International Journal of Forecasting 37 (2021) 1509–1519

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Optimal combination of Arctic sea ice extent measures: A dynamic factor modeling approach☆

Francis X. Diebold [a], Maximilian Göbel [b], Philippe Goulet Coulombe [a,*], Glenn D. Rudebusch [c], Boyuan Zhang [a]

[a] University of Pennsylvania, United States of America
[b] ISEG - Universidade de Lisboa, Portugal
[c] Federal Reserve Bank of San Francisco, United States of America

ARTICLE INFO

Keywords:
Climate modeling
Nowcasting
Model averaging
Ensemble averaging
Denoising

ABSTRACT

The diminishing extent of Arctic sea ice is a key indicator of climate change as well as being an accelerant for future global warming. Since 1978, Arctic sea ice has been measured using satellite-based microwave sensing; however, different measures of Arctic sea ice extent have been made available based on differing algorithmic transformations of raw satellite data. We propose and estimate a dynamic factor model that combines four of these measures in an optimal way and accounts for their differing volatility and cross-correlations. We then use the Kalman smoother to extract an optimal combined measure of Arctic sea ice extent. It turns out that almost all weight is put on the NSIDC Sea Ice Index, confirming and enhancing confidence in the Sea Ice Index and the NASA Team algorithm on which it is based.

1. Introduction

Climate change is one of the most pressing issues of our time, with many severe economic, environmental, and geopolitical consequences. Recently, the application of time series analytical methods to this topic and, more broadly, "climate econometrics" have emerged as a vibrant area of research in the literature, as highlighted, for example, in Hillebrand et al. (2020) and the references therein. One important issue that these methods can address is the loss of Arctic sea ice. The loss of Arctic sea ice is a vital focus point in climate studies. It is both an ongoing conspicuous *effect* of climate change and a *cause* of additional climate change via feedback loops. In particular, reduced Arctic sea ice boosts solar energy absorption via decreased albedo due to its darkening color (see, for example, Diebold and Rudebusch (2019), Pistone et al. (2019) and Stroeve et al. (2012)) and increased methane release due to melting permafrost (see, for example, Vaks et al. (2020)).[1]

There are, however, several alternative measures of Arctic sea ice extent that are based on different processing methodologies of the underlying satellite-based microwave measurement data, and the best choice among these measures is not obvious (Bunzel et al., 2016). In this paper, we study four such sea ice extent (*SIE*) measures, which we denote as the Sea Ice Index (SIE^S), Goddard Bootstrap (SIE^G), JAXA (SIE^J), and Bremen (SIE^B). The top panel of Fig. 1 provides time series plots of these four measures of Arctic *SIE* for the satellite measurement era, which started in 1978. The four measures appear to be

☆ **Acknowledgments**: We thank the editor, an associate editor, and two referees for comments and/or assistance, as well as Gladys Teng and the Penn Climate Econometrics Research Group. The views expressed here are those of the authors and do not necessarily represent those of others in the Federal Reserve System.

* Corresponding author.
E-mail address: gouletc@sas.upenn.edu (P. Goulet Coulombe).

[1] For a broad and insightful overview of the evolution and causes of reduced Arctic sea ice cover, see Shalina et al. (2020).

https://doi.org/10.1016/j.ijforecast.2020.10.006
0169-2070/

almost identical because their scale is dominated by large seasonal swings. However, the effects of seasonality can be removed by plotting each month separately for the four series, as done in the lower twelve panels of Fig. 1. Of course, the Arctic *SIE* measures all trend down in every month, with steeper trends in the low-ice "summer" months (e.g., August, September, October). (Note the different axis scales for different months.) There are also systematic differences across indicators. SIE^G, for example, tends to be high, and SIE^J tends to be low, while SIE^S and SIE^B are intermediate. However, the deviations between various pairs of measures are not rigid; that is, they are not simply parallel translations of each other. Instead, there are sizable time-varying differences among the various measures.

All of this suggests treating the various measures as noisy indicators of latent true sea ice extent, which in turn suggests the possibility of blending them into a single combined indicator with less measurement error. Indeed, some prominent studies have used simple equally weighted averages of competing indicators, with precisely that goal. For example, a recent report on the state of the cryosphere (IPCC, 2019) uses a simple average of three indicators.[2] Simple averages, however, are often suboptimal. Optimality generally requires the use of *weighted* averages that give, for example, less weight to noisier indicators. Motivated by these considerations, in this paper, we propose and explore a dynamic factor state-space model that combines the various published indicators into an optimal measure of sea ice extent, which we extract using the Kalman smoother.

This paper proceeds as follows: in Section 2, we describe the four leading Arctic sea ice extent indicators in this study and the satellites, sensors, and algorithms used to produce them; in Section 3, we propose a basic dynamic-factor state-space model for sea ice extent and use it to obtain optimal extractions of latent extent; finally, Section 4 concludes the paper.

2. Four Arctic sea ice extent indicators

Sea ice extent (*SIE*) indicators are constructed from satellite measurements of the earth's surface using passive microwave sensing, which is unaffected by cloud cover or a lack of sunlight. Several steps are necessary to convert raw reflectivity observations into final *SIE* measurements. First, for a polar region divided into a grid of individual cells, various sensors record a brightness reading or "brightness temperature" for each cell. An algorithm then transforms these brightness readings into fractional surface coverage estimates – sea ice concentration (*SIC*) values – for each grid cell. Finally, *SIE* is calculated by summing the area of all the cells with at least 15 percent ice surface coverage.[3] This rounding up of *SIE* is effectively a bias correction since determining the edge between ice and water can be especially difficult in the summer, when, for example, melting pools on summer ice surfaces can be mistaken for ice-free open water (Meier & Stewart, 2019).

Different algorithms for processing the raw measurements are important in shaping the final *SIE* estimates. In addition, the *SIE* series are not based on identical raw data because they use somewhat different satellites and sensors Comiso (2007) and Comiso, Meier et al. (2017). In this section, we review some aspects of the satellites, sensors, and algorithms that underlie the *SIE* measures.

2.1. Satellites and sensors

Table 1 summarizes the operative dates of the various satellites and sensors that are relevant for Arctic sea ice measurement. The first multi-frequency sensor equipped on a satellite was the Scanning Multichannel Microwave Radiometer (SMMR) launched in 1978 (Cavalieri et al., 1996). Starting in 1987, later sensors – the Special Sensor Microwave Imager (SSM/I) and the Special Sensor Microwave Imager/Sounder (SSMIS) – offered higher-resolution images.[4] In 2002 and 2012, respectively, the Advanced Microwave Scanner Radiometer for EOS (AMSR-E) and Advanced Microwave Scanner Radiometer 2 (AMSR2) sensors were launched, and provided further improvements in resolution (Comiso, Meier et al., 2017).[5] Given the inclinations of the satellite orbits and the spherical shape of the earth, all of the satellites share an inability to observe the Arctic "pole hole" – a circular region at the very top of the world. The size of the pole hole varies across sensors, but historically there has been full confidence that the area covered by the pole hole fulfills the 15 percent *SIC* requirement (Meier & Stewart, 2019).

Table 1 also describes the underlying source data for our four *SIE* indicators. These measures use algorithms to transform the raw satellite brightness data into *SIC* and *SIE* values. We now turn to a more detailed discussion of these algorithms to highlight the differences across *SIE* indicators.

2.2. From raw measurement to SIE: Algorithmic transformations

Once brightness data have been recorded by satellite sensors, an algorithm converts the measurements into estimates of *SIC*. Here, we discuss the algorithms and other details of the various *SIE* indicators.

2.2.1. Sea ice index

Updated on a daily basis and distributed by the National Snow and Ice Data Center (NSIDC), the Sea Ice Index (SII or SIE^S) combines two separate Sea Ice indicators: (1) the Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data

[2] See the notes for their Figure 3.3, page 3–13.

[3] Parkinson and Cavalieri (2008) discuss reasons for using a 15 percent cutoff.

[4] For a detailed discussion of sensor characteristics, see https://nsidc.org/ancillary-pages/smmr-ssmi-ssmis-sensors.

[5] Early in the sample, operational problems prevented data delivery for several days throughout 1986 and between December 1987 and January 1988 (Comiso, 2017). For more recent technical difficulties, see https://www.nrl.navy.mil/WindSat/Description.php.

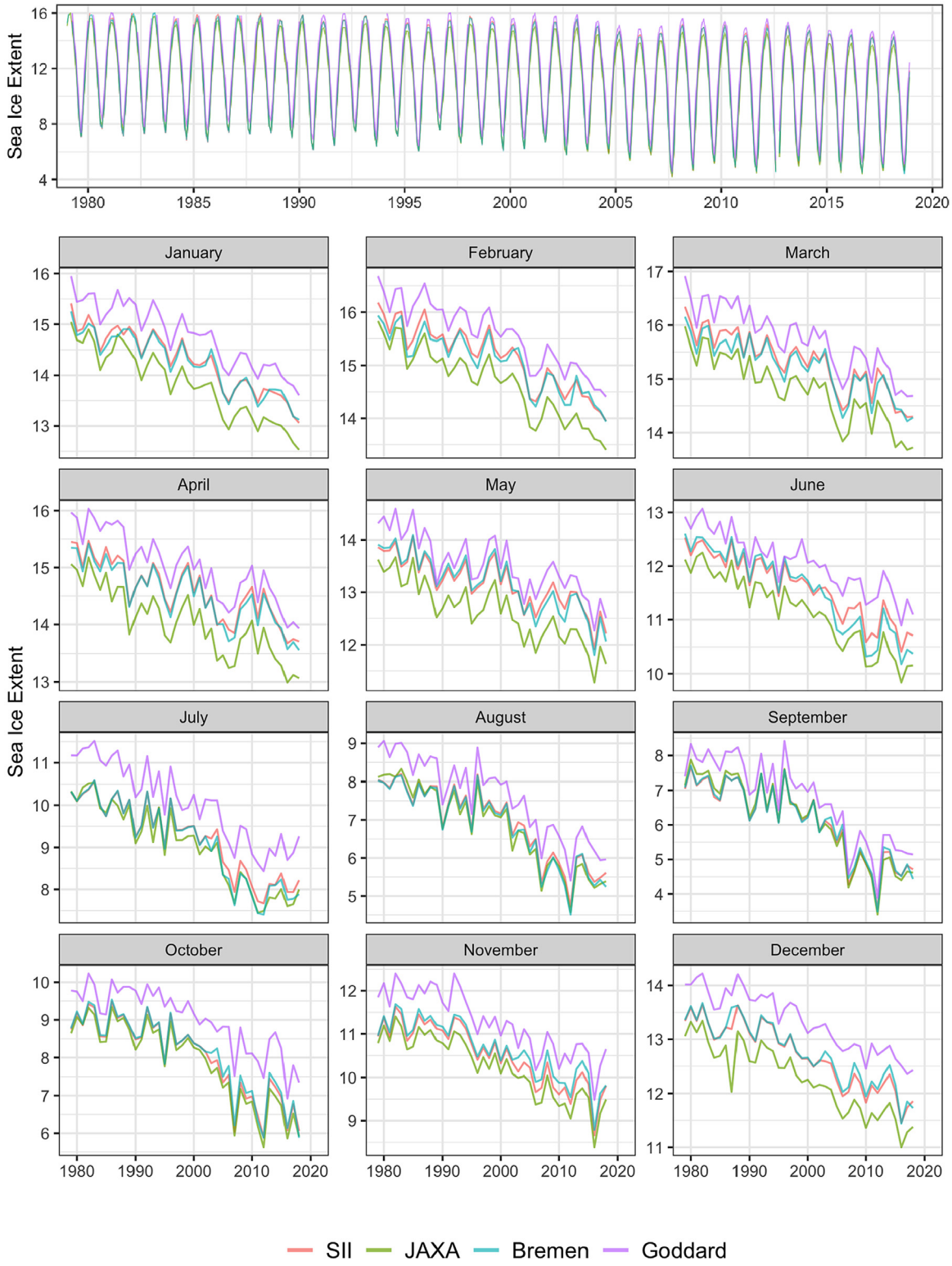

Fig. 1. Four sea ice extent indicators.
Notes: We show the Sea Ice Index (SII), Japan Aerospace Exploration agency (JAXA), University of Bremen (Bremen), and Goddard Bootstrap (Goddard). Units are millions of square kilometers.

Table 1
Satellites, sensors, and algorithms.

Satellite/Sensor	NASA Team		Goddard Bootstrap		JAXA Bootstrap		ASI	
	Start	End	Start	End	Start	End	Start	End
Nimbus-7 SMMR	10/26/1978	08/20/1987	11/01/1978	07/31/1987	11/1978	07/1987	1972	...
DMSP-F8 SSM/I	08/21/1987	12/18/1991	08/01/1987	12/17/1991	07/1987	...	...	...
DMSP-F11 SSM/I	12/19/1991	09/29/1995	12/18/1991	05/09/1995	...	...	...	...
DMSP-F13 SSM/I	09/30/1995	12/31/2007	05/10/1995	12/31/2007	...	06/2002	...	12/2010
DMSP-F17 SSMIS	01/01/2008	12/31/2017	01/01/2008	present			10/2011	07/2012
DMSP-F18 SSMIS	01/01/2018	present						
EOS/Aqua AMSR-E					06/2002	10/2011	2003	10/2011
Coriolis WindSat					10/2011	07/2012		
GCOM-W1 AMSR2					07/2012	present	07/2012	present

Notes: NASA Team and Goddard Bootstrap dates from (Fetterer et al., 2017). JAXA Bootstrap dates from https://kuroshio.eorc.jaxa.jp/JASMES/climate/index.html. ASI dates from https://seaice.uni-bremen.de/sea-ice-concentration/time-series/.

(NASA Team) (Cavalieri et al., 1996) – produced at the Goddard Space Flight Center – and (2) the Near-Real-Time DMSP SSMIS Daily Polar Gridded Sea Ice Concentrations (NRTSI) (Maslanik & Stroeve, 1999) – produced by the NSIDC itself.[6] A time lag of about one year between the *SIE* estimates by the NASA Team and its publication in the NSIDC database requires the NRTSI to complement the SII.

The NRTSI follows the NASA Team algorithm as closely as possible, but inconsistencies between the two series cannot be ruled out entirely (Fetterer et al., 2017). In particular, the two sub-indicators use brightness temperatures from different providers.[7] These raw readings can be distorted by weather effects, making open water look like sea ice cover. Therefore, post-calculation quality checks apply land and ocean masks to remove erroneous and implausible ice cover. However, the NASA Team and NRTSI do not apply the exact same filters (Fetterer et al., 2017). The former algorithm additionally screens the data manually for falsely detected ice formation (Cavalieri et al., 1996), which can enhance accuracy but also reduce the transparency of the final measurements.

As Table 1 shows, the SII obtains the raw data from different generations of satellites and sensors. To make the data comparable, a linear least-squares model of the brightness temperatures, as reported by the two distinct sensors for an overlapping period of operation, is intended to adjust the reference points of 100 percent sea ice and 100 percent open water. These *tie points* then remain fixed over the lifetime of the new system (Cavalieri et al., 2012).

2.2.2. Goddard bootstrap

Another sea ice indicator, distributed by the NSIDC, relies on *SIC* estimates from the Goddard Bootstrap algorithm (SIE^G).[8] Despite the NASA Team and the Goddard Bootstrap algorithms having both been developed at the *NASA Goddard Space Flight Center*, there are some differences between the two approaches. These arise mostly from the calibration of the tie points: while the NASA Team adjusts these reference points for 100 percent open water and 100 percent ice only when a new satellite or sensor becomes operational, the Goddard Bootstrap algorithm adjusts these reference points on a daily basis to account for varying weather conditions (Comiso, Meier et al., 2017). Differing weather filters and sensitivities to varying physical temperatures also lead to differences in the final measurements (Comiso et al., 1997). In contrast to the NASA Team, the strength of the Goddard Bootstrap algorithm is the identification of melting sea ice. Therefore, the Goddard Bootstrap algorithm provides more accurate estimates of the edge of the ice cover (Goldstein et al., 2018).

Although differences between SIE^S and SIE^G are generally assessed to be small, they cannot necessarily be neglected (Goldstein et al., 2018).[9] The differences between the NASA Team and Goddard Bootstrap algorithms particularly occur during the melting season, when the former generally reports larger deviations from ship or radar observations. However, the relative accuracy of the two algorithms is not clear cut since the Goddard Bootstrap algorithm is highly sensitive to physical temperature and underestimates *SIC* during winter periods in the higher latitudes of the Arctic region.

2.2.3. Japan aerospace exploration agency

As listed in Table 1, both SIE^S and SIE^G rely on the same set of instruments, which have been criticized for their low spatial resolution (Goldstein et al., 2018). The Japan Aerospace Exploration Agency (JAXA) sea ice measure (SIE^J) uses an adapted version of the Goddard Bootstrap algorithm to derive *SIC* measures from satellite readings with a higher spatial resolution (Comiso, Gersten et al., 2017).[10] However, readings from these high-resolution satellites have only been available since 2000, so their

[6] See https://doi.org/10.7265/N5K072F8.

[7] Maslanik and Stroeve (1999) takes data from the National Oceanic and Atmospheric Administration Comprehensive Large Array-data Stewardship System (NOAA CLASS); (Cavalieri et al., 1996) uses data processed at the NASA Goddard Space Flight Center.

[8] For a detailed description of the algorithm, see Comiso, Gersten et al. (2017). For the data, see https://doi.org/10.5067/7Q8HCCWS4I0R.

[9] See also https://nsidc.org/support/faq/nasa-team-vs-bootstrap-algorithm.

[10] For a description and data, see https://kuroshio.eorc.jaxa.jp/JASMES/climate/index.html.

data must be merged with observations from older sensors to extend the data coverage to 1978 (Comiso, Meier et al., 2017). SIE^J also distinguishes itself from SIE^S and SIE^G by using a 5-day moving average of observations to compensate for potentially missing data.

2.2.4. University of Bremen

Using observations delivered by the high-resolution AMSR-E sensor, a group of researchers at the University of Bremen developed the ARTIST Sea Ice (ASI) algorithm (Spreen et al., 2008) to estimate daily *SIC*.[11] The time-series (SIE^B) uses different algorithms for different sensors. Until the launch of the AMSR-E sensor in 2003, SIE^B used the NASA Team algorithm to transform brightness readings into *SIC* values. From then on, the NASA Team algorithm was replaced by the ASI.[12]

3. Optimal extraction of latent extent

The four sea ice indicators discussed above differ in terms of the raw data sources and the algorithms used to process the raw data. They can be viewed as distinct indicators of an unobserved or latent "true" sea ice extent, SIE^*. Blending such a set of noisy indicators can produce a single series with less measurement error. Here, we formalize this intuition in a state-space dynamic-factor model, from which we can extract an optimal composite estimate of SIE^* from the four component indicators.[13]

3.1. A dynamic factor model

We work in a state-space environment, modeling each of the four indicators (SIE^S, SIE^J, SIE^B, and SIE^G) as driven by latent true sea ice extent, SIE^*, with an additive measurement error.[14] As discussed previously, some indicators present level shifts with respect to one another, most often resulting from how they respectively deal with tie points. It is thus preferable to constrain both the trend and dynamics to follow a common factor, while leaving level offsets unconstrained. The measurement equation is:

$$\begin{pmatrix} SIE_t^S \\ SIE_t^J \\ SIE_t^B \\ SIE_t^G \end{pmatrix} = \begin{pmatrix} c_S \\ c_J \\ c_B \\ c_G \end{pmatrix} + \begin{pmatrix} \lambda_S \\ \lambda_J \\ \lambda_B \\ \lambda_G \end{pmatrix} SIE_t^* + \begin{pmatrix} \varepsilon_t^S \\ \varepsilon_t^J \\ \varepsilon_t^B \\ \varepsilon_t^G \end{pmatrix}, \tag{1}$$

where

$$\varepsilon_t = (\varepsilon_t^S, \varepsilon_t^J, \varepsilon_t^B, \varepsilon_t^G)' \sim iid\,(0,\ \Sigma)\,, \tag{2}$$

with

$$\Sigma = \begin{pmatrix} \sigma_{SS}^2 & \cdot & \cdot & \cdot \\ \sigma_{JS}^2 & \sigma_{JJ}^2 & \cdot & \cdot \\ \sigma_{BS}^2 & \sigma_{BJ}^2 & \sigma_{BB}^2 & \cdot \\ \sigma_{GS}^2 & \sigma_{GJ}^2 & \sigma_{GB}^2 & \sigma_{GG}^2 \end{pmatrix}. \tag{3}$$

Moreover, $c_i = 0$ when we normalize $\lambda_i = 1$ for $i \in \{S, J, B, G\}$.

The transition equation is:

$$SIE_t^* = \rho\, SIE_{t-1}^* + TREND_t + SEASONAL_t + \eta_t, \tag{4}$$

where $\eta_t \sim iid(0, \sigma_{\eta\eta}^2)$ is orthogonal to ε_t at all leads and lags. Various modeling approaches are distinguished by their treatment of $TREND_t$ and $SEASONAL_t$. We follow (Diebold & Rudebusch, 2019) and allow for 12 monthly deterministic seasonal effects, each of which is endowed with a (possible) deterministic quadratic trend.[15] This results in a blended deterministic "trend/seasonal", given by:

$$TREND_t + SEASONAL_t = \sum_{i=1}^{12} a_i\, D_{it} + \sum_{j=1}^{12} b_j\, D_{jt} \cdot TIME_t + \sum_{k=1}^{12} c_k\, D_{kt} \cdot TIME_t^2, \tag{5}$$

where D_i indicates month i and *TIME* indicates time. Hence, the full transition equation is:

$$SIE_t^* = \rho SIE_{t-1}^* + \sum_{i=1}^{12} a_i\, D_{it} + \sum_{j=1}^{12} b_j\, D_{jt} \cdot TIME_t + \sum_{k=1}^{12} c_k\, D_{kt} \cdot TIME_t^2 + \eta_t. \tag{6}$$

The model is already in a state-space form, and one pass of the Kalman filter, initialized with the unconditional state mean and covariance matrix, provides the 1-step prediction errors necessary to construct the Gaussian pseudo-likelihood, which we maximize using the EM algorithm, and standard errors are calculated from the analytic Hessian matrix.[16] Following estimation, we use the Kalman smoother to obtain the best linear unbiased extraction of SIE^* from the estimated model. The smoother

[11] Monthly data from https://seaice.uni-bremen.de/data/amsr2/today/extent_n_19720101-20181231_amsr2.txt.

[12] See https://seaice.uni-bremen.de/sea-ice-concentration-amsr-eamsr2/time-series/.

[13] Dynamic factor analysis is closely related to principal component analysis, but the dynamic factor model provides a fully specified probabilistic modeling framework in which estimation and factor extraction via the Kalman filter are statistically efficient. Under certain conditions, the two approaches coincide in large samples, but those conditions include a very large number of indicators. Those conditions are violated in our case since we have only four indicators; therefore, the dynamic factor analysis is preferred. For a much more complete account, see Stock and Watson (2011).

[14] The approach parallels (Aruoba et al., 2016), who extract latent "true" U.S. GDP from noisy expenditure-side and income-side estimates.

[15] We emphasize that our model is meant to be a simple benchmark, and that many potentially important variations and extensions are possible. For example, one could alternatively entertain stochastic (as opposed to deterministic) trend and seasonality. A simple approach would be separate month-by month modeling so that there is no seasonality, whether with one unit root, as in (for month m) $TREND_{m,t} = d_m + TREND_{m,t-1} + u_{m,t}$, or two unit roots, as in $TREND_{m,t} = d_{mt} + TREND_{m,t-1} + u_{m,t}$, where $d_{m,t} = d_{m,t-1} + v_{m,t}$.

[16] Note that we do not assume Gaussian shocks, and that it is not necessary to assume Gaussian shocks as we can still maximize the Gaussian likelihood even if the shocks are not truly Gaussian. The resulting (pseudo-)MLE still has good properties of consistency, asymptotic normality, and so forth.

averages across indicators, but it desirably produces optimally weighted averages rather than simple averages. The smoother also averages over time, using data both before and after time t to estimate SIE_t^*, which is also necessary for optimal extraction due to the serial correlation in SIE^*. For details, see Harvey (1989).

One or more restrictions are necessary for identification. The standard approach is to normalize a factor loading, which amounts to an unbiasedness assumption. Normalizing $\lambda_S=1$, for example, amounts to an assumption that SIE^S is unbiased for SIE^*. Whether such an unbiased indicator truly exists (and, if so, which) is of course an open question – one can never know for sure. SIE^S and SIE^G are the most widely used indicators (Meier et al., 2014), Peng et al. (2013), so it is natural to consider normalizing on λ_S or λ_G. We explore both.

3.2. Estimated measurement equation

The estimated measurement equation (1), normalized with $\lambda_S=1$, is:

$$\begin{pmatrix} SIE_t^S \\ SIE_t^J \\ SIE_t^B \\ SIE_t^G \end{pmatrix} = \begin{pmatrix} 0 \\ \underset{[0.025]}{0.225} \\ \underset{[0.020]}{0.043} \\ \underset{[0.034]}{1.040} \end{pmatrix} + \begin{pmatrix} 1 \\ \underset{[0.002]}{0.950} \\ \underset{[0.002]}{0.995} \\ \underset{[0.003]}{0.961} \end{pmatrix} SIE_t^* + \begin{pmatrix} \varepsilon_t^S \\ \varepsilon_t^J \\ \varepsilon_t^B \\ \varepsilon_t^G \end{pmatrix}, \tag{7}$$

where standard errors appear beneath each estimated loading. All indicators are estimated to load heavily on SIE^*, with all $\hat{\lambda}'s$ very close to 1. SIE^J and SIE^G load least heavily ($\hat{\lambda}_J=0.950$, $\hat{\lambda}_G=0.961$), in accordance with their generally less abrupt trend in Fig. 1. SIE^B loads with an estimated coefficient that is marginally different from 1 but significant at the 5% level. Of course, the SIE^S loading is 1 by construction. While Bremen's level offset (with respect to SII) is arguably negligible, those of Jaxa and especially Goddard are sizable. Hence, the estimation results agree with Fig. 1, with SIE^S and SIE^B more in the center of the range, and SIE^J and SIE^G being more extreme.

Alternatively, the estimated measurement equation normalized with $\lambda_G=1$ is:

$$\begin{pmatrix} SIE_t^S \\ SIE_t^J \\ SIE_t^B \\ SIE_t^G \end{pmatrix} = \begin{pmatrix} \underset{[0.039]}{-1.081} \\ \underset{[0.038]}{-0.803} \\ \underset{[0.042]}{-1.033} \\ 0 \end{pmatrix} + \begin{pmatrix} \underset{[0.003]}{1.040} \\ \underset{[0.003]}{0.988} \\ \underset{[0.003]}{1.034} \\ 1 \end{pmatrix} SIE_t^* + \begin{pmatrix} \varepsilon_t^S \\ \varepsilon_t^J \\ \varepsilon_t^B \\ \varepsilon_t^G \end{pmatrix}. \tag{8}$$

The estimated loadings in Eqs. (7) and (8), corresponding to $\lambda_S=1$ and $\lambda_G=1$, respectively, are effectively identical up to the normalization.

Now consider the associated measurement error covariance matrix (3). The estimate for the $\lambda_S=1$ normalization is:

$$\hat{\Sigma} = \begin{pmatrix} \underset{[0.0042]}{0.0003} & \cdot & \cdot & \cdot \\ \underset{[0.0043]}{0.0010} & \underset{[0.0049]}{0.0236} & \cdot & \cdot \\ \underset{[0.0044]}{0.0004} & \underset{[0.0045]}{0.0025} & \underset{[0.0048]}{0.0146} & \cdot \\ \underset{[0.0044]}{-0.0025} & \underset{[0.0048]}{0.0081} & \underset{[0.0046]}{0.0002} & \underset{[0.0056]}{0.0361} \end{pmatrix}, \tag{9}$$

with the implied estimated correlation matrix:

$$\hat{R} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ 0.366 & 1 & \cdot & \cdot \\ 0.206 & 0.134 & 1 & \cdot \\ -0.747 & 0.276 & 0.008 & 1 \end{pmatrix}. \tag{10}$$

Note that $\hat{\sigma}_{GG}^2$ is much higher than any of $\hat{\sigma}_{SS}^2$, $\hat{\sigma}_{JJ}^2$, and $\hat{\sigma}_{BB}^2$, potentially due to different indicators using different methods to determine tie points; i.e., reference points of brightness for 100% sea ice and 100% open water. The choice is crucial for the accurate measurement of SIC within grid cells. Tie points, moreover, need not be constant because brightness readings are sensitive to weather effects and atmospheric forcings (Ivanova et al., 2015). Dynamic tie-point calibration is potentially desirable because it can decrease the bias of SIC measurements (Comiso, Meier et al., 2017). The latest version of the Goddard Bootstrap algorithm, in particular, calibrates tie points daily. One would expect, however, that the bias reduction from dynamic tie-point calibration may come at the cost of potential discontinuities that increase measurement error variance. Our results confirm this conjecture. Our estimate of $\hat{\sigma}_{GG}^2$ is about 13 times that of $\hat{\sigma}_{BB}^2$ (which uses constant tie points).

Alternatively, the estimated measurement error covariance matrix for the $\lambda_G=1$ normalization is:

$$\hat{\Sigma} = \begin{pmatrix} \underset{[0.0041]}{0.0003} & \cdot & \cdot & \cdot \\ \underset{[0.0041]}{0.0008} & \underset{[0.0048]}{0.0233} & \cdot & \cdot \\ \underset{[0.0042]}{0.0003} & \underset{[0.0044]}{0.0022} & \underset{[0.0047]}{0.0144} & \cdot \\ \underset{[0.0042]}{-0.0025} & \underset{[0.0046]}{0.0079} & \underset{[0.0045]}{0.0001} & \underset{[0.0055]}{0.0361} \end{pmatrix}, \tag{11}$$

with the implied estimated correlation matrix:

$$\hat{R} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ 0.311 & 1 & \cdot & \cdot \\ 0.161 & 0.121 & 1 & \cdot \\ -0.782 & 0.272 & 0.002 & 1 \end{pmatrix}. \tag{12}$$

3.3. Estimated transition equation

Now let us move to the transition equation (6). Using the $\lambda_S=1$ normalization, we obtain $\hat{\rho}=0.704$ [0.041] and the trend/seasonal parameter estimates ($\hat{a}_i$, $\hat{b}_j$, and $\hat{c}_k$), as reported in the $\lambda_S=1$ columns of Table 2. The trends for all the months are highly significant and slope downwards. The trends for the summer months (August-November) display a notable negative, and are generally statistically significant, curvature, whereas in the non-summer months the quadratic trend terms are generally small and statistically insignificant. For the $\lambda_G=1$ normalization, we obtain $\hat{\rho}=0.719$ [0.042] and the trend/seasonal parameter estimates ($\hat{a}_i$, $\hat{b}_j$, and $\hat{c}_k$), as reported in the $\lambda_G=1$ columns of Table 2. The $\lambda_S=1$ and $\lambda_G=1$ results are very similar.

3.4. Extracted latent sea ice extent

In Figs. 2 ($\lambda_S=1$) and 3 ($\lambda_G=1$), we show the optimal latent sea ice extent extractions ($\widehat{SIE^*}$) in black, together with the four raw indicators in color, by month. First,

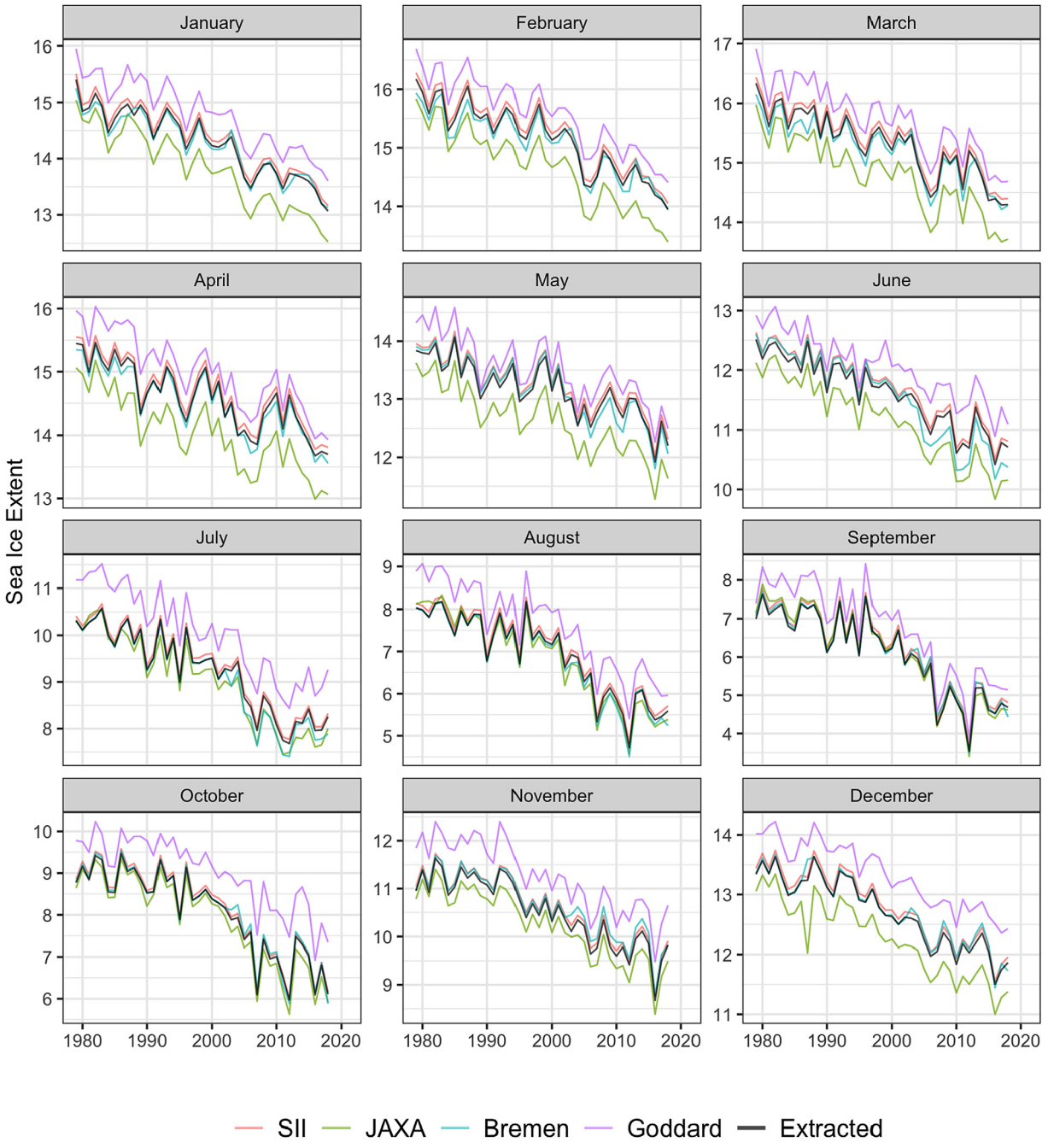

Fig. 2. Extracted sea ice extent and four raw indicators, by month, $\lambda_S=1$.
Notes: We show sea ice extent extracted assuming $\lambda_S=1$, together with four raw indicators: Sea Ice Index (SII), Japan Aerospace Exploration agency (JAXA), University of Bremen (Bremen), and Goddard Bootstrap (Goddard). Units are millions of square kilometers.

consider Fig. 2. Of course $\widehat{SIE}^*(\lambda_S=1)$ is centered on SIE^S due to the $\lambda_S=1$ normalization. Moreover, $\widehat{SIE}^*(\lambda_S=1)$ is always very close – almost identical – to SIE^S (and close to SIE^B because SIE^B tends to be very close to SIE^S).[17]

Now consider Fig. 3 ($\lambda_G=1$). Due to the different normalization, $\widehat{SIE}^*(\lambda_G=1)$ is centered not on SIE^S but rather on SIE^G, so $\widehat{SIE}^*(\lambda_G=1)$ is shifted upward relative to $\widehat{SIE}^*(\lambda_S=1)$. Moreover, the location of $\widehat{SIE}^*(\lambda_G=1)$ relative to SIE^G clearly varies by month. In winter months, it tends to be greater than SIE^G, whereas in summer months it tends to be less than SIE^G. Note in particular that the variation of $\widehat{SIE}^*(\lambda_G=1)$ around SIE^G is noticeably greater

[17] Indeed, when making Fig. 2, we added a tiny constant (0.1) to SIE^S to make it easier to distinguish SIE^S from $\widehat{SIE}^*(\lambda_S=1)$.

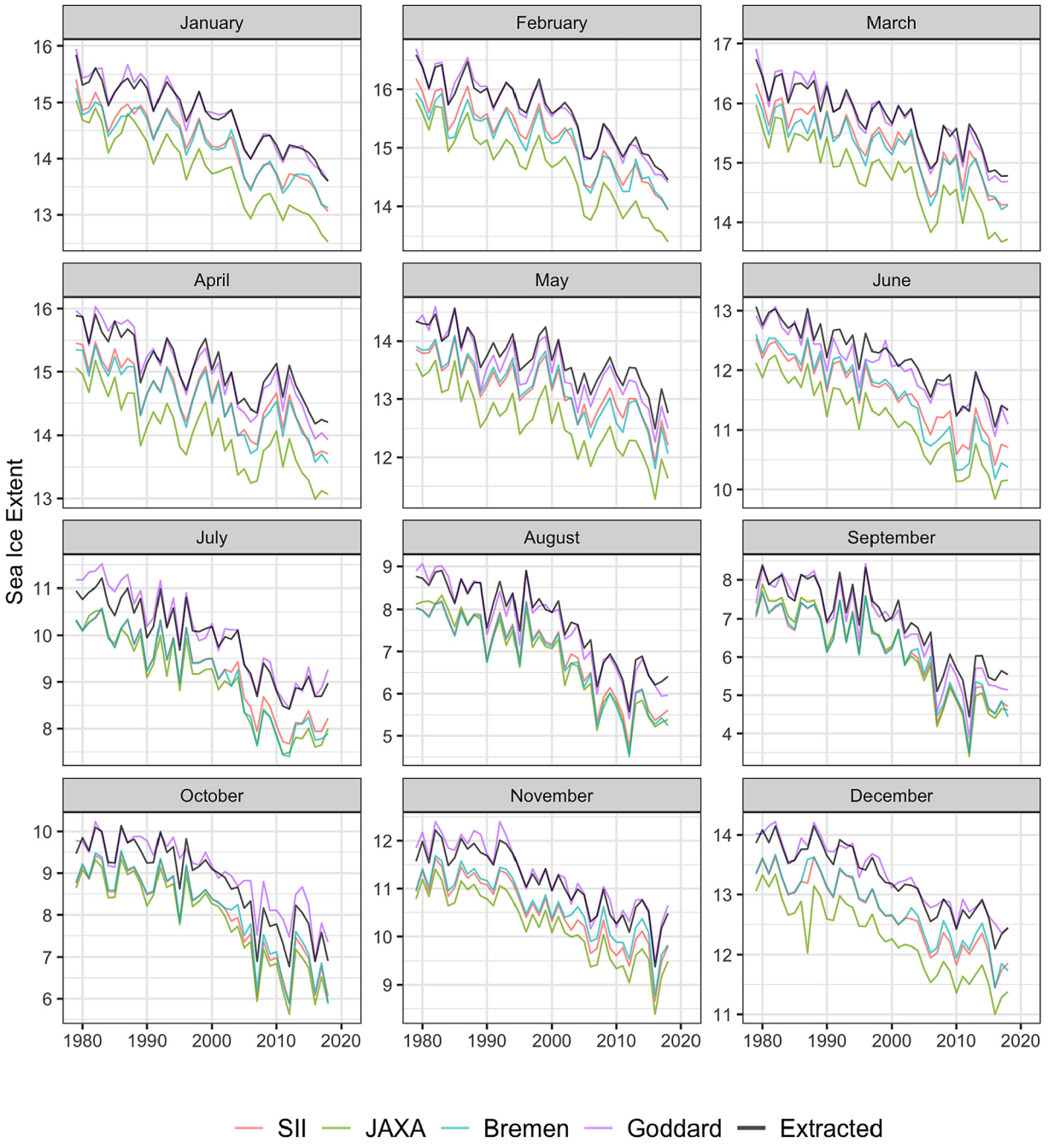

Fig. 3. Extracted latent sea ice extents and four raw indicators, by month, $\lambda_G=1$.
Notes: We show sea ice extent extracted assuming $\lambda_G=1$, together with four raw indicators: Sea Ice Index (SII), Japan Aerospace Exploration agency (JAXA), University of Bremen (Bremen), and Goddard Bootstrap (Goddard). Units are millions of square kilometers.

than the variation of $\widehat{SIE}^*(\lambda_S=1)$ around SIE^S. Clearly, $\widehat{SIE}^*(\lambda_G=1)$ is influenced more by movements in other indicators (SIE^S, SIE^J, and SIE^B) than $\widehat{SIE}^*(\lambda_S=1)$ is.

The fact that $\widehat{SIE}^*(\lambda_S=1)$ and $\widehat{SIE}^*(\lambda_G=1)$ are different, both in terms of level and variation around the level, limits their usefulness for research focusing on level since the level depends entirely on identifying assumptions. However (and crucially), in an important sense $\widehat{SIE}^*(\lambda_S=1)$ and $\widehat{SIE}^*(\lambda_G=1)$ are highly similar: *The model and identification scheme makes* $\widehat{SIE}^*(\lambda_S=1)$ *and* $\widehat{SIE}^*(\lambda_G=1)$ *identical up to a linear transformation.* This is clear in Fig. 4, which plots the two competing extracted factors on a month-by-month basis. Regressions of $\widehat{SIE}^*(\lambda_G=1)$ on $\widehat{SIE}^*(\lambda_S=1)$ yield highly significant intercepts not far from zero, highly significant slopes near 1.0, and R^2 values above 0.999 for each month.

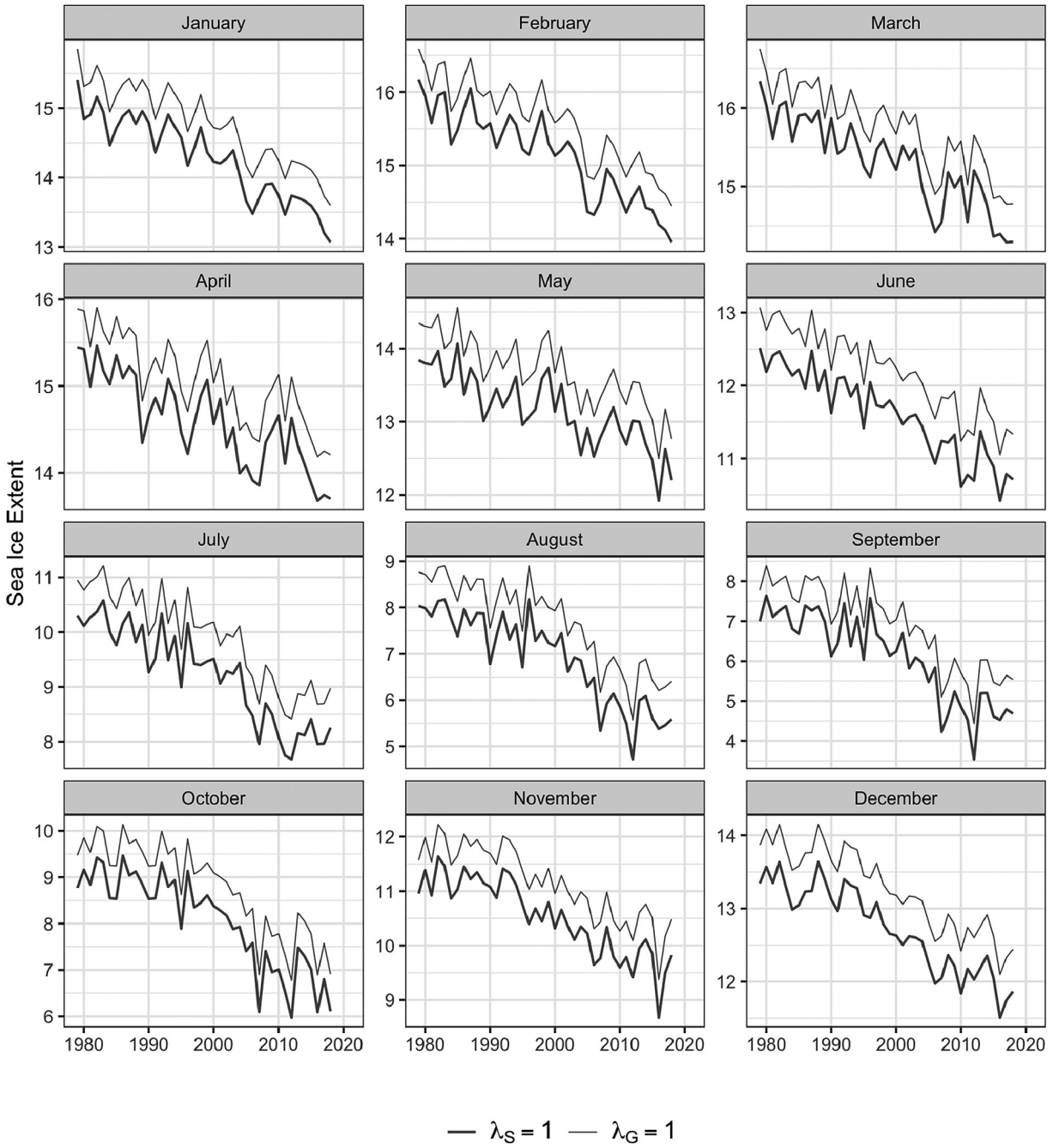

Fig. 4. Extracted sea ice extent, by month. Notes: Units are millions of square kilometers.

Because $\widehat{SIE^*}(\lambda_S=1)$ and $\widehat{SIE^*}(\lambda_G=1)$ are identical up to a linear transformation, it makes no difference which $\widehat{SIE^*}$ we use for research focusing on linear *relationships* between SIE^* and other aspects of climate (for instance, various radiative forcings). The obvious choice, therefore, is $\widehat{SIE^*}(\lambda_S=1)$, which is just SIE^S itself, dispensing with the need to estimate the factor model.

4. Summary and conclusion

We propose a dynamic factor model for four leading Arctic sea ice extent indicators. We estimate the model and use it in conjunction with the Kalman smoother to produce a statistically optimal combination of the individual indicators, effectively "averaging out" the individual measurement errors. We explore two identification

Table 2
Trend/seasonal parameter estimates.

	$\lambda_S=1$			$\lambda_G=1$		
	a_i	b_j	c_k	a_i	b_j	c_k
Jan	5.287 [0.583]	1.412 [1.005]	−4.736 [2.025]	5.245 [0.602]	0.876 [0.953]	−3.608 [1.930]
Feb	5.296 [0.632]	−0.31 [0.995]	−1.735 [2.031]	5.171 [0.653]	−0.267 [0.948]	−1.616 [1.936]
Mar	4.858 [0.666]	−1.227 [0.997]	0.99 [2.033]	4.738 [0.686]	−1.153 [0.950]	1.011 [1.938]
Apr	4.066 [0.672]	−1.719 [1.003]	2.187 [2.027]	3.975 [0.692]	−1.62 [0.956]	2.133 [1.932]
May	2.977 [0.644]	0.367 [1.012]	−2.25 [2.026]	2.939 [0.665]	0.398 [0.965]	−2.176 [1.931]
Jun	2.732 [0.580]	−1.307 [1.011]	−1.072 [2.029]	2.725 [0.603]	−1.225 [0.963]	−1.021 [1.934]
Jul	1.657 [0.525]	−1.497 [1.015]	−3.053 [2.031]	1.711 [0.550]	−1.411 [0.967]	−2.884 [1.935]
Aug	0.719 [0.444]	0.511 [1.029]	−5.634 [2.041]	0.837 [0.472]	0.532 [0.981]	−5.338 [1.945]
Sep	1.715 [0.350]	−1.066 [1.026]	−2.894 [2.064]	1.829 [0.380]	−0.993 [0.978]	−2.667 [1.968]
Oct	3.923 [0.324]	1.274 [1.026]	−5.929 [2.061]	3.958 [0.354]	1.264 [0.978]	−5.583 [1.965]
Nov	4.976 [0.393]	−1.711 [1.025]	3.678 [2.089]	4.947 [0.421]	−1.639 [0.976]	3.715 [1.992]
Dec	5.504 [0.484]	−0.791 [1.032]	0.056 [2.036]	5.423 [0.510]	−0.738 [0.984]	0.133 [1.941]

Notes: The b_j are $\times 10^3$ and the c_k are $\times 10^6$.

strategies that correspond to two different factor loading normalizations. The corresponding two extracted combined measures (latent factors) are identical up to a linear transformation, so either one can be used to explore relationships between sea ice extent and other variables. Interestingly, however, the extracted factor for one of the normalizations puts all the weight on the Sea Ice Index. Therefore, the Sea Ice Index alone is a statistically optimal "combination", and one can simply use it alone with no loss, dispensing with the need to estimate the factor model. There is no gain from combining the Sea Ice Index with other indicators, confirming and enhancing confidence in the Sea Ice Index and the NASA Team algorithm on which it is based, and similarly lending credibility – in a competition against very sophisticated opponents – to the NSIDC's claim that the Sea Ice Index is the "final authoritative SMMR, SSM/I, and SSMIS passive microwave sea ice concentration record" (Fetterer et al., 2017).

References

Aruoba, S. B., Diebold, F. X., Nalewaik, J., Schorfheide, F., & Song, D. (2016). Improving GDP measurement: A measurement-error perspective. *Journal of Econometrics*, *191*, 384–397.

Bunzel, F., Notz, D., Baehr, J., Müller, W. A., & Fröhlich, K. (2016). Seasonal climate forecasts significantly affected by observational uncertainty of arctic sea ice concentration. *Geophysical Research Letters*, *43*, 852–859.

Cavalieri, D. J., Parkinson, C. L., DiGirolamo, N., & Ivanoff, A. (2012). *Intersensor calibration between F13 SSMI and F17 SSMIS for global sea ice data records*. NASA National Snow and Ice Data Center Distributed Active Archive Center, https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20120009376.pdf. (Accessed 31 August 2019).

Cavalieri, D. J., Parkinson, C. L., Gloersen, P., & Zwally, H. J. (1996). *Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS passive microwave data, Version 1*. NASA National Snow and Ice Data Center Distributed Active Archive Center, http://dx.doi.org/10.5067/8GQ8LZQVL0VL. (Accessed 29 August 2019).

Comiso, J. C. (2007). Enhanced sea ice concentrations from passive microwave data. https://nsidc.org/sites/nsidc.org/files/files/data/pm/Bootstrap_Algorithm_Revised07.pdf. (Accessed 30 August 2019).

Comiso, J. C. (2017). Bootstrap sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS, Version 3. https://doi.org/10.5067/7Q8HCCWS4I0R. (Accessed 29 August 2019).

Comiso, J. C., Cavalieri, D. J., Parkinson, C. L., & Gloersen, P. (1997). Passive microwave algorithms for sea ice concentration: A comparison of two techniques. *Remote Sensing of Environment*, *60*(3), 357–384.

Comiso, J. C., Gersten, R. A., Stock, L. V., Turner, G. J., & Cho, K. (2017). Positive trend in the antarctic sea ice cover and associated changes in surface temperature. *Journal of Climate*, *30*(6), 2251–2267.

Comiso, J. C., Meier, W. N., & Gersten, R. (2017). Variability and trends in the Arctic Sea ice cover: Results from different techniques. *Journal of Geophysical Research: Oceans*, *122*(8), 6883–6900.

Diebold, F. X., & Rudebusch, G. D. (2019). Probability assessments of an ice-free arctic: Comparing statistical and climate model projections. Working Paper, arXiv:1912.10774 [stat.AP, econ.EM].

Fetterer, F., Knowles, K., Meier, W. N., Savoie, M., & Windnagel, A. K. (2017). Sea ice index, version 3. https://doi.org/10.7265/N5K072F8. (Accessed 29 August 2019.

Goldstein, M. A, Lynch, A. H., Zsom, A., Arbetter, T., Chang, A., & Fetterer, F. (2018). The step-like evolution of arctic open water. *Nature Scientific Reports*, *8*, 16902.

Harvey, A. C. (1989). *Forecasting, structural time series models, and the Kalman filter*. Cambridge University Press.

Hillebrand, E., Pretis, F., & Proietti, T. (2020). Econometric models of climate change. *Journal of Econometrics*, *214*, 1–294, (special issue).

IPCC (2019). In H.-O. Pörtner, D. C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, M. Nicolai, A. Okem, J. Petzold, B. Rama, & N. Weyer (Eds.), *IPCC special report on the ocean and cryosphere in a changing climate*. https://www.ipcc.ch/srocc/.

Ivanova, N., Pedersen, L. T., Tonboe, R. T., Kern, S., Heygster, G., Lavergne, T., Sørensen, A., Saldo, R., Dybkjær, G., Brucker, L., & Shokr, M. (2015). Inter-comparison and evaluation of sea ice algorithms: Towards further identification of challenges and optimal approach using passive microwave observations. *The Cryosphere*, *9*(5), 1797–1817. http://dx.doi.org/10.5194/tc-9-1797-2015.

Maslanik, J., & Stroeve, J. (1999). Near-real-time DMSP SSMIS daily polar gridded sea ice concentrations, Version 1. https://doi.org/10.5067/U8C09DWVX9LM. (Accessed 29 August 2019).

Meier, W. N., Peng, G., Scott, D. J., & Savoie, M. H. (2014). Verification of a new NOAA/NSIDC passive microwave aea-ice concentration climate record. *Polar Research*, *33*(1), 21004. http://dx.doi.org/10.3402/polar.v33.21004.

Meier, W. N., & Stewart, J. S. (2019). Assessing uncertainties in sea ice extent climate indicators. *Environmental Research Letters*, *14*(3), Article 035005. http://dx.doi.org/10.1088/1748-9326/aaf52c.

Parkinson, C. L., & Cavalieri, D. J. (2008). Arctic sea ice variability and trends, 1979-2006. *Journal of Geophysical Research: Oceans*, *113*(C7), http://dx.doi.org/10.1029/2007JC004558, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007JC004558.

Peng, G., Meier, W. N., Scott, D. J., & Savoie, M. H. (2013). A long-term and reproducible passive microwave sea ice concentration data record for climate studies and monitoring. *Earth System Science Data*, *5*(2), 311–318. http://dx.doi.org/10.5194/essd-5-311-2013.

Pistone, K., Eisenman, I., & Ramanathan, V. (2019). Radiative heating of an ice-free arctic ocean. *Geophysical Research Letters*, *46*(13), 7474–7480. http://dx.doi.org/10.1029/2019GL082914.

Shalina, E. V., Johannessen, O. M., & Sandven, S. (2020). Changes in arctic sea ice cover in the twentieth and twenty-first centuries. In O. M. Johannessen, L. P. Bohylev, E. V. Shalina, & S. Sandven (Eds.), *Sea Ice in the arctic: Past present and future* (pp. 93–166). Springer Nature.

Spreen, G., Kaleschke, L., & Heygster, G. (2008). Sea ice remote sensing using AMSR-E 89-GHz channels. *Journal of Geophysical Research: Oceans*, *113*(C2), http://dx.doi.org/10.1029/2005JC003384.

Stock, J. H., & Watson, M. W. (2011). Dynamic factor models. In M. P. Clements, & D. F. Hendry (Eds.), *Oxford handbook of economic forecasting*. Oxford University Press.

Stroeve, J. C., Serreze, M. C., Holland, M. M., Kay, J. E., Malanik, J., & Barrett, A. P. (2012). The Arctic's rapidly shrinking sea ice cover: A research synthesis. *Climatic Change*, *110*(3), 1005–1027. http://dx.doi.org/10.1007/s10584-011-0101-1.

Vaks, A., Mason, A. J., Breitenbach, S. F. M., Kononov, A. M., Osinzev, A. V., Rosensaft, M., Borshevsky, A., Gutareva, O. S., & Henderson, G. M. (2020). Palaeoclimate evidence of vulnerable permafrost during times of low sea ice. *Nature*, *577*(7789), 221–225. http://dx.doi.org/10.1038/s41586-019-1880-1.