections. In our analysis, we set the initial  $\varepsilon$  to 0.25 on the nominal 25-km grid and increase it by 0.25 as needed.

In this way, we create a final bias-corrected contour that does not contain any self-intersections. It is also possible that a bias-corrected prediction will go through land. When this is the case, we simply remove the land contained within the region enclosed by the contour, so that the contour line goes through the interior edge of the land region. Examples of both of these adjustments to a contour are given in [Fig. 7](#). Finally, sections of sea ice that have been expanded such that they now overlap are merged into a single section of sea ice.

#### d. Implementation and software

We have made our implementation of contour shifting available in the IceCast R package (Director et al. 2017). The software package includes a one-line automated execution function that takes in observations and prediction and outputs bias-corrected predictions. Additionally, all intermediate functions and their code are accessible within the package. Interested users can use these intermediate functions to gain more flexibility in how contour shifting is implemented or to extend or modify this method. A vignette, or tutorial within the package, discusses the IceCast package in more detail and provides examples of its use.

Within the package two main functions are used to execute contour shifting. The first function maps observations and predictions, and the second estimates the bias and creates the new prediction based on results from the first function. On a typical laptop, mapping the prediction and observation for a single year takes around a minute. Given the mappings, estimating the bias correction and creating a new prediction for a single year takes about a minute. In other words, computation time to bias correct a single year of predictions given  $N$  years of observations and predictions is roughly  $N + 1$  min. However, if multiple years are bias corrected, the computation time per year is notably reduced since the mappings for each previous year used to learn the bias need to be calculated only once. There is some variability in computation time depending on the forecast month. Months with less sea ice tend to run faster than average, since there are fewer regions with sea ice to map and bias correct. The opposite is true for months with more sea ice. While contour shifting does have greater computational cost than existing methods for bias correction of the total sea ice area or extent, it provides substantially more information about where sea ice is expected to be located. Further, relative to the cost of running a full dynamical forecasting system, the computation time is trivial.

### 3. Results

#### a. Experimental setup

To evaluate this bias correction approach, we consider retrospective predictions of monthly sea ice concentration from the CM2.5 FLOR model produced by the National Oceanic and Atmospheric Administration's GFDL from 1981 to 2013 (Vecchi et al. 2014; Msadek et al. 2014). This fully coupled global climate model has approximate atmospheric resolution of  $50 \text{ km} \times 50 \text{ km}$  and approximate ocean resolution of  $1^\circ \times 1^\circ$ . This prediction system has 12 ensemble members associated with it. The predictions were downloaded from the Earth

System Grid (National Center for Atmospheric Research 2017) and have been converted from their native grid to the nominal 25-km polar stereographic grid used by NASA Bootstrap using weights obtained from the spherical coordinate remapping and interpolation package (Jones 1997). For total sea ice, this model is known to be of high quality for forecasting the mean state (Msadek et al. 2014). However, this does not mean that there is not spatial bias, since for total sea ice estimates overestimation of sea ice in one region can compensate for underestimation of sea ice in another region. This type of error, referred to as misplacement error, has been shown to form a substantial fraction of the total error in many model predictions (Goessling et al. 2016).

For our analysis, we consider monthly forecasts that have been initialized on the first day of the month and run for a year. This gives 12 initialization times per year. To obtain the prediction for the 15% sea ice edge corresponding to the ensemble forecast, we average the concentration in each grid box over the ensemble members to create a field of the average concentration. Then, we find the 15% sea ice edge for this averaged field.

To assess contour shifting's effect on prediction accuracy, we apply it to forecasts for all 12 months at four different lead times. We evaluate all forecast months because sea ice has varying behavior and potential predictability during different parts of the year, such as differing persistence by season (Blanchard-Wrigglesworth et al. 2011a). Since forecasts are for monthlong periods, there can be some ambiguity in the definition of lead time. We report lead times rounding up to the nearest whole month. For example, a prediction for the monthly sea ice concentration for February made on 1 January is described as being made at a lead time of 2 months. For any prediction, we use all data prior to the forecast year to estimate the bias correction. Results are reported for 2001–13. This corresponds to years where there were at least 20 years of data available prior to the forecast year from which to build the bias model.

Predictions from this model are compared to observations of the monthly sea ice concentration obtained from the NASA satellites *Nimbus-7* SMMR and DMSP SSM/I-SSMIS. The observations have been processed by the bootstrap algorithm and are distributed by the National Snow and Ice Data Center (Comiso 2000). The original land mask for the observations has been replaced with the simpler land mask (derived from approximately  $1^\circ$ ) used in the predictions. We define a grid box to contain sea ice if it has at least 15% sea ice concentration.

#### b. Spatial error reduction

In Fig. 8, we map the initial GFDL prediction and the bias-corrected forecast of the sea ice contour on top of

## Predicted vs. Bias-Corrected Contours (3.5 Month Lead Time)

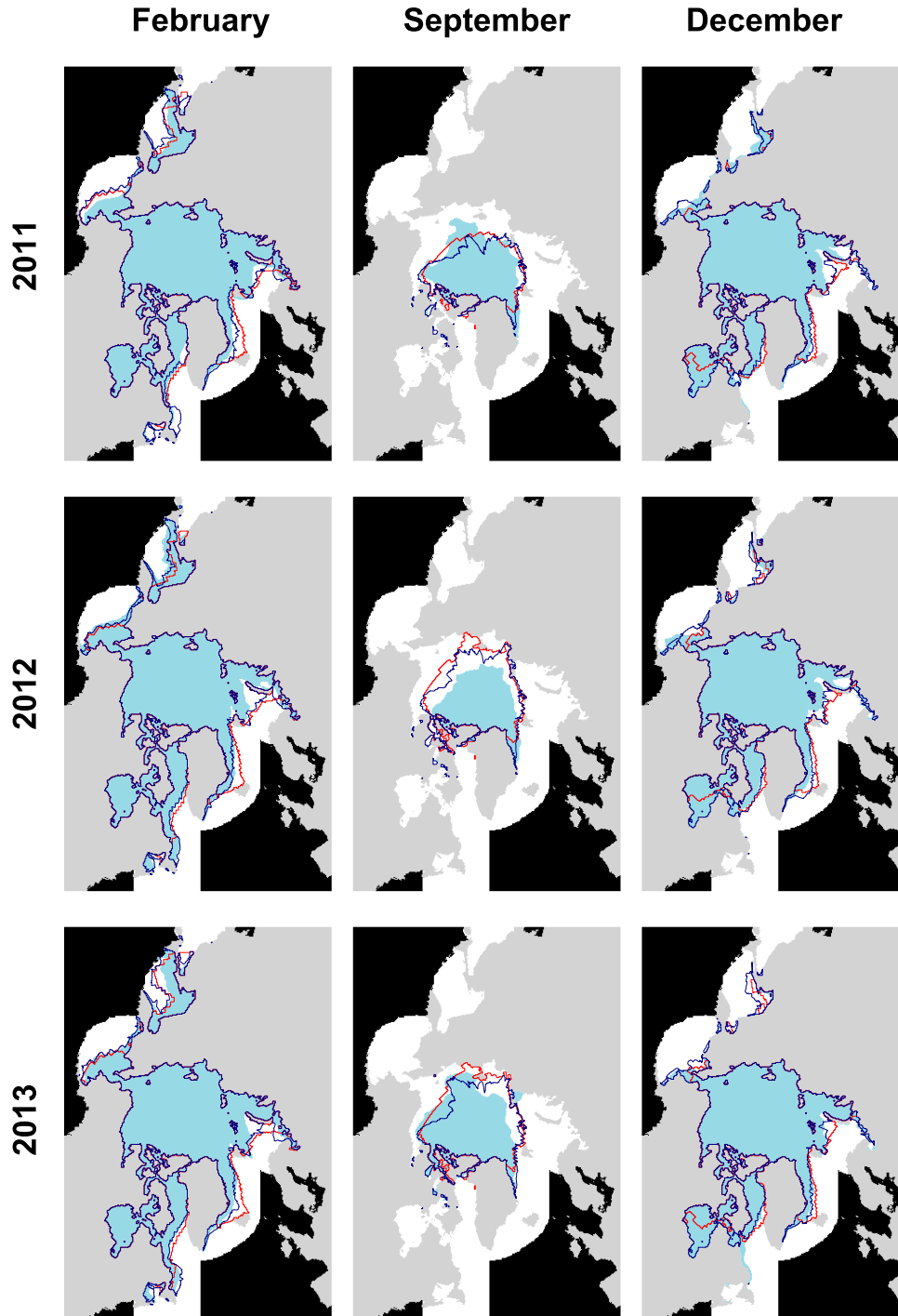


FIG. 8. Predictions for the 15% sea ice concentration contour obtained from the GFDL model with a 4-month lead time (red) and bias-corrected using contour shifting (navy blue) for February, September, and December from 2011 to 2013. The observed ice region is plotted in light blue. The black regions are part of the “nonregional ocean” in the ocean mask described in [section 2a](#) and are excluded in this analysis. The remaining white area represents grid boxes with sea ice concentration less than 15%.

	2 Months	4 Months	7 Months	12 Months	Mean
Jan	4.6	5.3	4.4	3.8	4.5
Feb	4.6	5.0	5.1	5.1	4.9
Mar	2.9	3.6	3.9	3.7	3.5
Apr	3.6	3.5	2.9	2.6	3.1
May	3.2	4.1	5.0	3.9	4.0
June	0.6	1.2	1.7	1.8	1.3
Jul	2.0	1.4	1.6	2.3	1.8
Aug	2.5	2.5	1.7	1.4	2.0
Sep	1.5	2.5	1.9	0.9	1.7
Oct	2.4	1.7	2.1	4.1	2.6
Nov	2.2	2.2	2.1	2.6	2.3
Dec	7.5	7.6	7.8	8.9	8.0
Mean	3.1	3.4	3.4	3.4	3.3

□ (-0.8, 0.8]
□ (0.8, 2.4]
□ (2.4, 4]
■ (4, 5.6]
■ (5.6, 7.2]
■ (7.2, 8.9]

FIG. 9. The reduction in the mean IIEE ( $10^5 \text{ km}^2$ ) over the test years 2001–13 for bias-corrected predictions compared to the original GFDL predictions at four different lead times. For each year, the IIEE is defined to be the sum of all areas where sea ice is predicted but not observed or where sea ice is observed but not predicted. Results are rounded to one decimal place.

the observed sea ice for several example months and years. In many cases, the bias-corrected contour clearly follows the observed sea ice boundary more closely than the uncorrected dynamical model prediction does, indicating that bias correction is improving the forecast.

To quantitatively assess the improvement due to contour shifting, we use the integrated ice-edge error (IIEE) measure proposed by Goessling et al. (2016). The IIEE is an aggregate measure of the amount of error in a prediction. It is obtained by adding up the areas where ice is predicted but not observed and the areas where ice is observed but not predicted. For all months and lead times, we see a reduction in the average IIEE after bias correction has been applied. In Fig. 9, we report the mean error reduction of this measure for all forecasts months at four different lead times. Figure 10 reports the mean percent change in the IIEE. For all months and lead times, we obtain an average reduction in the IIEE of  $3.3 \times 10^5 \text{ km}^2$ , or 21.3%. In Fig. 11, we plot the average IIEE for the years 2001–13 organized by forecast month for the unadjusted GFDL forecasts and for the bias-corrected forecasts. From this we can see the increased prediction skill that bias correction provides for each month. For example, for March, we see that the average IIEE for bias-corrected forecasts at all lead times is less than the lowest average IIEE for the unadjusted forecasts. This means that a bias-corrected forecast issued 12 months in advance is more accurate on average than an unadjusted forecast issued 2 months in advance. Similarly, the average IIEE for unadjusted September forecasts issued 2 months in advance is  $13.8 \times 10^5 \text{ km}^2$ . This is greater than the average IIEE for the bias-corrected forecast at a 2-month lead time ( $12.3 \times 10^5 \text{ km}^2$ ) and a 4-month lead time ( $13.4 \times 10^5 \text{ km}^2$ ). This indicates that on average for September a

	2 Months	4 Months	7 Months	12 Months	Mean
Jan	31.8	35.2	30.0	25.9	30.7
Feb	29.0	29.7	31.0	30.7	30.1
Mar	19.9	23.2	23.0	21.8	22.0
Apr	23.7	23.8	20.0	16.6	21.0
May	25.0	32.2	35.0	27.6	29.9
June	4.5	8.7	13.4	14.5	10.3
Jul	11.7	7.7	8.1	11.2	9.7
Aug	14.2	14.4	10.4	8.6	11.9
Sep	10.2	14.4	9.3	5.2	9.8
Oct	15.5	13.9	15.4	26.1	17.7
Nov	16.2	15.4	15.3	19.8	16.7
Dec	44.1	45.3	45.6	49.0	46.0
Mean	20.5	22.0	21.4	21.4	21.3

□ (-5, 5]
□ (5, 15]
■ (15, 25]
■ (25, 35]
■ (35, 45]
■ (45, 55]

FIG. 10. As in Fig. 9, except with the percent change in mean error area. Results are again rounded to one decimal place.

bias-corrected forecast issued 4 months in advance is more accurate than an unadjusted forecast issued both 2 and 4 months in advance. In summary, bias-correcting contours obtained from dynamical sea ice forecasts with contour shifting leads to more accurate forecasts of the sea ice contours. Again, we note that a portion of the IIEE is from irreducible forecast error.

To compare performance across months and lead times, we look at the percent error reduction results in Fig. 10. We note that the amount of error reduction in a given month is not directly related to the IIEE in the original forecast. Rather, the amount of error that can be corrected is a function of how much systematic error a prediction has. Contour shifting will provide the most error reduction when there is a notable amount of error that follows consistent patterns. Contour shifting will provide less error reduction when the unadjusted forecast is quite accurate or its error is highly variable.

The largest percent-error reductions occur for forecasts in the winter and in May. These months have large, consistent errors across years. For example, the month of December has a high amount of error overall and the errors are repeated across years. As can be seen in Fig. 8, underestimation of sea ice in Hudson Bay and overestimation of sea ice off the coast of Greenland north of Iceland and in the Barents Sea occur in most years. Thus these errors can be well corrected by contour shifting.

On the other hand, the smallest percent-error reductions occur from June to September. The adjusted IIEE values are only marginally better than the unadjusted IIEE values. This suggests that random, rather than systemic, errors dominate during these months and/or that our bias correction method is not identifying systematic errors. Mechanistically, this is plausible. In early summer, the heterogeneous surface of the Arctic is especially affected by the ice-albedo feedback (Blanchard-Wrigglesworth et al. 2011a) and in late summer and early fall there is the highest proportion of thin ice, which leads to more

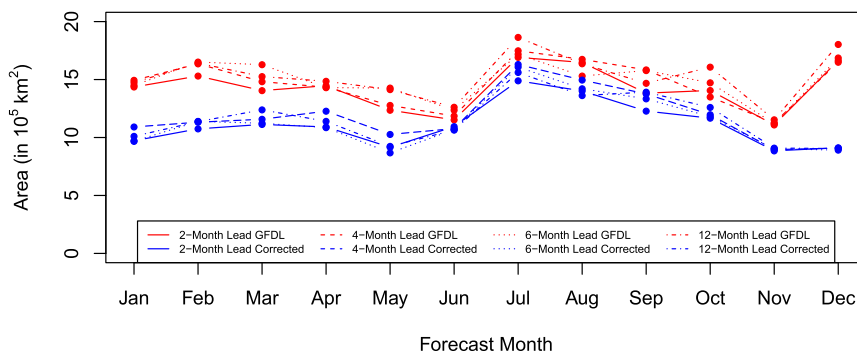


FIG. 11. Average IIEE for the years 2001–13 by month for the unadjusted GFDL model output (red) and bias corrected with contour shifting (blue). The different line types indicate the lead time.

variability (Holland et al. 2011). Variability makes it less likely that a consistent discrepancy between models and observations would occur and makes it so that more data would be needed to detect any bias that is there. For example, looking at consecutive September observation and predictions in the second column of Fig. 8, the only repeated error pattern is a small underestimation of sea ice off the northeastern part of Greenland. The interior of the Arctic Ocean is variable and not substantially affected by bias correction.

No clear patterns emerge in performance across lead times, suggesting that contour shifting can be applied to obtain more accurate forecasts for lead times up to at least a year. We also find little effect of lead time on IIEE even after bias correction. This suggests that initial conditions strongly affect forecast accuracy at these time scales. This result aligns with that of Msadek et al. (2014) for this dynamical model, who infer that initial conditions are a primary factor in predicting sea ice extent at lead times up to a year.

### c. Total sea ice area

While the key contribution of contour shifting is the ability to bias correct a contour, this technique can also be employed to bias correct the total sea ice area. Sea ice area refers to the total area of grid boxes with sea ice concentration of at least 15%. To obtain a bias-corrected estimate of the total sea ice area with contour shifting, one simply needs to sum the areas contained within the bias-corrected contour. This approach gives comparable performance to existing bias correction techniques focused exclusively on correcting total sea ice area estimates.

To illustrate this, we implement the trend bias correction technique introduced in Kharin et al. (2012) and the initial condition technique introduced in Fučkar et al. (2014). The approach of Kharin et al. (2012)

requires regressing observed total sea ice area on time and regressing predicted total sea ice area on time. The difference in the regression lines at the forecast time is then used to estimate and correct for the expected difference in total sea ice area between the observation and the prediction at the forecast time. The initial condition approach is similar, except that instead of regressing on time, the observed and predicted total sea ice areas are regressed on information describing their initial conditions. In our implementation, we regress on the observed monthly sea ice areas for the month prior to initialization.

In Fig. 12, we plot the root-mean-square error (RMSE) for predictions of the total sea ice area obtained versus the forecast month for four lead times. We do this for the GFDL model alone and with the various bias correction techniques. In Table 1, we also summarize this information by lead time. We found RMSE to be lower for the bias-corrected forecasts than for the unadjusted GFDL predictions for all the bias correction techniques, highlighting the importance of bias correction for obtaining accurate sea ice forecasts. For all lead times, contour shifting had lower RMSE than all the other methods considered. Contour shifting performed best around the time when sea ice reached its minimum. This differs from the Kharin et al. (2012) technique, which performed best around June. Overall this analysis shows that contour shifting not only improves the prediction of the sea ice contour but also leads to more accurate estimates of the total sea ice area.

## 4. Discussion

We have proposed and implemented contour shifting, a novel technique for bias correcting a prediction of a contour of sea ice at a particular concentration, a typical definition of the sea ice edge. Retrospective comparisons



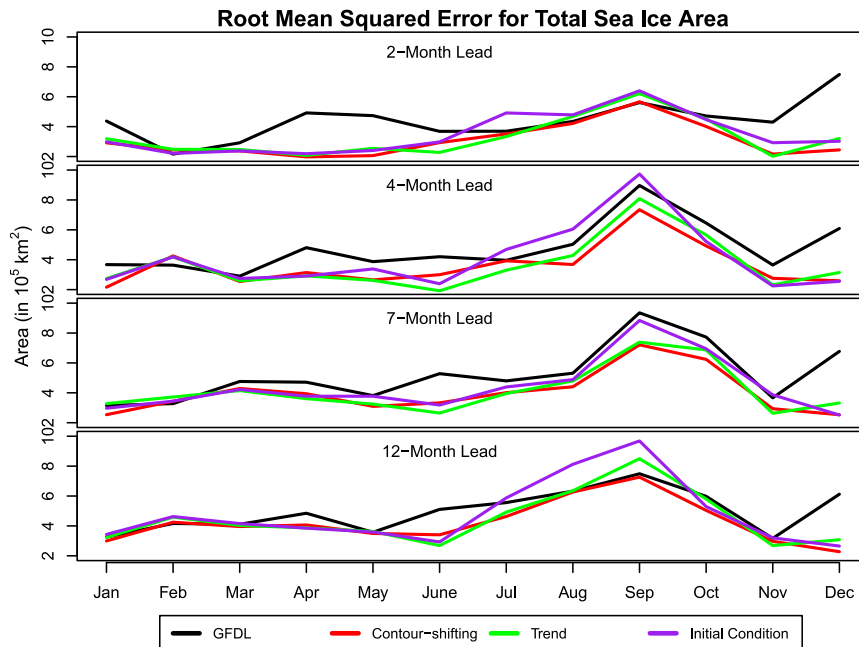


FIG. 12. The RMSE for the total sea ice area prediction at four different lead times using four different methods. The black line (GFDL) is for the unadjusted dynamical model. The other lines are for bias-corrected versions of these forecasts. The red line (contour shifting) is for the bias correction method described in this paper. The green line (trend) is for the bias correction method given in [Kharin et al. \(2012\)](#), and the purple line (initial condition) is for the bias correction method given in [Fučkar et al. \(2014\)](#).

of observations and predictions are used to model how the contour obtained from a dynamical model typically differs from the contour observed. New predicted contours are then adjusted to correct for the expected discrepancy. This approach goes beyond existing bias correction methods by addressing the spatial configuration of the sea ice rather than just summary measures of the total amount of sea ice.

In our experiments, contour shifting reduced error in all forecast months and lead times considered. The total area where the prediction and the observations disagree, or IIEE, was reduced by  $3.3 \times 10^5 \text{ km}^2$ , or 21.3%, on average. Contour shifting also produced error reductions in forecasts of the total sea ice area that were better than existing bias correction methods that focus only on forecasting this summary measure. The methodology of contour shifting is not specific to the 15% sea ice concentration threshold used in this paper. After categorizing grid boxes as being above or below a different threshold, a bias-corrected contour for another threshold can be obtained in the same way as the 15% threshold.

We have evaluated spatial prediction accuracy using the IIEE measure ([Goessling et al. 2016](#)). Since contour shifting leads to substantial error reduction, which evaluation criterion is used is not of great importance. Any reasonable criterion summarizing areas of discrepancy between observations and predictions would

likely show improvement of forecast accuracy with bias correction. However, as [Dukhovskoy et al. \(2015\)](#) highlight, there are a number of properties that an ideal sea ice measure should have. As further spatial bias correction methods are developed, more nuanced measures could be useful for distinguishing among them.

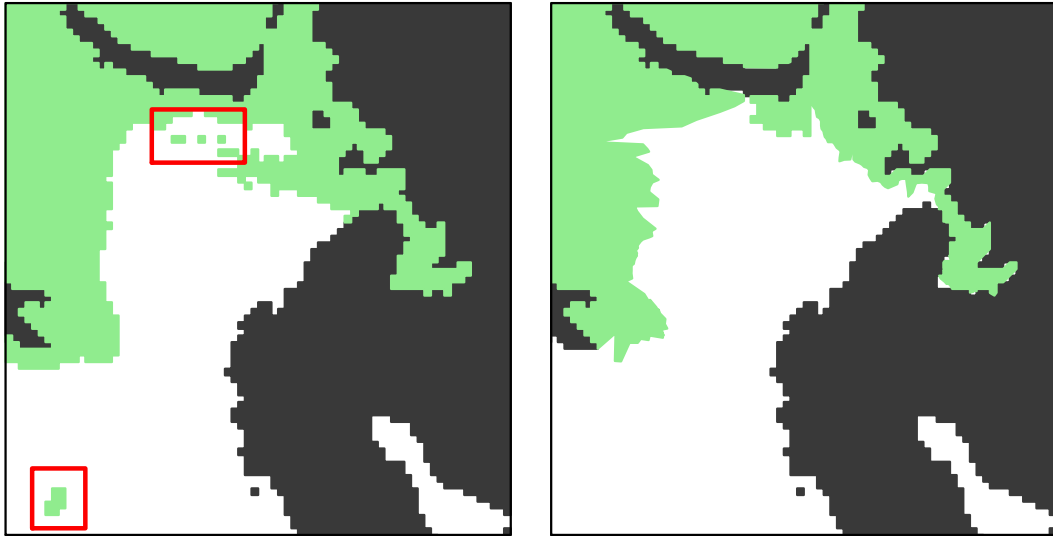
#### a. Method comparisons

A simple alternative to contour shifting for obtaining spatial predictions of sea ice would be to independently apply an existing bias correction technique to each grid

TABLE 1. Root-mean-square error for the total sea ice area predictions for years 2001–13 for all forecast months. GFDL refers to the unadjusted dynamical model. The other rows are for bias-corrected versions of these forecasts. Contour shifting is the bias correction method described in this paper. Trend refers to the bias correction method given in [Kharin et al. \(2012\)](#), and initial condition refers to the bias correction method given in [Fučkar et al. \(2014\)](#). Bold indicates the lowest value in each column.

Method	Lead time (months)				Mean
	2	4	7	12	
GFDL	4.41	4.77	5.22	4.98	4.84
Contour shifting	<b>3.06</b>	<b>3.58</b>	<b>4.00</b>	<b>4.22</b>	<b>3.71</b>
Trend	3.25	3.65	4.13	4.45	3.87
Initial condition	3.47	4.07	4.40	4.78	4.18

### Bias-Corrected Predictions, March 2001, Kara & Barents Seas Trend by Grid Box      Contour-shifting



### Bias-Corrected Predictions, July 2001, Barents & Greenland Seas Trend by Grid Box      Contour-shifting

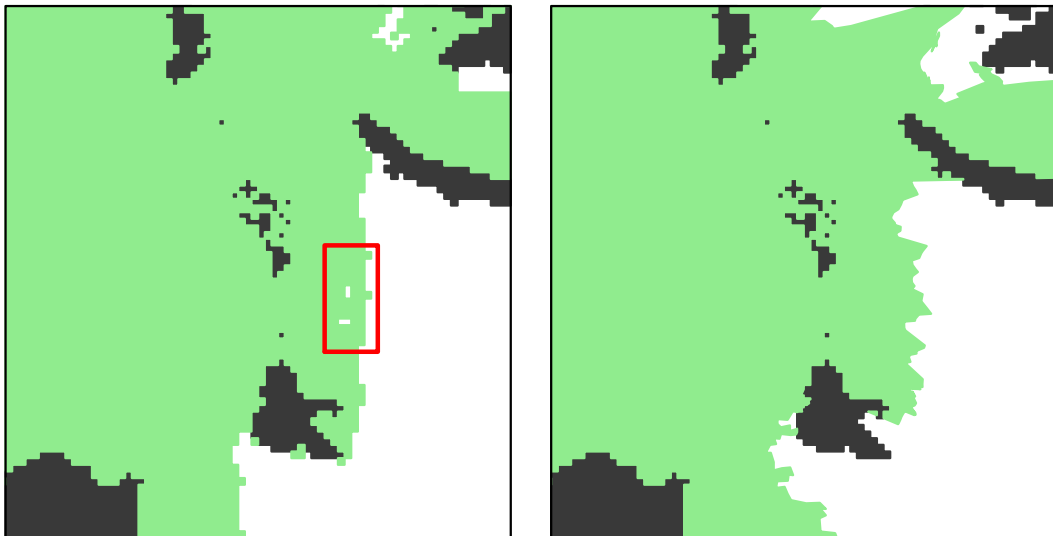


FIG. 13. Predictions of regions with sea ice concentration of at least 15% (green) for March 2001 and July 2001 bias corrected using a variation of the [Kharin et al. \(2012\)](#) approach where logistic regression is used in place of linear regression and concentrations are (left) rounded to be within the interval  $[0, 1]$  and (right) bias corrected with contour shifting. The red boxes (left) highlight disconnected sets of a few grid boxes containing sea ice within open ocean and holes of a small number of grid boxes within large contiguous sections of sea ice.

box. This does notably reduce the IIEE compared to an unadjusted prediction. However, bias correcting each grid box independently occasionally leaves scattered grid boxes or groups of grid boxes that contain sea ice in an area that is otherwise open water or vice versa. We

show examples of this occurring in [Fig. 13](#) using predictions obtained with a slightly modified version of the bias correction proposed by [Kharin et al. \(2012\)](#). Specifically, we modify the [Kharin et al. \(2012\)](#) approach to account for the range of values that concentrations can

take by replacing linear regression with logistic regression and rounding adjusted concentrations down to 1 or up to 0 as needed.

While the artifacts in Fig. 13 are not that common, they illustrate why assuming independence of grid boxes is problematic, especially for binary data. The modeling and computation needed to ensure these artifacts never occur is substantial. A statistical spatial model over the field would be appropriate, but the presence of land and other physical features and the size of the grid complicate the design of an appropriate and computationally tractable covariance structure. Working instead with the lower-dimensional structure of a contour avoids this limiting frame. It also opens up the possibility of extensions to this method, such as covariate-based time series models or probabilistic modeling, which would be computationally challenging to model directly on a full grid.

In section 3a, we define the 15% contour with respect to the concentration level estimates obtained from the ensemble mean for each grid box. However, this is not the only way the sea ice contour could have been defined. For example, in the Sea Ice Outlook produced by the Sea Ice Prediction Network, each grid box in each ensemble member is defined as containing sea ice or not, based on its concentration. The average of these values is used as a summary of each grid box (e.g., [Sea Ice Prediction Network 2014](#)). While common, in this case this approach gives a slightly larger average IIEE for the GFDL forecasts than the technique described in section 3a. This suggests that further investigation is needed to determine how best to convert ensemble model output into a single binary prediction for each grid box. Regardless, even for these alternative definitions of the contour, contour shifting still produces substantial error reductions over unadjusted forecasts.

Connecting two contours with mapping lines has some parallels to the objective curves proposed by [Strong \(2012\)](#) for measuring the marginal ice zone. His approach to matching up curves cannot be applied here since it requires full knowledge of both the predicted and observed curves which we do not have at the time a forecast is issued. However, the similar aspects of these methods highlight the need for techniques that relate contours systematically. We expect that other systematic approaches to matching contours could be used in the general framework of contour shifting with success comparable to the method we have proposed.

#### *b. Limitations and future work*

Expected changes in sea ice concentration leave some questions open about how spatial bias correction will perform in the future. There is currently limited analysis of how the predictability of sea ice will change over time.

Using “perfect model” experiments (which test the ability of a model to predict its own sea ice state), [Holland et al. \(2011\)](#) and [Tietsche et al. \(2013\)](#) both infer that potential skill in predicting sea ice extent will decline. The resulting increase in irreducible error may reduce the performance of bias correction, since biases are more difficult to identify and model in a system with more variability. Additionally, perfect model experiments provide no insight into how biases between models and observations might change. If, in the future, errors are less consistent, the performance of bias correction will decrease, even if the total area in error does not increase. The general pattern of the mapping vectors in the model and observations may also change over time or could be affected by long-term variability. We have been able to obtain substantial error reductions on average using the simple assumption that the time series for the components of all mapping vectors follow a consistent linear trend. However, were the relationships between the model and the observation to change or were the relationships found to be more complicated on multidecadal time scales, we might need more complex time series models, such as state-space models, to respond to changes in the trend and to accurately represent nonlinear structure (e.g., [Brockwell and Davis 2016](#)). Covariate information could also prove useful.

Also, as the concentration of sea ice decreases, non-contiguous sections of sea ice that do not border land and holes in the sea ice may become more common. While an ideal bias correction method would provide a way to correct these features, our analysis does not address these cases. This is because there are currently insufficient examples of observations and predictions exhibiting these phenomena across space to be able to confidently develop a model for their bias. As more data become available, methodological extensions to contour shifting to address biases in these features could be developed and assessed. For example, it would likely be feasible to build a model to identify and correct for regions where the size and/or frequency of holes are over- or underestimated.

Given the extensive development effort and large computational cost required to run dynamical forecasting systems, extracting all possible information from their predictions is of considerable importance. Without bias correction methods that account for the spatial nature of climate data, predictions from these systems can only produce biased spatial information or bias-corrected summary measures. This work represents a first step to overcoming this limitation for sea ice. With contour shifting, we can now use dynamical model output to correct sea ice concentration contours so they have reduced systematic differences from observations.

**Acknowledgments.** Contributions by HMD, AER, and CMB to this study were supported by NOAA's Climate Program Office Climate Variability and Predictability Program through Grant NA15OAR4310161. AER's work was also partly supported by the Center for Advanced Study in the Behavioral Sciences at Stanford University. HMD's contribution to this work was also based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant DGE-1256082. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We thank Edward Blanchard-Wrigglesworth and Yongfei Zhang for helpful discussions and three anonymous reviewers for useful comments that improved the manuscript.

## REFERENCES

- Arisido, M. W., C. Gaetan, D. Zanchettin, and A. Rubino, 2017: A Bayesian hierarchical approach for spatial analysis of climate model bias in multi-model ensembles. *Stochastic Environ. Res. Risk Assess.*, doi:10.1007/s00477-017-1383-2, in press.
- Blanchard-Wrigglesworth, E., K. C. Armour, C. M. Bitz, and E. DeWeaver, 2011a: Persistence and inherent predictability of Arctic sea ice in a GCM ensemble and observations. *J. Climate*, **24**, 231–250, doi:10.1175/2010JCLI3775.1.
- , C. M. Bitz, and M. M. Holland, 2011b: Influence of initial conditions and climate forcing on predicting Arctic sea ice. *Geophys. Res. Lett.*, **38**, L18503, doi:10.1029/2011GL048807.
- , R. I. Cullather, W. Wang, J. Zhang, and C. M. Bitz, 2015: Model forecast skill and sensitivity to initial conditions in the seasonal sea ice outlook. *Geophys. Res. Lett.*, **42**, 8042–8048, doi:10.1002/2015GL065860.
- , and Coauthors, 2017: Multi-model seasonal forecast of Arctic sea-ice: forecast uncertainty at pan-Arctic and regional scales. *Climate Dyn.*, **49**, 1399–1410, doi:10.1007/s00382-016-3388-9.
- Brockwell, P. J., and R. A. Davis, 2016: *Introduction to Time Series and Forecasting*. 3rd ed. Springer, 425 pp.
- Collins, M., 2002: Climate predictability on interannual to decadal time scales: The initial value problem. *Climate Dyn.*, **19**, 671–692, doi:10.1007/s00382-002-0254-8.
- Comiso, J. C., 2000: Bootstrap sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS, version 2. National Snow and Ice Data Center, accessed 1 July 2016, [http://nsidc.org/data/docs/daac/nsidc0079\\_bootstrap\\_seaice.gd.html](http://nsidc.org/data/docs/daac/nsidc0079_bootstrap_seaice.gd.html).
- , C. L. Parkinson, R. Gersten, and L. Stock, 2008: Accelerated decline in the Arctic sea ice cover. *Geophys. Res. Lett.*, **35**, L01703, doi:10.1029/2007GL031972.
- Day, J., S. Tietsche, and E. Hawkins, 2014: Pan-Arctic and regional sea ice predictability: Initialization month dependence. *J. Climate*, **27**, 4371–4390, doi:10.1175/JCLI-D-13-00614.1.
- Director, H. M., A. E. Raftery, and C. M. Bitz, 2017: IceCast: Apply statistical post-processing to improve sea ice predictions, version 1.1.0. R package, <https://CRAN.R-project.org/package=IceCast>.
- Douglas, D. H., and T. K. Peucker, 1973: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica*, **10**, 112–122, doi:10.3138/FM57-6770-U75U-7727.
- Dukhovskoy, D. S., J. Ubnoske, E. Blanchard-Wrigglesworth, H. R. Hiester, and A. Proshutinsky, 2015: Skill metrics for evaluation and comparison of sea ice models. *J. Geophys. Res. Oceans*, **120**, 5910–5931, doi:10.1002/2015JC010989.
- Fučkar, N. S., D. Volpi, V. Guemas, and F. J. Doblas-Reyes, 2014: A posteriori adjustment of near-term climate predictions: Accounting for the drift dependence on the initial conditions. *Geophys. Res. Lett.*, **41**, 5200–5207, doi:10.1002/2014GL060815.
- García-Serrano, J., and F. J. Doblas-Reyes, 2012: On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast. *Climate Dyn.*, **39**, 2025–2040, doi:10.1007/s00382-012-1413-1.
- Goddard, L., and Coauthors, 2013: A verification framework for interannual-to-decadal predictions experiments. *Climate Dyn.*, **40**, 245–272, doi:10.1007/s00382-012-1481-2.
- Goessling, H. F., S. Tietsche, J. J. Day, E. Hawkins, and T. Jung, 2016: Predictability of the Arctic sea-ice edge. *Geophys. Res. Lett.*, **43**, 1642–1650, doi:10.1002/2015GL067232.
- Guemas, V., and Coauthors, 2016: A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales. *Quart. J. Roy. Meteor. Soc.*, **142**, 546–561, doi:10.1002/qj.2401.
- Hazeleger, W., V. Guemas, B. Wouters, S. Corti, I. Andreu-Burillo, F. J. Doblas-Reyes, K. Wyser, and M. Caian, 2013: Multiyear climate predictions using two initialization strategies. *Geophys. Res. Lett.*, **40**, 1794–1798, doi:10.1002/grl.50355.
- Holland, M. M., D. A. Bailey, and S. Vavrus, 2011: Inherent sea ice predictability in the rapidly changing Arctic environment of the Community Climate System Model, version 3. *Climate Dyn.*, **36**, 1239–1253, doi:10.1007/s00382-010-0792-4.
- Huber, P. J., and E. M. Ronchetti, 2011: *Robust Statistics*. Springer, 380 pp.
- Huntington, H. P., and Coauthors, 2015: Vessels, risks, and rules: Planning for safe shipping in Bering Strait. *Mar. Policy*, **51**, 119–127, doi:10.1016/j.marpol.2014.07.027.
- Jones, P. W., 1997: A user's guide for SCRIP: A spherical coordinate remapping and interpolation package. Los Alamos National Laboratory Rep., 29 pp.
- Kharin, V. V., G. J. Boer, W. J. Merryfield, J. F. Scinocca, and W. S. Lee, 2012: Statistical adjustment of decadal predictions in a changing climate. *Geophys. Res. Lett.*, **39**, L19705, doi:10.1029/2012GL052647.
- Krikken, F., M. Schmeits, W. Vlot, V. Guemas, and W. Hazeleger, 2016: Skill improvement of dynamical seasonal Arctic sea ice forecasts. *Geophys. Res. Lett.*, **43**, 5124–5132, doi:10.1002/2016GL068462.
- Maraun, D., 2016: Bias correcting climate change simulations—a critical review. *Curr. Climate Change Rep.*, **2**, 211–220, doi:10.1007/s40641-016-0050-x.
- Meehl, G. A., and Coauthors, 2014: Decadal climate prediction: An update from the trenches. *Bull. Amer. Meteor. Soc.*, **95**, 243–267, doi:10.1175/BAMS-D-12-00241.1.
- Messner, J. W., G. J. Mayr, A. Zeileis, and D. S. Wilks, 2014: Heteroscedastic extended logistic regression for postprocessing of ensemble guidance. *Mon. Wea. Rev.*, **142**, 448–456, doi:10.1175/MWR-D-13-00271.1.
- Msadek, R., G. A. Vecchi, M. Winton, and R. G. Gudgel, 2014: Importance of initial conditions in seasonal predictions of Arctic sea ice extent. *Geophys. Res. Lett.*, **41**, 5208–5215, doi:10.1002/2014GL060799.
- National Center for Atmospheric Research, 2017: Earth system grid at NCAR. NCAR, <https://www.earthsystemgrid.org/home.html>.

- National Snow and Ice Data Center, 2017: Region mask for the Northern Hemisphere. National Snow and Ice Data Center, accessed 22 June 2017, [http://nsidc.org/data/polar-stereo/tools\\_masks.html](http://nsidc.org/data/polar-stereo/tools_masks.html).
- Parkinson, C. L., and D. J. Cavalieri, 2008: Arctic sea ice variability and trends, 1979–2006. *J. Geophys. Res.*, **113**, C07003, doi:10.1029/2007JC004558.
- Sea Ice Prediction Network, 2014: Sea Ice Outlook 2014: August Report. SIPN, <https://www.arcus.org/sipn/sea-ice-outlook/2014/august>.
- Smith, L. C., and S. R. Stephenson, 2013: New trans-Arctic shipping routes navigable by midcentury. *Proc. Natl. Acad. Sci. USA*, **110**, E1191–E1195, doi:10.1073/pnas.1214212110.
- Stephenson, S. R., L. W. Brigham, and L. C. Smith, 2014: Marine accessibility along Russia's northern sea route. *Polar Geogr.*, **37**, 111–133, doi:10.1080/1088937X.2013.845859.
- Stroeve, J. C., V. Kattsov, A. Barrett, M. Serreze, T. Pavlova, M. Holland, and W. N. Meier, 2012a: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations. *Geophys. Res. Lett.*, **39**, L16502, doi:10.1029/2012GL052676.
- , M. C. Serreze, M. M. Holland, J. E. Kay, J. Malanik, and A. P. Barrett, 2012b: The Arctic's rapidly shrinking sea ice cover: a research synthesis. *Climatic Change*, **110**, 1005–1027, doi:10.1007/s10584-011-0101-1.
- , L. C. Hamilton, C. M. Bitz, and E. Blanchard-Wrigglesworth, 2014: Predicting September sea ice: Ensemble skill of the SEARCH sea ice outlook 2008–2013. *Geophys. Res. Lett.*, **41**, 2411–2418, doi:10.1002/2014GL059388.
- Strong, C., 2012: Atmospheric influence on Arctic marginal ice zone position and width in the Atlantic sector, February–April 1979–2010. *Climate Dyn.*, **39**, 3091–3102, doi:10.1007/s00382-012-1356-6.
- Tietsche, S., D. Notz, J. H. Jungclaus, and J. Marotzke, 2013: Predictability of large interannual Arctic sea-ice anomalies. *Climate Dyn.*, **41**, 2511–2526, doi:10.1007/s00382-013-1698-8.
- , and Coauthors, 2014: Seasonal to interannual Arctic sea ice predictability in current global climate models. *Geophys. Res. Lett.*, **41**, 1035–1043, doi:10.1002/2013GL058755.
- van Oldenborgh, G. J., F. J. Doblas-Reyes, B. Wouters, and W. Hazeleger, 2012: Decadal prediction skill in a multi-model ensemble. *Climate Dyn.*, **38**, 1263–1280, doi:10.1007/s00382-012-1313-4.
- Vecchi, G., and Coauthors, 2014: On the seasonal forecasting of regional tropical cyclone activity. *J. Climate*, **27**, 7994–8016, doi:10.1175/JCLI-D-14-00158.1.
- Venables, W. N., and B. D. Ripley, 2002: *Modern Applied Statistics with S*. 4th ed. Springer, 498 pp.
- Wilks, D., 2009: Extending logistic regression to provide full-probability-distribution MOS forecasts. *Meteor. Appl.*, **16**, 361–368, doi:10.1002/met.134.