Response to Anonymous Referee #1

In the following, we present the referee comments in black, our point-by-point response and changes in the manuscript in blue, and literature references at the end of the document.

The manuscript "Predictability of Arctic Sea Ice Drift in Coupled Climate Models" discusses potential predictability of sea ice motion in four climate models from Eulerian and Lagrangian perspectives. The authors identify the potential predictability horizon and identify the wind variability as the main source of uncertainty. The role of initial ice thickness was found to be small. The manuscript is very well written, with good presentation of methodology and results, with deeply though discussions. It provides an important contribution to understanding of sea ice predictability in general. My minor comments only concern few overly complicated explanations in the text that should be simplified for less experience readers.

We thank Referee #1 for reviewing our work and for the valuable feedback and constructive comments, which helped to improve the quality of the manuscript, particularly clearing up overly technical explanations in the introduction and methods section.

Line 30: Although 'errors' and 'uncertainty' are well established terms, it is better to provide here a concise and clear definition of these terms (as well as 'accuracy' and 'skill') as understood by the authors for avoiding ambiguity in the rest of the manuscript.

We thank Referee #1 for this valuable suggestion. We changed the manuscript accordingly as follows. We replaced lines 30-38, that is,

"Initialized predictions inevitably come with errors and uncertainty. Errors arise from physical models being simplified representations of reality, incomplete knowledge of the initial conditions, and inevitable chaotic error growth (Lorenz, 1969, 1975), which gave rise to ensemble forecasting. Here, we therefore differentiate between *errors*, which one should strive to reduce, and *uncertainty*, which must be represented (and communicated) adequately. Both of them act on the forecast accuracy, or skill.

An initialized forecast is commonly considered skillful as long as its accuracy is higher than some chosen benchmark, for example a climatological reference forecast.",

with

"Initialized predictions inevitably come with errors and uncertainty. In this work, we differentiate between *errors*, that originate from *how* a forecast is made, e.g. from physical models being simplified representations of reality, incomplete knowledge of the initial conditions, or truncation errors in numerical models; and (inherent) *uncertainty* due to the inevitable growth of infinitesimal perturbations in the initial conditions, a property of the predicted (chaotic-deterministic) system (Lorenz, 1969, 1975). Errors should be reduced as much as possible, while uncertainty must be represented and communicated adequately. Both act on the forecast accuracy, or skill.

Accuracy refers to the "degree to which forecasts correspond to observations" (Murphy and Winkler, 1992), often described by the mean squared error of an ensemble with respect to a "true" value. An initialized forecast is commonly considered *skillful* as long as its accuracy is higher than some chosen benchmark accuracy, for example from a climatological reference forecast. Forecast skill, often expressed via "skill scores", can therefore be understood as a relative measure of accuracy.".

Line 69: What is "climatological uncertainty"? The following explanation "the uncertainty of an ensemble forecast constructed from independent years simulated by the same model with constant mean climate and variance" seems very short and hard to understand.

We agree that the sentence is hard to digest. We suggest simplifying the sentence as presented below – as it is still part of the introduction – and refer the reader to Section 2.3 ("Measures of predictability") and Hawkins et al. (2016) for a more detailed explanation of how a control simulation should be set up in general, and how it is set up in the used simulations.

We suggest replacing

"Analogous to the forecast skill horizon, here we call a variable (potentially) predictable up to a certain lead time as long as the uncertainty of an initialized ensemble forecast due to chaotic error growth is smaller than the expected climatological uncertainty, that is, the uncertainty of an ensemble forecast constructed from independent years simulated by the same model with constant mean climate and variance."

by

"Analogous to the forecast skill horizon, here we call a variable (potentially) predictable up to a certain lead time if the uncertainty of an initialized ensemble is smaller than the expected climatological uncertainty. This climatological uncertainty, or variability, is usually derived from a control simulation with constant climate, see Sect. 2.1, Sect. 2.3, and Hawkins et al. (2016).".

Does it mean that a model is initialized at some point of time, then it is run for several years (and external forcing is the same every year), then an ensemble is constructed from individual years, then the uncertainty of a predictand in this ensemble is computed and used as a reference?

Yes, this is correct.

Is there a reference to justify building the "climatological uncertainty" this way?

We suggest Collins (2002) and Hawkins et al. (2016). The first study provides a general overview of the problem of initial value (climate) predictability and presents the normalized root mean squared error (NRMSE) as a metric for predictability which is quite often used in predictability studies; the second one provides an introduction on how model experiments for predictability studies should be designed.

So – yes, it is common to build the climatological uncertainty this way. For scalar quantities the climatological standard deviation is often used, or its climatological expectation value.

For how many years should the model run?

For the perfect model approach used in this study, the control simulations should be in equilibrium, which requires some spin-up time beforehand. Furthermore, the control run serves as a background climatology. For a statistically robust assessment of this climatology, the control run must be "sufficiently" long.

In our case, all models were run for at least 100 years for spin-up, and then at least for 200 more years as control simulation. A sample size of 200 years is comparably large, considering that statistics of the real climate mostly use 30 years. However, as per Referee #1's next question, the spin-up time was chosen too short for some of the participating models.

What if the model doesn't stabilize around a constant climate and the "climate uncertainty" continues to grow with the number of years?

This is a very important aspect that does not become clear in our manuscript yet, which is why we suggest the changes presented below.

Day et al. (2016) report that some of the participating models still exhibited linear trends in sea ice extent and volume, for instance. Thusly, the climate is only approximately constant, and these trends needed to be removed in their study, and related studies.

For sea ice drift speed however, we found linear trends of monthly mean ice speeds in January and July to be negligible, they were smaller than 3.1×10^{-2} cm s^{-1} decade^{-1} for the examined models, and trends of the standard deviation of monthly mean ice speeds are in the same order of magnitude. We therefore did not remove any trend from either variable, as we expect the non-trivial removal of a trend in ice thickness, for instance, to complicate the analysis of drift speeds and ice velocity predictability in relation to the initial ice state later on.

We added in line 149 (Section 2.1, "APPOSITE data set"):

"However, Day et al. (2016) report that most participating models were not in equilibrium after the spin-up period; there was significant drift regarding sea ice extent and volume that needed to be removed prior to the analysis of predictability of these variables. We find trends in monthly mean ice drift speeds to be negligible (smaller than 3.1×10^{-2} cm s^{-1} decade^{-1}) so that no trends were removed in this study.".

We added in line 464 (Section 5, "Discussion"):

"While monthly mean ice speeds for January and July did not exhibit noteworthy linear trends, we again mention that the models were not in an equilibrium state after the spin-up period. This might have a more meaningful effect on possible future studies on the relation of ice speed (predictability) and the mean ice state.".

If my understanding is wrong, a better explanation, possibly with a scheme, is worth adding here. On such a scheme the error, uncertainty, accuracy and predictability can be visually shown for easier understanding by readers not well familiar with the topic.

From our perspective, the understanding of Referee #1 is very accurate. Still, we hope to have cleared up open questions by the changes in the manuscript and would refrain from adding a scheme if the editor and Referee #1 do not object.

Line 208–211: It is difficult to understand how the measure of uncertainty is computed. "variance ellipse", "semi-major axis" In which space? Dimensionality of this space? Can an equation be added here?

For the velocity covariance, the respective ellipse described by the covariance matrix is in u-v-space, and the semi-major axis of the ellipse has the dimension of a velocity (in m s^{-1}, for instance). For the position vectors, the measure has the dimension of a length (e.g. in km), and the variance is calculated in a local Cartesian coordinate system, obtained by a coordinate transformation of the geographical positions.

We agree that this should be presented more concisely by providing equations. We therefore added the following part in the Appendix of the manuscript and shortened lines 205-211 as per suggestion of Referee #2 from

"To account for the bivariate character of position and velocity vectors, we chose a different approach here, which we exemplify in the following for velocity vectors. For a given ensemble of velocity vectors at a given position and lead time, we determine the variance ellipse. Our measure for the uncertainty is then the length of the semi-major axis, which is the spectral norm of the covariance matrix of the velocity vectors. This also enables an analysis of the axis ratio and thus the anisotropy of the uncertainty. The uncertainty of initialized forecasts is then given by the mean of all available initializations (at least eight, due to the filtering)."

to

"To account for the bivariate nature of velocity vectors, we describe ensemble spread at a given lead time by the corresponding covariance matrix Σ. Our measure for uncertainty is then the spectral norm of Σ, which is also the length of the semi-major axis of the ellipse described by Σ (see Appendix A). One can thus use Σ for analyzing the anisotropy of uncertainty as well.".

We added in the Appendix:

"We estimate the uncertainty of the ensemble mean of Eulerian velocity vectors by the spectral norm, that is, the square root of the largest eigenvalue, of the covariance matrix in u-v-space. This is equivalent to the semi-major axis a of the covariance ellipse and can be computed as follows.

Let u_j and v_j be the sea ice velocity components of the jth member of an ensemble of size N_{mem} at a fixed lead time, and $\bar{u}$, $\bar{v}$ be the respective ensemble means. The covariance matrix Σ is given by

$$\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{uv} \\ \sigma_{uv} & \sigma_v^2 \end{bmatrix}, \tag{1}$$

where σ_u^2 and σ_v^2 are the variance in the direction of u and v, respectively, and σ_{uv} the covariance:

$$\sigma_u^2 = \frac{1}{N_{\text{mem}}} \sum_{j=1}^{N_{\text{mem}}} (u_j - \bar{u})^2, \tag{2}$$

$$\sigma_v^2 = \frac{1}{N_{\text{mem}}} \sum_{j=1}^{N_{\text{mem}}} (v_j - \bar{v})^2, \tag{3}$$

$$\sigma_{uv} = \frac{1}{N_{\text{mem}}} \sum_{j=1}^{N_{\text{mem}}} (u_j - \bar{u})(v_j - \bar{v}). \tag{4}$$

The largest eigenvalue of Σ, i.e. the length of the semi-major axis a of the variance ellipse, can then be obtained via

$$a^2 = \frac{1}{2}(\sigma_u^2 + \sigma_v^2) + \sqrt{\frac{1}{4}(\sigma_u^2 + \sigma_v^2)^2 - (\sigma_u \sigma_v - \sigma_{uv}^2)}, \tag{5}$$

$$b^2 = \frac{1}{2}(\sigma_u^2 + \sigma_v^2) - \sqrt{\frac{1}{4}(\sigma_u^2 + \sigma_v^2)^2 - (\sigma_u \sigma_v - \sigma_{uv}^2)}, \tag{6}$$

where b is the length of the semi-minor axis for the sake of completeness. The value of a represents the (direction of) maximum variability within the bivariate data, and we therefore consider it an appropriate measure for the uncertainty of the ensemble mean.

For the Lagrangian target positions we follow the same approach, except that we project the (spherical) geographical coordinates onto a (Cartesian) x'-y'-plane before as follows. Let λ_j' and ϕ_j' be longitudes and latitudes from a trajectory ensemble at a fixed lead time in a rotated coordinate system, such that the North Pole of the rotated system represents the center of mass (barycenter) of the positions $[\lambda_j', \phi_j']$. The projection is then readily obtained by

$$x_j' = R\ \cos\phi_j'\ \cos\lambda_j', \tag{7}$$

$$y_j' = R\ \cos\phi_j'\ \sin\lambda_j', \tag{8}$$

with the Earth's radius $R = 6371$ km. Then one can plug in x_j' and y_j' for u_j and v_j in the framework above. Note that, due to the coordinate rotation, it holds $\bar{x}' = \bar{y}' = 0$.".

Line 229-231: This sentence is also difficult to digest. How a plane can be tangential to a point (barycenter)? Please add an equation.

We agree that this sentence needs to be revised. We therefore simplified the bulky version

"We follow the same approach for the position vectors, only that we first use an orthographic azimuthal projection onto the plane tangential to the barycenter of the point cloud given by a single ensemble prediction from a given initial position, for obtaining two-dimensional Cartesian coordinates (in km)"

to

"We follow the same approach for the position vectors, only that these vectors in geographical coordinates are projected onto a local Cartesian coordinate system (with units km) before, see Appendix A.",

and therewith refer to the equations added in response to the previous comment.

Line 352: What does it mean "normalized uncertainty reaches the climatological uncertainty"? Wasn't the normalization done to the climatological uncertainty? (eq. 4)? Shouldn't it read "uncertainty reaches the climatological uncertainty, i.e. normalized uncertainty reaches 1"?

We thank Referee #1 for this attentive remark. This is correct, the normalized uncertainty cannot reach the climatological uncertainty. We therefore changed

"…the normalized uncertainty reaches the climatological uncertainty …"

simply to

"…the uncertainty reaches the climatological uncertainty …".

Line 410: A reference to Fig. 10 should be added.

We added the missing reference in line 413, changing

"For each initial position, we calculate the correlation coefficient for initial ice thickness and the target position uncertainty at 45 d lead time."

to

"For each initial position, we calculate the correlation coefficient for initial ice thickness and the target position uncertainty at 45 d lead time (see Fig. 10).".

References

Collins, M. (2002). Climate predictability on interannual to decadal time scales: the initial value problem. *Climate Dynamics*, 19(8):671–692.

Day, J. J., Tietsche, S., Collins, M., Goessling, H. F., Guemas, V., Guilliory, A., Hurlin, W., Ishii, M., Keeley, S., Matei, D., Msadek, R., Sigmond, M., Tatebe, H., and Hawkins, E. (2016). The Arctic Predictability and Prediction on Seasonal-to-Interannual TimEscales (APPOSITE) data set version 1. *Geosci. Model Dev.*, 9:2255–2270.

Hawkins, E., Tietsche, S., Day, J. J., Melia, N., Haines, K., and Keeley, S. (2016). Aspects of designing and evaluating seasonal-to-interannual Arctic sea-ice prediction systems. *Quarterly Journal of the Royal Meteorological Society*, 142(695):672–683.