Dynamics and Statistics of the Climate System, 2018, 1–23
doi:10.1093/climsys/dzy001
Advance Access Publication Date: XX XXXX XXXX
Research Article

Research Article

Data-adaptive harmonic decomposition and prediction of Arctic sea ice extent

Dmitri Kondrashov[a,*], **Mickaël D. Chekroun**[a] and **Michael Ghil**[a,b]

[a]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA 90095-1565, USA and [b]Geosciences Department and Laboratoire de Météorologie Dynamique (CNRS and IPSL), École Normale Supérieure and PSL Research University, F-75231 Paris Cedex 05, France.

*Correspondence to Dmitri Kondrashov, Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA 90095-1565, USA; E-mail: dkondras@atmos.ucla.edu

Received 16 October 2017; Revised 28 November 2017; Accepted 16 January 2018

Abstract

Decline in the Arctic sea ice extent (SIE) is an area of active scientific research with profound socio-economic implications. Of particular interest are reliable methods for SIE forecasting on subseasonal time scales, in particular from early summer into fall, when sea ice coverage in the Arctic reaches its minimum. Here, we apply the recent data-adaptive harmonic (DAH) technique of Chekroun and Kondrashov, (2017), *Chaos*, **27** for the description, modeling and prediction of the Multisensor Analyzed Sea Ice Extent (MASIE, 2006–2016) data set. The DAH decomposition of MASIE identifies narrowband, spatio-temporal data-adaptive modes over four key Arctic regions. The time evolution of the DAH coefficients of these modes can be modelled and predicted by using a set of coupled Stuart–Landau stochastic differential equations that capture the modes' frequencies and amplitude modulation in time. Retrospective forecasts show that our resulting multilayer Stuart–Landau model (MSLM) is quite skilful in predicting September SIE compared to year-to-year persistence; moreover, the DAH–MSLM approach provided accurate real-time prediction that was highly competitive for the 2016–2017 Sea Ice Outlook.

Key words: Coupled stochastic oscillators, Hankel matrices, phase and power spectra, Arctic sea ice prediction, stochastic modelling.

1. Introduction and motivation

Decline in Arctic Sea ice extent (SIE) is an area of active scientific research with profound climatic and socio-economic implications, both negative—on global temperatures—and positive—by facilitating navigation in polar waters. Seasonal forecasting of SIE is very challenging due to the high variability in both ocean and atmosphere over the Arctic in summer, as well as the shortness of observational data sets and the inadequacies of physics-based models in simulating sea ice dynamics (Blanchard-Wrigglesworth *et al.*, 2015; Jung *et al.*, 2016). The Sea Ice Prediction Network (SIPN, http://www.arcus.org/sipn) is a collaborative network to facilitate and improve Arctic

© The Author(s) 2018. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

sea ice prediction by combining physics-based and data-driven statistical models in Sea Ice Outlook of September SIE (Stroeve *et al.*, 2015).

Memory effects that appear explicitly in the governing equations have been advocated for the understanding of various climatic phenomena and their natural variability. The delayed oscillator mechanism (Suarez and Schopf, 1988; Battisti and Hirst, 1989) of the El Nino–Southern Oscillation (ENSO) is among the best known (Neelin *et al.*, 1998), but delayed mechanisms have been proposed even earlier in energy balance models to represent the effect of slow growth and decay of ice sheets on reflected solar radiation (Bhattacharya *et al.*, 1982; Roques *et al.*, 2014). Fundamentally, either appearing in the form of discrete or distributed delays into differential equations (Chekroun *et al.*, 2016), the presence of such terms is responsible of nonlinear oscillations that may interact with other factors such as the seasonal forcing to give rise to complex dynamical behaviour (Jin *et al.*, 1994; Tziperman *et al.*, 1994; Chekroun *et al.*, 2018).

Memory effects have been also discussed in the modelling of the Arctic sea ice dynamics. There, they manifest under the form of nonzero time-lagged correlations, such as positive correlations of sea ice anomalies between Spring and Fall that give rise to the sea ice re-emergence mechanism (see Bushuk *et al.*, 2015). Long-term memory present in the sea ice thickness and its role in the dynamics and summer forecasting of sea ice concentration have also been emphasized by Day *et al.* (2014). Among the statistical models for sea ice prediction, Wang *et al.* (2016) have used long-range information in their autoregressive forecast model for daily sea ice concentrations over the Arctic in summer.

More generally, low-order data-driven stochastic linear (Penland, 1996; Newman, 2013) and nonlinear analytic approaches (Chekroun *et al.*, 2011a; Wouters and Lucarini, 2013; Franzke *et al.*, 2015) have demonstrated their relevance for the modelling and prediction of various climatic fields, and several promising directions are emerging (Chekroun *et al.*, 2011a; Majda and Harlim, 2013; Harlim *et al.*, 2014; Lu *et al.*, 2017a, b) to further improve the modelling/prediction skills or assimilation of data within such models. The recent Multilayer Stochastic Model (MSM) framework introduced by Kondrashov *et al.* (2015) emphasizes the ubiquitous role of the triad of nonlinear, stochastic and memory effects in the derivation of data-driven closure models aimed at simulating and predicting the main dynamical features of the targeted spatio-temporal field, be it as an output of a high-end geophysical model or as a set of observations. As discussed in Kondrashov *et al.* (2015), the MSM framework allows for envisioning the data-driven inverse modelling of partial observations of a complex system by means of a broad class of nonlinear and random differential equations, including (possibly nonlinear) terms with retarded arguments (Kondrashov *et al.*, (2015), Proposition 3.3). This framework establishes, thus within a data-driven context, clear connections with the Mori–Zwanzig (MZ) formalism of statistical physics (Kondrashov *et al.*, (2015), Sections 4 and 5); the latter formalism showing that a closed system of equations for a subset of the (original) system's variables consists of a Markovian term, a non-Markovian (memory) term and a stochastic noise term (see Chorin *et al.*, 2002; Chorin and Hald, 2006; Wouters and Lucarini, 2013; Venkataramani *et al.*, 2017). As shown in various contexts, in the absence of scale separation, these non-Markovian terms may turn to be particularly useful for the capture of certain dynamical features (see e.g. Wilks, 2005; Crommelin and Vanden-Eijnden, 2008; Wouters and Lucarini, 2013; Venkataramani *et al.*, 2017 and Chekroun *et al.*, 2015 (Chaps. 6 & 7)). Their benefits, though, are often conditioned to a good approximation of the Markovian terms as well (see Chekroun *et al.*, 2014, 2015, 2017).

By exploiting the layered structure of an MSM, it was shown in Kondrashov *et al.* (2015) that simple estimation techniques, such as multilevel regression, allow for an estimation of the Markovian terms as well as (possibly) complicated memory and noise terms, bypassing thus the challenges encountered in the standard ways of deriving approximations of the optimal MZ closure.

The MSM approach also generalizes various multilevel regression inverse models that include Empirical Model Reduction (EMR) (Kravtsov *et al.*, 2005, Kravtsov *et al.*, 2009; Strounine *et al.*, 2010; Kondrashov *et al.*, 2013), while allowing a greater flexibility in the choice of the nonlinear predictors and ensuring a stable global asymptotic behaviour via the existence of random attractors (Chekroun *et al.*, 2011b; also see *Kondrashov et al.* (2015), [Theorem 3.1 and Corollary 3.2]). However, if the amount of available data are not large enough and the analysis leads to predictors that may require too many coefficients, alternatives must be found.

The data-adaptive harmonic (DAH) method introduced in Chekroun and Kondrashov (2017) proposes such an alternative. Roughly speaking, given a time-evolving data set, it consists of expanding this data set against a set of data-adaptive modes allowing for a reduction of the modelling effort to elemental stochastic models; the objects to be modelled being the transformed coordinates rather than that of the original data set.

By relying on a general class of integral operators built from temporal cross-correlations contained in the data set, it was shown in Chekroun and Kondrashov (2017) that such modes can be easily extracted for multivariate time series issued from a broad class of mixing dynamical systems, provided that the decay of correlations are sufficiently well resolved for the amount of available data. The associated transformed coordinates, obtained by projection onto these modes, are then efficiently modelled within a family of MSMs—namely, multilayer Stuart–Landau models (MSLMs)—which necessitate only few coefficients to be estimated, helping thus to cope with the possible shortness of the data (see Section VIII of Chekroun and Kondrashov (2017).

This study proposes to illustrate in the context of Arctic sea ice modelling that—by adopting the data-adaptive change of basis techniques of Chekroun and Kondrashov (2017)—transformed coordinates can be extracted that reduce the learning effort to a simple class of MSMs with few coefficients to estimate, a feature especially useful due to the shortness of Arctic sea ice records.

This study also follows in the steps of a decade-long successful effort of real-time ENSO forecasting continuously monitored at the International Research Institute for Climate and Society (IRI). Barnston *et al.* (2012) provided a review of this effort and identified the present authors' EMR–ENSO model (Kondrashov *et al.*, 2005; Chekroun *et al.*, 2011a) as highly competitive against other statistical and physics-based models, thus confirming the importance of memory effects in that predictive model (see also Chen *et al.*, 2016). Moreover, Kondrashov and Berloff (2015) have demonstrated that data-adaptive Multichannel Singular Spectrum Analysis (M-SSA) (Ghil *et al.*, 2002) is particularly suited for the inverse modelling within the MSM framework since the M-SSA basis utilizes time-lagged information and therefore implicitly uses memory effects.

The main goal of this article is thus to develop a novel data-driven sea ice model that includes memory effects for participation in the SIPN-coordinated summertime projections for September SIE (Stroeve *et al.*, 2015). Section 2 introduces the Arctic sea ice data set used in this study, namely the Multisensor Analyzed Sea Ice Extent (MASIE) data. Sections 3 and 4 present key mathematical and numerical details about the DAH decomposition, and DAH-based MSLMs, respectively, as well as the usage of the latter models for prediction. Section 5.1 demonstrates how DAH–MSLM helps decompose the complex regional dynamics in the MASIE data set by harmonic patterns that enables an efficient stochastic modelling. Section 5.2 discusses retrospective forecasts of September MASIE in 2013–2016 by our DAH–MSLM model that obey SIPN submission protocols. Finally, discussion of real-time prediction results in 2016 as well as 2017 and concluding remarks follow in Section 6.

2. Data

There are two kinds of data sets related to real-time monitoring of Arctic sea ice. The widely used Sea Ice Index (SII) data set of the National Snow & Ice Center relies on satellite passive microwave measurements as its only source; it provides 35 years of daily sea ice concentrations from 1979 through the present. The passive microwave satellite observations are automatically processed by the NASA Team (Cavalieri *et al.*, 1996) and bootstrap algorithms (Comiso, 2014) to create daily maps of sea ice concentrations; each of the two algorithms has its own biases and limitations.

On the other hand, the MASIE in the Northern Hemisphere (Fetterer *et al.*, 2010) uses several satellite data sources including passive microwave, but it is also based on visual analysis and other data sources and undergoes a form of manual data fusion. The main purpose of MASIE is to provide accurate daily ice conditions to support navigation and operational forecast models, and it is available since 2006. Another difference between SII and MASIE is spatial resolution; the nominal resolution of the latter is 4 km, much higher than the 25 km of the former. The underlying data for the MASIE are aggregated further in space, and the final product provides information about Northern Hemisphere–wide SIE, as well as for 16 smaller Arctic regions http://nsidc.org/data/masie/.

Meier *et al.* (2015) performed comprehensive comparisons of the SII and MASIE data sets and found that MASIE shows higher Arctic-wide extent values throughout most of the year, largely because of the limitations of passive microwave sensors in some conditions, such as surface melt. During the periods, however, that lead to the peak of the melt season, i.e. May–June, as well as at the end of the melt season and the beginning of freeze-up, i.e. late September–October, MASIE tends to indicate lesser ice extents than estimated by SII.

Since our main purpose here is to develop a data-driven model for SIPN submission, MASIE was chosen because it is updated and readily available in real time. Also, to the best of our knowledge, MASIE is currently not used by empirical models submitted to SIPN.

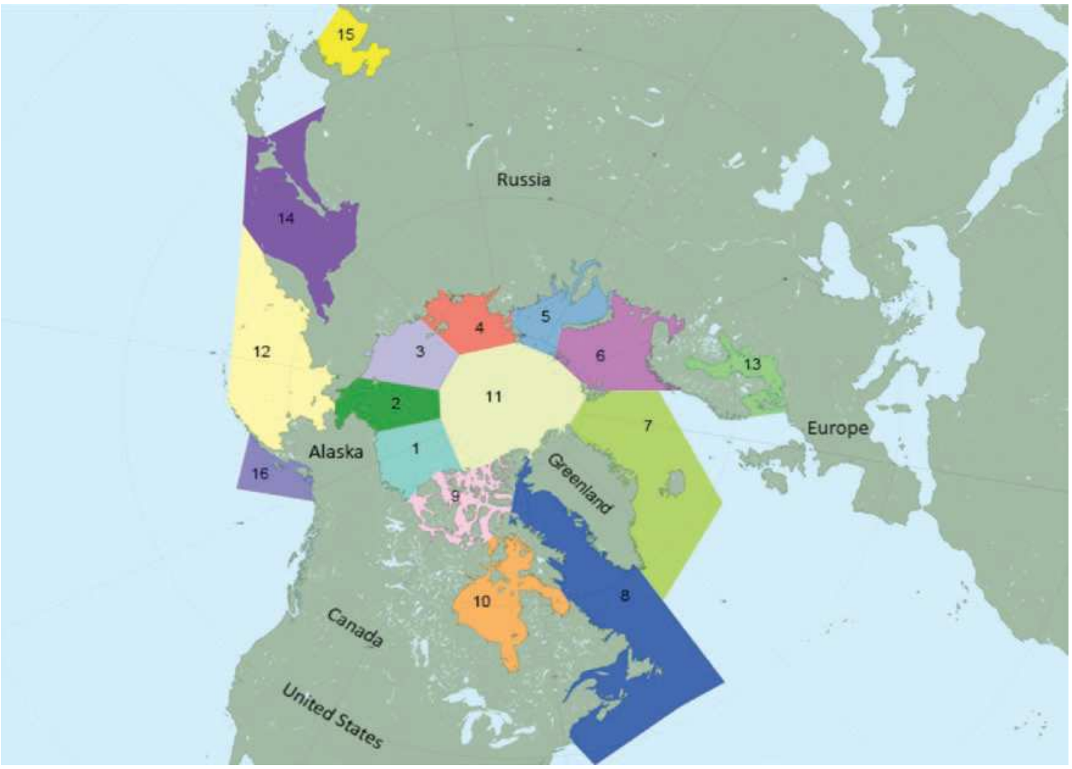

Figure 1. Map of the 16 MASIE regions: 1 – Beaufort Sea, 2 – Chukchi Sea, 3 – East Siberian Sea, 4 – Laptev Sea, 5 – Kara Sea, 6 – Barents Sea, 7 – Greenland Sea, 8 – Baffin Bay and Gulf of St. Lawrence, 9 – Canadian Archipelago, 10 – Hudson Bay, 11 – Central Arctic, 12 – Bering Sea, 13 – Baltic Sea, 14 – Sea of Okhotsk, 15 –Yellow Sea and 16 – Cook Inlet. See text for more details. (Image courtesy of the National Snow & Ice Data Center, University of Colorado, Boulder).

We have used MASIE time series for the North Hemisphere from 1 January 2006 through 27 September 2016. The original daily data were aggregated into a weekly resolution of 52 weeks in each calendar year, as well as combined into four large Arctic regions by adding the underlying subset of 16 regional time series (Fetterer *et al.*, 2010); see Figure 1. The resulting four sectors considered in this article are organized as follows:

- **Central Arctic** (**CEN**),
- **Russia-adjacent** (**RUS**)**:** Laptev Sea, Kara Sea, Barents Sea, Greenland Sea and Baltic Sea,
- **USA-adjacent** (**USA**)**:** Beaufort Sea, Chukchi Sea, East Siberian Sea, Bering Sea, Sea of Okhotsk, Yellow Sea, Cook Inlet,
- **Canada-adjacent** (**CAN**)**:** Baffin Bay, Canadian Archipelago and Hudson Bay.

This grouping is made roughly speaking according to the geographical proximity of the regions, but is also aimed at achieving, in terms of variance, a balanced contribution to the SIE, resulting, in certain cases, into a combination of non-neighbouring regions. At the end, four time series, of 558 weekly data points each, are obtained by adding up the time series of the corresponding subregions. This study focuses on the analysis, modelling and prediction of these four time series that we will refer hereafter as the MASIE data set (see Fig. 2a).

Weekly anomalies of these time series, after removal of the seasonal cycle, are shown in Figure 2b for the full time length of the available MASIE data set and in Figure 3 for the last four corresponding calendar years, i.e. from 2013 to 2016. As shown by these figures, the time evolution of these anomalies does not reveal any recurrent features such as in other climatic fields like ENSO, making its modelling and prediction particularly challenging (cf. Chekroun *et al.*, 2011a).

The total anomaly is dominated by large deviations in the fall, such as in September 2012, with its historic minimum for both the SII and MASIE data sets, while September 2006 exhibited a large positive anomaly. In addition, there were substantial negative and positive anomalies in October 2011 and 2013, respectively. Moreover, the relative contributions of the four large Arctic regions change in time, involve delays and sometimes act in opposite ways.

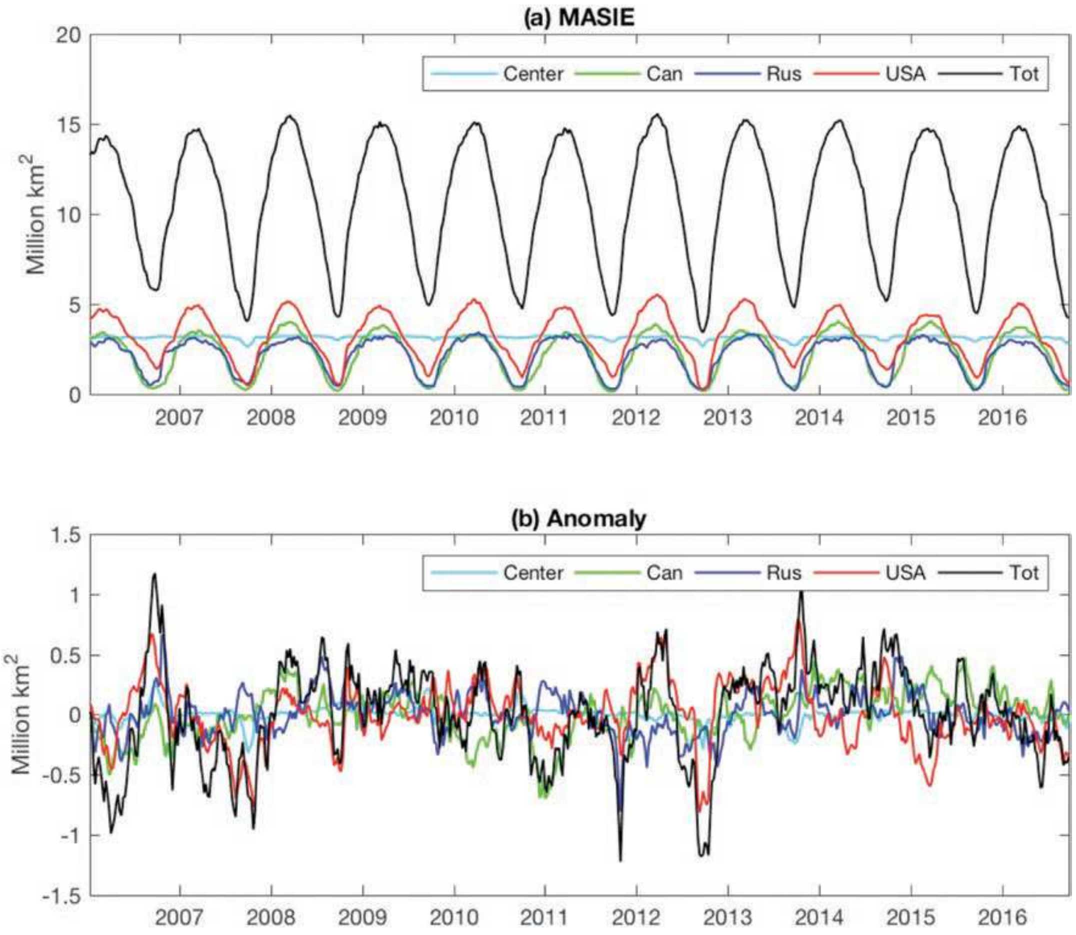

Figure 2. MASIE data set after the CEN-RUS-USA-CAN grouping. (a) Time series for January 2006–September 2016 of the total MASIE (black) and its four regional components with pronounced seasonal cycle and (b) associated anomalies exhibiting complex mixture of multiple temporal scales.

Furthermore, the variability is maximal in the USA region and it is minimal in the Central region, which remains ice covered throughout the year. In general, variability is larger at the peak of the melt season, i.e. in late summer and early fall, than throughout the rest of the calendar year.

The time evolution of the regional MASIE anomalies differs drastically from year to year, due to the complex and nonlinear interactions of air–sea–ice processes. In particular, the year 2015 is strongly dominated by the CAN region, while 2013 and 2016 are dominated by RUS and USA and 2014 by USA and CAN (see Fig. 3). As we will show in Section 5.2, these regional differences greatly influence the prediction skill of the total MASIE data set's SIE, combined over all regions.

3. Data-adaptive harmonic decomposition

The data-adaptive harmonic decomposition (DAHD) (Chekroun and Kondrashov, 2017) is a signal processing technique that allows for a decomposition of the power and phase spectra via data-adaptive modes within a time-embedded phase space. Unlike other techniques exploiting time-embedding—such as M-SSA (Ghil *et al.*, 2002) or its nonlinear version (Giannakis and Majda, 2012)—DAHD exploits a combination of integral operator and semigroup techniques (Engel and Nagel, 2006) that help decompose the original signal into elementary signals that, as illustrated in Section 5 below, are narrowband for each separate discrete Fourier frequency, while being data-adaptive.

The mathematical details of the approach are provided in (Chekroun and Kondrashov, 2017) within a general framework, including the case of multivariate time series issued from a mixing dynamical system, either stochastic or

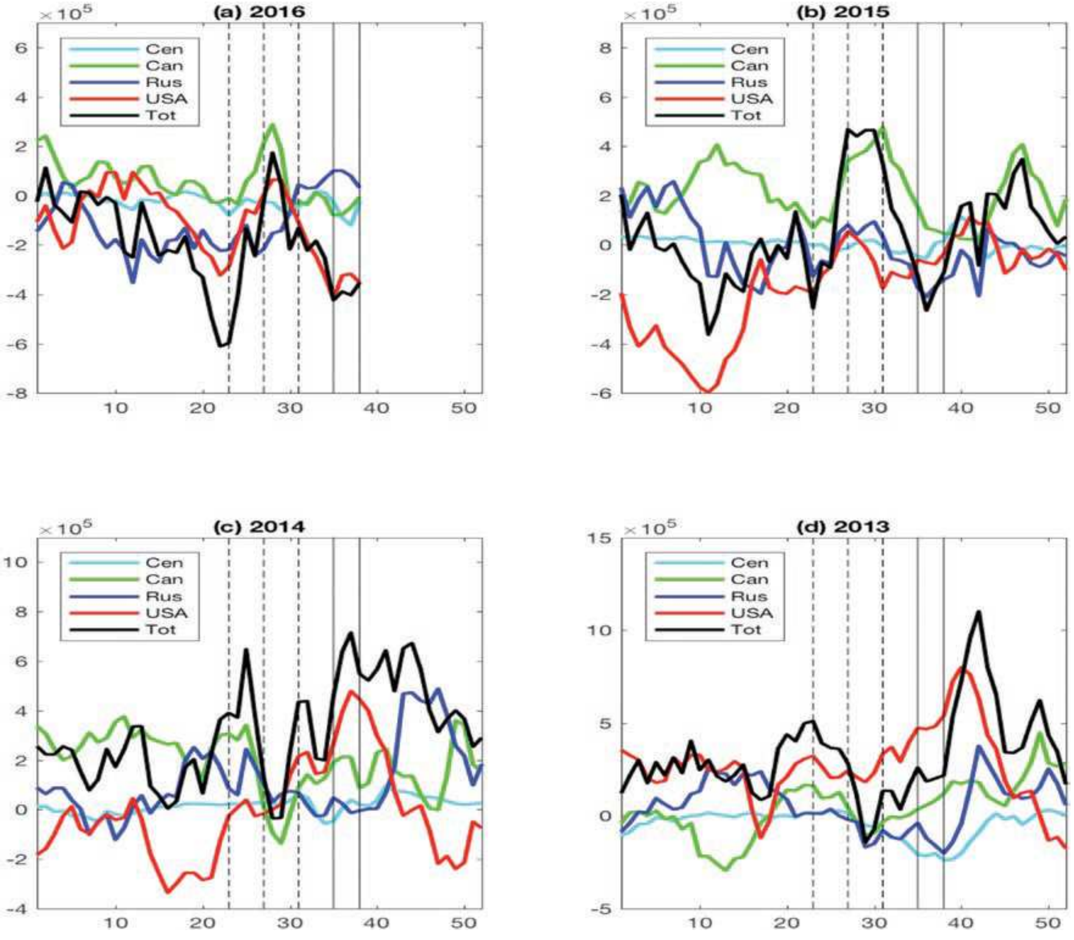

Figure 3. Time evolution of the regional and total MASIE anomalies for the 2013–2016 period: *x*-axis – calendar weeks, *y*-axis – SIE anomalies in km². Thin black solid lines mark the target prediction interval corresponding to the SIPN September Outlook (weeks 35–38), while the thin dashed black lines mark the times when June, July and August (weeks 23, 27 and 31) predictions are issued, respectively (see text for details).

deterministic. Central to the approach is the spectral analysis of a class of integral operators whose kernels are built from correlation functions in a quite different way than found in principal component analysis and its generalizations. For the sake of simplicity, we recall first from (Chekroun and Kondrashov, 2017) how such an integral operator is constructed in the case of a one-dimensional time series $X(t)$. Given the two-sided autocorrelation function (ACF), ρ (of $X(t)$), estimated on the interval $I = [-\tau/2, \tau/2]$, such an operator takes the form

$$\mathcal{L}_\rho(\Psi)(r) := \frac{1}{\tau}\left(\int_{-\frac{\tau}{2}}^{\frac{\tau}{2}-r} \rho(s+r)\Psi(s)\,\mathrm{d}s + \int_{\frac{\tau}{2}-r}^{\frac{\tau}{2}} \rho(r+s-\tau)\Psi(s)\,\mathrm{d}s\right),\quad r \in I, \tag{3.1}$$

and acts on any square-integrable function Ψ on the interval I. The parameter $\tau > 0$ characterizes the embedding window but is chosen in practice so that $\rho(t)$ sufficiently decay over $[-\tau/2, \tau/2]$.

At a practical level, the discretization of the operator $\mathcal{L}_\rho$ defined by (3.1) leads to Hankel matrices built from temporal correlations in a different way than found in M-SSA and alike (see Section 3.1).

For multivariate time series, the ACF, ρ, is replaced by time-lagged cross-correlations, and operators such as given by (3.1) are grouped into a block operator whose discretization results into block Hankel matrices (see (3.4) and [Section VI-D] of Chekroun and Kondrashov (2017)). The aforementioned DAH modes (DAHMs) are then obtained as eigenvectors of such a block Hankel matrix, while the corresponding eigenvalues provide a notion of energy contained into the signal that although allowing for a reconstruction of the signal is not equivalent to variance (see

Remark V.1-(ii) of Chekroun and Kondrashov (2017)). We summarized hereafter the main properties of these spectral elements in a general context before presenting applications to the MASIE data set in Section 5.

3.1. DAH eigenelements and power spectrum

Given a multivariate time series constituted of d 'channels' sampled uniformly at a unit of time δt, i.e. $\mathbf{X}(t_n) = (X_1(t_n), \cdots, X_d(t_n))$, with $t_n = n\delta t$, $n = 1, \cdots, N$, the first step to determine the DAH spectra consists of estimating the two-sided cross-correlation coefficients $\rho_k^{(p,q)}$ between channels p and q at lag k up to a maximum lag $M-1$, i.e. $-M+1 \leqslant k \leqslant M-1$.

As shown in Chekroun and Kondrashov (2017) ([Section VI-D]), the discretization of the operator $\mathcal{L}_\rho$ given by (3.1) with $\rho = \rho^{(p,q)}$ leads to the following Hankel matrix $\mathbf{H}^{(p,q)}$,

$$\mathbf{H}^{(p,q)} = \begin{pmatrix} \rho_{-M+1}^{(p,q)} & \rho_{-M+2}^{(p,q)} & \cdots & \rho_0^{(p,q)} & \rho_1^{(p,q)} & \cdots & \rho_{M-1}^{(p,q)} \\ \rho_{-M+2}^{(p,q)} & \iddots & \iddots & \iddots & \iddots & \iddots & \rho_{-M+1}^{(p,q)} \\ \vdots & \iddots & \iddots & \iddots & \iddots & \iddots & \rho_{-M+2}^{(p,q)} \\ \rho_0^{(p,q)} & \iddots & \iddots & \iddots & \rho_{-M+1}^{(p,q)} & \iddots & \vdots \\ \rho_1^{(p,q)} & \iddots & \iddots & \iddots & \rho_{-M+2}^{(p,q)} & \iddots & \rho_0^{(p,q)} \\ \vdots & \rho_{M-1}^{(p,q)} & \rho_{-M+1}^{(p,q)} & \iddots & \iddots & \iddots & \vdots \\ \rho_{M-1}^{(p,q)} & \rho_{-M+1}^{(p,q)} & \rho_{-M+2}^{(p,q)} & \cdots & \rho_0^{(p,q)} & \cdots & \rho_{M-2}^{(p,q)} \end{pmatrix}. \tag{3.2}$$

Equivalently, this matrix can be viewed as a left-circulant matrix formed from the $(2M-1)$-dimensional row $\mathbf{r} = (\rho_{-M+1}^{(p,q)}, \cdots, \rho_0^{(p,q)}, \cdots, \rho_{M-1}^{(p,q)})$, i.e.:

$$\mathbf{H}^{(p,q)} = l\text{-circ}(\rho_{-M+1}^{(p,q)}, \cdots, \rho_{-1}^{(p,q)}, \rho_0^{(p,q)}, \rho_1^{(p,q)} \cdots, \rho_{M-1}^{(p,q)})\,; \tag{3.3}$$

in other words, the rows of $\mathbf{H}^{(p,q)}$ are obtained by successive shifts to the left by one position, starting from $\mathbf{r}$ as a first row.

As mentioned above, by forming such a Hankel matrix for each (p, q) in $\{1, \cdots, d\}^2$, one can assemble the following block Hankel matrix $\mathfrak{C}$ constituted of d^2 blocks of size $(2M-1) \times (2M-1)$, each given according to

$$\begin{aligned} \mathfrak{C}^{(p,q)} &= \mathbf{H}^{(p,q)}, \text{ if } 1 \leqslant p \leqslant q \leqslant d, \\ \mathfrak{C}^{(p,q)} &= \mathbf{H}^{(q,p)}, \text{ else}\,. \end{aligned} \tag{3.4}$$

Note that because each of its building block, $\mathbf{H}^{(p,q)}$, is symmetric and because $\mathfrak{C}^{(p,q)} = \mathfrak{C}^{(q,p)}$, the grand matrix $\mathfrak{C}$ is itself symmetric. Hereafter, we use $M' = 2M-1$ for concision, reindexing the string $\{-M+1, \cdots, M-1\}$ to run from 1 to M' as necessary.

Chekroun and Kondrashov (2017) ([Theorem V.1]) provides then a useful characterization of the eigenvalues of $\mathfrak{C}$. It shows that the corresponding eigenvalues come in pairs of eigenvalues of opposite sign that can be grouped per Fourier frequency $f\,(\neq 0)$, and are actually given, at each frequency, as the singular values of a cross-spectral matrix depending on the data.

We recall from Chekroun and Kondrashov (2017) the main details concerning this latter property. First, denoting by $\widehat{\rho^{p,q}}(f)$ the Fourier transform at the frequency f of the cross-correlation function $\rho^{p,q}$, we consider the following $d \times d$ **cross-spectral matrix** $\mathfrak{S}(f)$ whose entries are given by

$$\mathfrak{S}_{p,q}^k = \begin{cases} \widehat{\rho^{p,q}}(f) \text{ if } q \geqslant p, \\ \widehat{\rho^{q,p}}(f) \text{ if } q < p\,. \end{cases} \tag{3.5}$$

Then, Chekroun and Kondrashov (2017) ([Theorem V.1]) show that for each **singular value** $\sigma_k(f)$ of $\mathfrak{S}(f)$, there exists, when $f \neq 0$, a pair of negative–positive eigenvalues $(\lambda_-^k(f), \lambda_+^k(f))$ of $\mathfrak{C}$ such that

$$\lambda_+^k(f) = -\lambda_-^k(f) = \sigma_k(f), \quad 1 \leqslant k \leqslant d, \tag{3.6}$$

i.e. $2d$ eigenvalues are associated with each Fourier frequency $f \neq 0$. The same theorem shows that d (but not paired) eigenvalues are associated with the frequency $f = 0$.

Another key property identified by Chekroun and Kondrashov (2017) ([Theorem V.1]) is that the eigenvectors of $\mathfrak{C}$, i.e. the aforementioned DAHMs, can also be grouped per Fourier frequency f in an even more explicit fashion. Indeed, Chekroun and Kondrashov (2017) ([Theorem V.1]) show that the eigenvectors of $\mathfrak{C}$ possess the following representation

$$\mathbf{W}_j = (\mathbf{E}_1^j, \cdots, \mathbf{E}_d^j)^{\mathrm{tr}}, \tag{3.7}$$

where each $\mathbf{E}_k^j$ is a M'-dimensional row vector that is explicitly associated with a Fourier frequency f according to

$$\mathbf{E}_k^j(s) = B_k^j \cos(fs + \theta_k^j), \quad 1 \leqslant s \leqslant M', \tag{3.8}$$

where the amplitudes B_k^j and the phases θ_k^j are both data dependent, for each k in $\{1, \cdots, d\}$.

According to this representation, each DAHM has thus a temporal (embedding) component with s ranging from 1 to M' and a spatial one with k ranging from 1 to d, and since $\mathfrak{C}$ is symmetric, the collection of DAHMs form an orthogonal set. We will sometimes refer to a DAHM snippet, an M'-long segment $\mathbf{E}_k^j$ that arises in the representation (3.7) of a DAHM.

Now thanks to (3.8) we can, given a Fourier frequency,

$$f_\ell = \frac{2\pi(\ell - 1)}{M' - 1}, \quad \ell = 1, \cdots, \frac{M' + 1}{2}, \tag{3.9}$$

determine the following subset of indices in $\{1, \cdots, dM'\}$,

$$\mathcal{J}(f_\ell) := \left\{ j : \text{ s.t. (3.8) holds with } f = f_\ell \right\}. \tag{3.10}$$

Note that for reasons similar to those mentioned above, $\mathcal{J}(f_\ell)$ is composed of $2d$ indices when $\ell \neq 0$ and of d indices if $\ell = 0$ such that the $\mathcal{J}(f_\ell)$'s form a partition of the total set of indices, $\{1, \cdots, dM'\}$.

Due to the simple form (3.8) of a DAHM snippet $\mathbf{E}_k^j$, the determination of the subset $\mathcal{J}(f_\ell)$ can be obtained by various means, for example by computing numerically the power spectral density p_k^j of $\mathbf{E}_k^j$ and grouping the j's for which the average power spectral density, $d^{-1} \sum_{k=1}^d p_k^j$, exhibits a dominant peak at the frequency f_ℓ.

The formula (3.6) is useful for interpretation purposes but not numerically used in practice. Instead, the eigenvalues of the $dM' \times dM'$ matrix $\mathfrak{C}$ are computed directly and listed as the set of eigenvalues $(\lambda_j)_{1 \leqslant j \leqslant dM'}$. This is where the grouping of indices obtained by the procedure described above is used in practice. It allows indeed for a rearrangement of these eigenvalues per Fourier frequency (without having to form the cross-spectral matrix $\mathfrak{S}(f)$ for each frequency) into a useful object called the **DAH power spectrum**. The latter consists of forming, for each ℓ ranging from 1 to $(M' + 1)/2$, the discrete set

$$\mathcal{P}_\ell := \{|\lambda_j|, : j \in \mathcal{J}(f_\ell)\}, \tag{3.11}$$

or in other words, the collection of the $\mathcal{P}_\ell$ for ℓ ranging from 1 to $(M' + 1)/2$, denotes the DAH power spectrum.

As illustrated in Chekroun and Kondrashov (2017) ([Section VII]), when a multivariate data set exhibits a dominant period, the DAH power spectrum displays a distinct peak located at the corresponding frequency. This property is particularly useful for the identification of patterns that, while evolving irregularly in time, manifest a characteristic frequency within a complex data set. We emphasize that such an identification of recurrent patterns can be achieved by various other methods exploiting time-embedding techniques such as the aforementioned M-SSA. The difference, however, is that the DAH power spectrum does not correspond to a decomposition of the variance but rather to a decomposition of another type of energy related to the cross-spectral matrix $\mathfrak{S}(f)$ defined by (3.5) (see [Theorem V.1 and Remark V.1-(ii) in Chekroun and Kondrashov (2017) for details]). This difference is essential as it turns out that the corresponding DAH eigenvalues and eigenmodes allow, unlike M-SSA, for the computation of another useful spectrum, namely the DAH phase spectrum that we describe in the next section.

3.2. DAH phase spectrum

As explained in Chekroun and Kondrashov (2017) ([Section V]), the DAH decomposition does not only provide a data-adaptive power spectrum but also yields a well-defined **DAH phase spectrum** obtained from the relation

$$\Phi_k^j(f_\ell) := \arg(\lambda_j \widehat{\mathbf{E}}_k^j(f_\ell)) - \arg(\overline{\widehat{\mathbf{E}}_k^j}(f_\ell)), \tag{3.12}$$

where $\arg(z)$ denotes the *principal value* (that we adopt to lie in $[0, 2\pi)$ here) of the argument of the complex number z, while $\widehat{\mathbf{E}}_k^j$ and $\overline{\widehat{\mathbf{E}}_k^j}$ denote, respectively, the discrete Fourier transform of the DAHM snippet, $\mathbf{E}_k^j$, and its conjugate. The DAH phase spectrum is then obtained as the collection of the following discrete sets,

$$\Phi_\ell := \{\Phi_k^j(f_\ell) \ : \ j \in \mathcal{J}(f_\ell)\}, \tag{3.13}$$

as ℓ varies from 1 to $(M'+1)/2$.

As illustrated in Chekroun and Kondrashov (2017) ([Section VI-E]), the DAH phase spectrum provides a measure of phase synchronization that would occur across the channels at a given frequency. In particular, it was noticed in Chekroun and Kondrashov (2017) that for periodic or chaotic data sets exhibiting recurrent patterns, the $\Phi_k^j(f)$ become distributed according to the lines across the frequencies located above a critical frequency f_c, lines that typically converge to one point as f approaches the Nyquist frequency f_N.

To the contrary, it has been noted that such a phase synchronization observed on the DAH phase spectrum is typically altered for noisy data sets, indicating that cross-spectral terms become non-negligible (see [Section VIII-B-2] Chekroun and Kondrashov (2017)). Instead, the $\Phi_k^j(f_\ell)$ align along bands within a diffuse background (see e.g. fig. 11 of Chekroun and Kondrashov, 2017).

Finally, we mention that there is a total of d^2 (possibly) distinct phase values $\Phi_k^j(f)$ per frequency since the right-hand side (RHS) of (3.12) yields the same answer for the given DAH pair due to the sign change of the associated λ_j and the phase-quadrature property satisfied by the θ_k^j's arising in the expression (3.8) of the DAHM snippets.

3.3. DAH expansion coefficients

Another useful property concerns the pair of DAHMs associated with a pair of DAH eigenvalues $(\lambda_j, \lambda_{j'})$, such that $\lambda_{j'} = -\lambda_j$ with j and j' that belong thus to the same subset $\mathcal{J}(f)$. For such a DAHM pair, the theory shows indeed that their corresponding phases satisfy $\theta_k^{j'} = \theta_k^j + \pi/2$, i.e. in each DAHM pair, the modes are shifted by one fourth of the period (see [Theorem IV.1] in Chekroun and Kondrashov (2017)). The DAHMs are thus always in exact phase quadrature, as for a sine-and-cosine pair in Fourier analysis, but in a data-adaptive fashion as encapsulated in the θ_k^j's and the B_k^j's (see Fig. 5 for an illustration in the context of the MASIE data set).

By analogy with M-SSA (Ghil *et al.*, 2002), the multivariate data set $\mathbf{X}$ can be projected onto the orthogonal set formed by the $\mathbf{W}_j$'s, to obtain the following **DAH expansion coefficients (DAHCs):**

$$\xi_j(t) = \sum_{s=1}^{M'} \sum_{k=1}^{d} X_k(t+s-1)\mathbf{E}_k^j(s), \tag{3.14}$$

where t varies from 1 to

$$N' = N - M' + 1. \tag{3.15}$$

Although the DAHCs are not formally orthogonal in time, the DAHC pair $(\xi_j(t), \xi_{j'}(t))$ associated with a DAHM pair $(\mathbf{W}_j, \mathbf{W}_{j'})$ is constituted of time series that are *nearly in phase quadrature*, a property that is all the more pronounced when the embedding window parameter M can be made sufficiently large to resolve the decay of temporal correlations contained in $\mathbf{X}$ (see Chekroun and Kondrashov, 2017). In other words, the larger M (subject to the length of the record), the more apparent is the phase-quadrature property exhibited by $\xi_j(t)$ and $\xi_{j'}(t)$ constituting a DAHC pair.

Furthermore, any subset **B** of the set ξ of DAHCs, as well as the full set ξ, can be convolved with its corresponding set of DAHMs, $\mathbf{W}_j$, to produce a partial or full reconstruction of the original data set, respectively. Thus, the following jth **reconstructed component** (**RC**) at time t and for channel k is defined as:

$$R_k^j(t) = \frac{1}{M_t}\sum_{s=L_t}^{U_t} \xi_j(t-s+1)\mathbf{E}_k^j(s),\ \ 1\leqslant s\leqslant M', \tag{3.16}$$

where L_t (resp. U_t) is a lower (resp. upper) bound in $\{1,\cdots,M'\}$, that is allowed to depend on time. The normalization factor M_t equals M', except near the ends of the time series, as in M-SSA (Ghil *et al.*, 2002), and the sum of all the RCs recovers the original time series.

Due to the ordering of the DAHMs in terms of Fourier frequency, the following **harmonic reconstruction component** (**HRC**) can also be formed. It consists of the sum of the RC pairs associated with a same Fourier frequency $f\neq 0$, namely

$$R_k(t;f) = \sum_{j\in\mathcal{J}(f)} R_k^j(t), \tag{3.17}$$

where $\mathcal{J}(f)$ denotes the set of indices given in equation (3.10). The interpretation of an HRC is then natural: it provides an unambiguous way to determine how a particular frequency f is expressed within the time domain.

Note that near the ends of the time series, the estimated values of both DAHCs and RCs are biased due to finite convolution as for M-SSA. Such biases are however not uniform across the frequencies and are not detrimental for the robust prediction of key frequencies, once the appropriate predictors are used for the DAHCs, as discussed in the next section (see also Section 5.2).

4. Multilayer Stuart–Landau models

We turn now to the key aspects of the modelling of a complex data set such as the MASIE data set, by exploiting the DAH eigenspectrum. Instead of addressing the direct modelling of the multivariate data set **X**, we aim instead at modelling the DAHCs, ξ_j, and then at using the reconstruction formulas (3.16) and (3.17). The DAHCs are thus the transformed coordinates mentioned in Section 1. As we explain below, the modelling of these transformed coordinates is simplified due to (i) the near-phase-quadrature property satisfied by any DAHC pair and (ii) its narrowband character.

To setup the ideas, consider a DAHC pair $(\xi_j(t), \xi_{j'}(t))$ associated with a pair of DAH eigenvalues $(\lambda_j, \lambda_{j'})$, such that $\lambda_{j'} = -\lambda_j$ with j and j' that belong thus to the same subset $\mathcal{J}(f)$ associated with a frequency f. Hereafter, we assume the time t to be a continuous parameter.

For such a DAHC pair, we form the complex time series, $\zeta_j(t) = \xi_j(t) + i\xi_{j'}(t)$ where $i^2 = -1$. As shown in Chekroun and Kondrashov (2017) ([Section VII]), we can infer that

$$\xi_j(t) = \mathrm{Re}\,(\zeta_j(0)e^{-ift}) + \sum_{k=1}^{d} B_k^j \int_0^t \cos\left(-(t-r)f+\theta_k^j\right)\left(X_k\left(r+\frac{\tau}{2}\right)-X_k\left(r-\frac{\tau}{2}\right)\right)\mathrm{d}r\,. \tag{4.1}$$

Since the convolution term in the RHS of (4.1) involves a cosine function oscillating at the frequency f, the frequencies $g\neq f$ contained in the time series $X_k(t+\tau/2) - X_k(t-\tau/2)$ are filtered out and the resulting DAHC, $\xi_j(t)$, is narrowband about the frequency f. This latter property is shared by all the DAHCs, $\xi_j(t)$, with j in $\mathcal{J}(f)$, and in particular for $\xi_{j'}(t)$.

Although the RHS of (4.1) provides an exact representation (in the case of an infinite time series) of a DAHC, its usage in practice is limited since it requires the knowledge of the original data set for ulterior time instants, which is incompatible with, e.g., a prediction purpose. Furthermore, we desire a model of the $\xi_j(t)$'s that avoids an explicit dependence on the data, X_k, as found in (4.1).

Guided from the fact that a DAHC pair $(\xi_j(t), \xi_{j'}(t))$ is constituted of time series that are nearly in phase quadrature and that $\xi_j(t)$ and $\xi_{j'}(t)$ are narrowband from what precedes, Chekroun and Kondrashov (2017) have shown that the class of Stuart–Landau (SL) oscillators driven by an additive noise (Selivanov *et al.*, 2012; Zakharova *et al.*, 2016) constitutes a natural class of models to emulate the behaviour of $\xi_j(t)$ and $\xi_{j'}(t)$, in particular their amplitude modulations.

Recall that an SL oscillator writes as

$$\dot{z} = (\mu + i\gamma)z - (1 + i\beta)|z|^2 z + \epsilon_t, \quad z \in \mathbb{C}. \tag{4.2}$$

where μ, γ and β are real parameters and ϵ_t is a noise term, not necessarily white.

With the appropriate parameter values of μ, γ and β as well as noise characteristics of ϵ_t, one can generate a solution $z(t)$ of (4.2) whose real and imaginary parts are also nearly in phase quadrature, modulated in amplitude and narrowband about the frequency f. As a consequence, it is reasonable to envision a good approximation of the complex DAHC $\zeta_j(t)$ by the $z(t)$ solving (4.2).

This idea may turn out however to be insufficient to emulate accurately the behaviour of a DAHC pair $(\xi_j(t), \xi_{j'}(t))$. Indeed, such a pair in general may not be regarded as isolated from the other DAHC pairs associated with the same frequency f, and their collective behaviour must be taken into account. The idea consists then of introducing an appropriate dynamical coupling between the corresponding individual SL oscillators to reproduce any global phase coherence that would be displayed by the DAHC pairs at a given frequency, as well as to take into account the associated temporal and spatial cross-pair correlations in the noise term ϵ_t.

The multilayer MSM framework of Kondrashov *et al.* (2015) is particularly suited to deal with these issues and applied to (4.2), and it leads to MSLMs such as introduced in Chekroun and Kondrashov (2017). In the simplest case of one layer used to model the noise, ϵ_t, an MSLM is given as the following system of SDEs:

$$\begin{aligned}
\dot{x}_j &= \beta_j(f)x_j - \alpha_j(f)y_j + \sigma_j(f)x_j(x_j^2 + y_j^2) + \sum_{\substack{i \neq j \\ i \in \mathcal{J}_d(f)}} b_{ij}^x(f)x_i + \sum_{\substack{i \neq j \\ i \in \mathcal{J}_d(f)}} a_{ij}^x(f)y_i + \epsilon_j^x, \\
\dot{y}_j &= \alpha_j(f)x_j + \beta_j(f)y_j + \sigma_j(f)y_j(x_j^2 + y_j^2) + \sum_{\substack{i \neq j \\ i \in \mathcal{J}_d(f)}} a_{ij}^y(f)x_i + \sum_{\substack{i \neq j \\ i \in \mathcal{J}_d(f)}} b_{ij}^y(f)y_i + \epsilon_j^y, \\
\dot{\epsilon}_j^x &= L_{11}^j(f)x_j + L_{12}^j(f)y_j + M_{11}^j(f)\epsilon_j^x + M_{12}^j(f)\epsilon_j^y + \\
&\qquad Q_{11}^{ij}(f)\dot{W}_1^j + Q_{12}^{ij}(f)\dot{W}_2^j + \sum_{\substack{i \neq j \\ i \in \mathcal{J}_d(f)}} \sum_{k=1}^{2} Q_{1k}^{ij}(f)\dot{W}_k^i, \\
\dot{\epsilon}_j^y &= L_{21}^j(f)x_j + L_{22}^j(f)y_j + M_{21}^j(f)\epsilon_j^x + M_{22}^j(f)\epsilon_j^y + \\
&\qquad Q_{21}^{ij}(f)\dot{W}_1^j + Q_{22}^{ij}(f)\dot{W}_2^j + \sum_{\substack{i \neq j \\ i \in \mathcal{J}_d(f)}} \sum_{k=1}^{2} Q_{2k}^{ij}(f)\dot{W}_k^i.
\end{aligned} \tag{MSLM}$$

Here $(x_j(t), y_j(t))$ is aimed at approximating a DAHC pair $(\xi_j(t), \xi_{j'}(t))$ associated with a frequency $f = f_\ell$, and the index j varies thus in the subset of indices, $\mathcal{J}_d(f_\ell)$, constituted by the indices of $\mathcal{J}(f)$, which correspond to d distinct pairs. The W_k^j's with $k = 1$ or $k = 2$, and $1 \leqslant j \leqslant d$, form $2d$ independent Brownian motions.

Following Kondrashov *et al.* (2015), all model coefficients are estimated starting from the main level of MSLM for each (x_j, y_j) pair, namely by successive multiple linear regression (MLR). Linear constraints on $\alpha_j(f)$, $\beta_j(f)$ and $\sigma_j(f)$ are imposed to ensure antisymmetry for the linear part of each (x_j, y_j) pair, as well as equal and non-positive values $\sigma_j(f) \leqslant 0$ to ensure asymptotic stability. The resulting regression residuals yield time series of $(\epsilon_j^x, \epsilon_j^y)$ that are modelled by means of linear dependencies that involve only $(\epsilon_j^x, \epsilon_j^y)$ on the one hand and the (x_j, y_j) on the other hand (see Eqns (MSLM)). The corresponding model's coefficients are estimated by MLR as well. The regression residuals of the ϵ_j^x- and ϵ_j^y-equations are approximated by pairwise-correlated white noise in the sums involving the W_k^j's. Note that the DAHCs associated with $f \equiv 0$ are modelled a bit differently. Indeed, since they are not paired and resemble to red noises in practice (see Chekroun and Kondrashov, 2017), only a linear MSM is used to model the DAHCs associated with the zero frequency (see Kondrashov *et al.*, 2015).

As needed, extra layers in an MSLM can be added until the regression residuals at the last layer are reasonably approximated by a white noise, according to the stopping criterion described in Kondrashov *et al.* (2015) ([Appendix A]). These layers allow typically for coping with complex state-dependencies—such as dependence on the past of $x_j(t)$ and $y_j(t)$—as well as with temporal correlations exhibited by the residual, ϵ_t, of the main layer for the x_j- and y_j-variables (see [Proposition 3.3 in [Kondrashov *et al.* (2015)]).

Because the SL oscillators are uncoupled across the frequencies, the DAH–MSLM approach is computationally efficient, totally parallelizable and for a variety of data sets, laptop-enabled in practice. First, the model coefficients can be estimated in parallel for each frequency, and the overall number of independent coefficients to estimate remains small and fixed for each (x_j, y_j)-pair.

Once all the resulting (few) MSLM coefficients have been estimated, for the simulation as well, no extra coupling across the frequencies is needed other than running the MSLMs across the frequencies by the same noise realization. The MSLMs are thus also run in parallel across the frequencies, being driven by the same white-noise realization in the last layer.

To simulate the DAHC pairs, Eqns (MSLM) are discretized in time and integrated numerically forward using an Euler-Maruyama scheme from initial states obtained according to the initialization procedure described in Appendix B of Kondrashov *et al.* (2015). Recalling that each DAHC is of length N' given by (3.15), in order to predict the original variable for K steps ahead, first an MSLM prediction of the DAHC pairs is performed for K steps ahead, and then extended DAHCs of length $N' + K$ are formed by concatenation followed by a convolution with the associated DAHMs according to (3.16). The RCs thus obtained are added in each frequency bin to yield the corresponding HRCs given by (3.17). The latter HRCs, in turn, are added across the frequencies to obtain the DAH–MSLM predictions in the original variables.

We turn next to the application of the DAH–MSLM approach for modelling and prediction of the MASIE data set.

5. Numerical results

5.1. Decomposition and modelling

First, we computed the DAH eigenspectrum of the MASIE anomaly time series in our four large Arctic regions (Fig. 2b) over the full time interval 2006–2016, such as extracted from the MASIE data set (see Section 2). To do so, we computed the two-sided cross-correlations between these time series up to lag $M = 39$ weeks and formed the matrix $\mathfrak{C}$ given by (3.4) with $d = 4$ and $M' = 2M - 1 = 77$ weeks. The latter value of M is optimal for the modelling and prediction results presented hereafter (see Section 5.2).

After computation of the eigenvalues λ_j of the matrix $\mathfrak{C}$, we re-arrange them according to the procedure described in Section 3.1 in order to form the DAH power spectrum $\mathcal{P}_\ell$ given in (3.11), over each Fourier frequency f_ℓ.

Figure 4a shows the resulting DAH power spectrum with the corresponding $|\lambda_j|$'s plotted as red filled circles. The DAH power spectrum is evenly spaced in frequency and the number of frequency bins $B_f = (M'+1)/2 = 39$. There are exactly $d = 4$ DAHM pairs at each equidistant frequency f, except at $f = 0$, where there are d unpaired modes. Note that in Figure 4 and the figures hereafter, the frequency values f that appear therein correspond to $f_\ell/(2\pi)$.

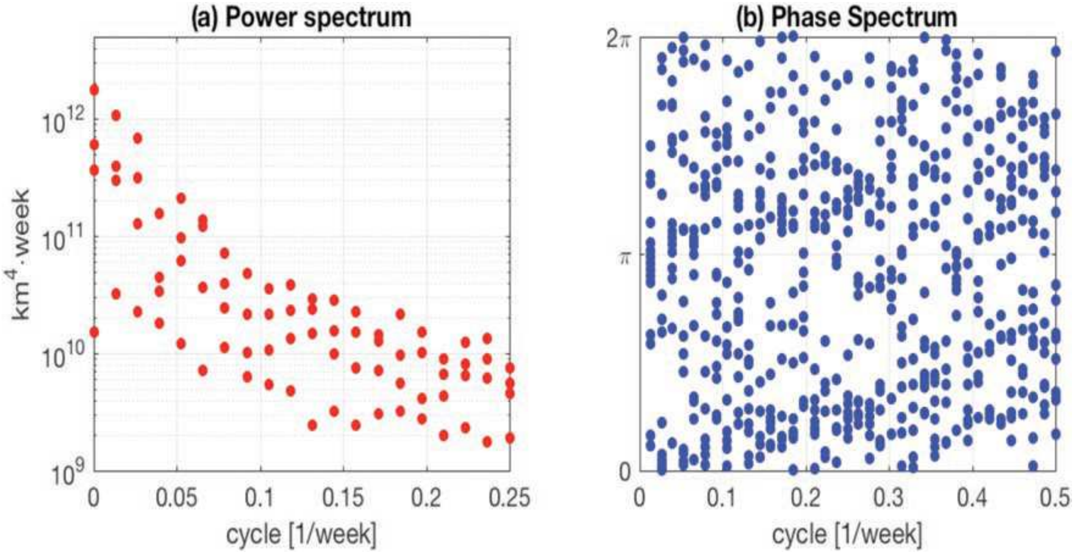

Figure 4. DAH power and phase spectra of the MASIE data set. (a) **DAH power spectrum** $\mathcal{P}_\ell$ given by (3.11) for the MASIE anomalies. Each $\mathcal{P}_\ell$ is constituted by four elements represented by red dots (equal to the MASIE data set's channels number) corresponding to the DAH pairs of eigenvalues λ_j per frequency bin f_ℓ given by (3.9) (see Section 3.1). Each red dot corresponds also a pair of DAHMs associated with the same temporal frequency (see Fig. 5). (b) **DAH phase spectrum** $\boldsymbol{\Phi}_\ell$ given by (3.13) for the MASIE anomalies, showing mostly a random character of the 'interactions' between the MASIE regions.

There is higher power at low frequencies, i.e. intraseasonal band due to both larger anomalies typically occurring in the melt season (see Section 2), as well as downward trend, but there are no distinct peaks above red-noise type background. Recalling the interpretation provided by (3.6), the separation between the DAH pairs at a given frequency f reflects how the singular values of the cross-spectral matrix $\mathfrak{C}(f)$ defined by (3.5) are distributed for our MASIE data set.

As explained in Section 3.2, the DAHMs allow us to define also a phase spectrum $\mathbf{\Phi}_\ell$ given by (3.13) and shown in Figure 2d, as ℓ varies from 1 to B_f. This phase spectrum reflects here the complex nature of sea ice dynamics across the spatio-temporal scales. In particular, there is a hint of parallel lines across the whole frequency band within a diffuse background that indicates a possible inherent noisy content. On the other hand, the identification of structures in the DAH phase spectrum such as discussed in Section 3.2 may be made complicated here due to uncertainties related to the shortness of MASIE observations.

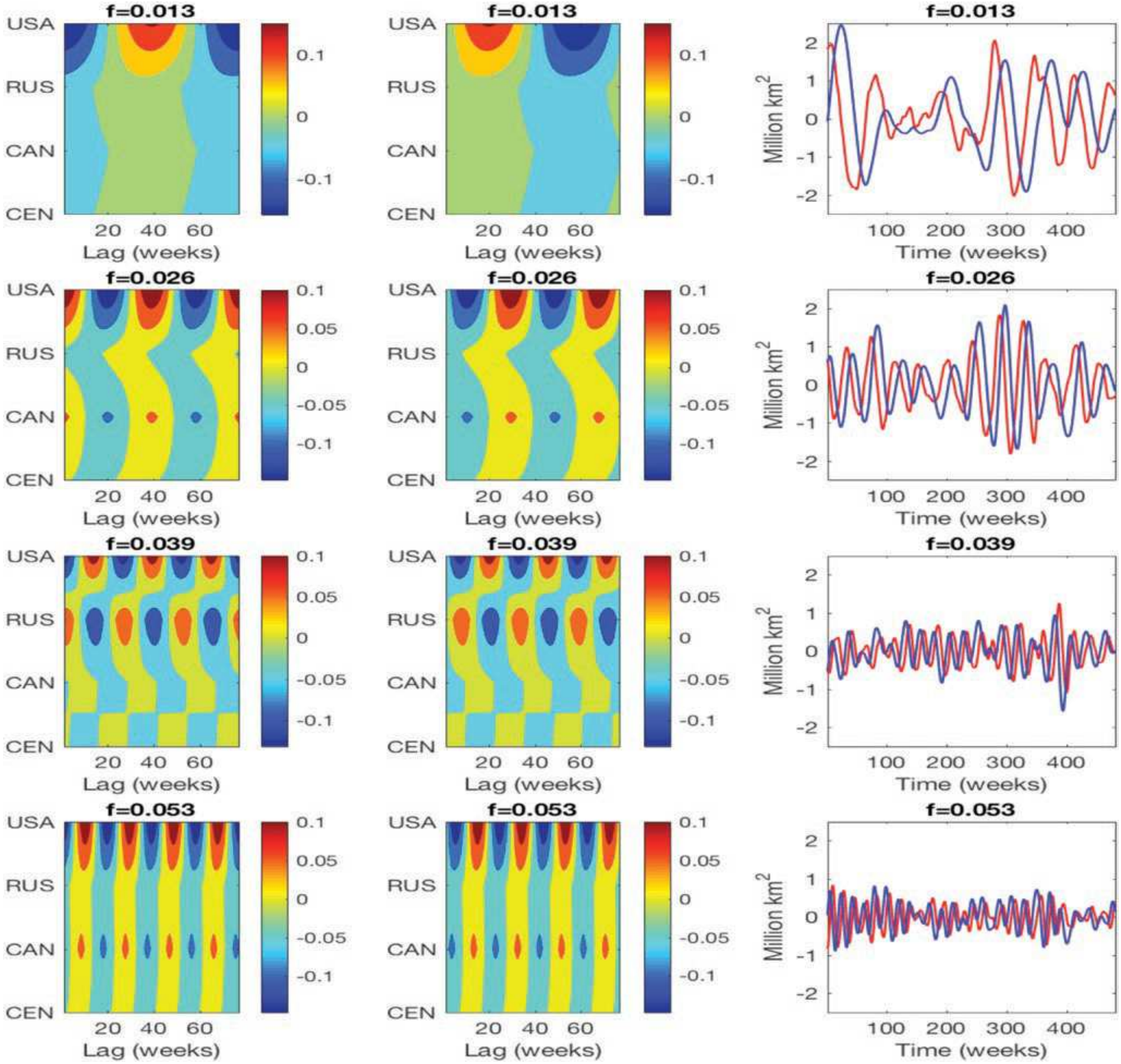

Figure 5. Left and centre panels: Leading **DAH mode (DAHM)** pairs corresponding to the largest $|\lambda_j|$ at the given frequency; the $|\lambda_j|$ being shown in the DAH power spectrum of Figure 4a. Each of the modes constituting a DAHM pair is time-shifted by a quarter of a period, i.e. in exact phase quadrature. The colour bar represents DAHM amplitude in non-dimensional units. Right panels: **DAH coefficients (DAHCs)** obtained by projection of the MASIE data set onto the DAHMs (see (3.14)). A DAHC pair is constituted of time series that are narrowband about the same temporal frequency of the associated DAHM pair, exhibiting however pronounced modulations in amplitude.

The spatio-temporal DAHM patterns shown in the left and centre columns of Figure 5 provide information on the evolution and mutual influence of regional MASIE anomalies in particular frequency bands. For example, the dominant harmonic patterns of the variability, i.e. those corresponding to the pair having largest $|\lambda_j|$ at a given frequency, convey different type of relationships between the regions, such as out-of-phase co-variability between the RUS and USA Arctic sectors at $f = 0.039$ and in-phase influence between the CAN and USA sectors at $f = 0.053$. Furthermore, these patterns may exhibit a time-lagged influence between different sectors, such as for the pairs shown in the two rows at the bottom of Figure 6, where slanting ridges are present; DAHM pairs at other frequencies exhibit also complex and non-trivial patterns.

As predicted by the theory and in particular (4.1), the time series of DAHCs obtained from (3.14) and constituting a pair, display a temporal behaviour narrowband about the characteristic frequency associated with the respective DAHM pair. This property is indeed clearly displayed by the time series $\xi_j(t)$ shown in the right columns

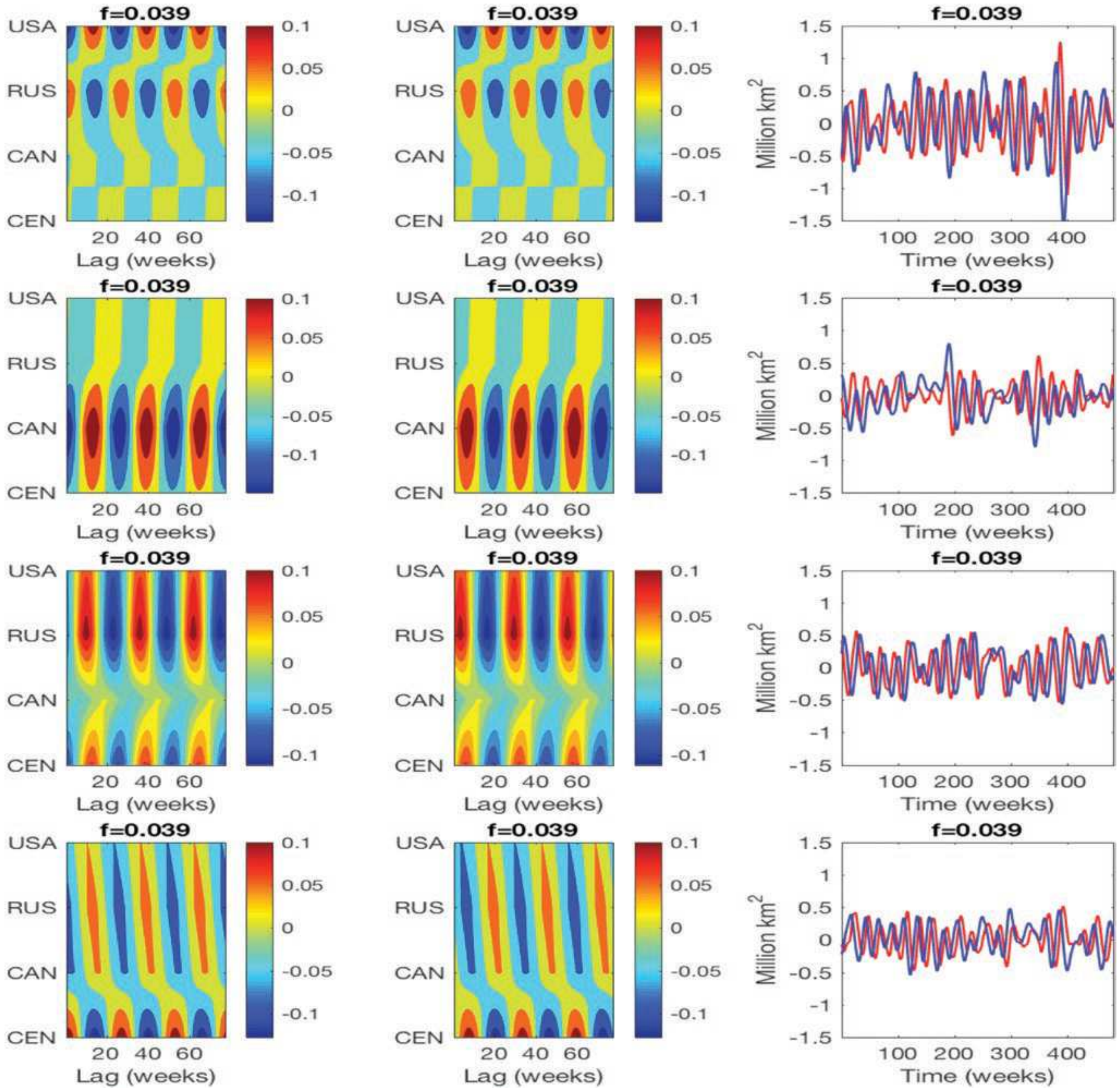

Figure 6. Same as in Figure 5, but for all DAHM pairs at a given frequency, ranked top-to-bottom in the DAH power spectrum, i.e. with $|\lambda_j|$ decreasing.

of Figures 5 and 6; time series that furthermore exhibit for each frequency, amplitude modulations and a shift of roughly a quarter of their associated period. The DAHCs are thus nearly in phase quadrature due to the sufficiently long window $M' = 77$ that help resolve here sufficiently well the decay of correlations. As mentioned earlier, this latter feature constitutes a key factor for a successful application of the DAH–MSLM approach for prediction purposes (see Section 4).

The DAH–MSLM model was trained on the full available 2006–2016 data for an estimation of $3 + 4(d - 1) = 15$ coefficients for the main layer of Eqns (MSLM) and of $12 + 4(d - 1) = 24$ coefficients for the second layer. The estimation of these coefficients is made from the DAHC pair of interest obtained by projection of the original data set according to (3.14), resulting thus into $2N' = 964$ DAH-processed observations over the full 11-year time interval. The model was then integrated from initial conditions corresponding to the first week of 2006 and forced by 10 different realizations of the white noise to obtain an ensemble of simulated DAHCs. The latter are convolved with DAHMs according to (3.16)–(3.17) and added first across frequency bands, and then across Arctic regions to obtain a simulated total MASIE anomaly time series whose length is equal to that of the observed one.

Given the limited amount of data available, Figure 7 shows good DAH–MSLM modelling skill in reproducing the statistics of the total MASIE anomaly, in terms of both ACFs and probability density functions (PDFs). Both the ACF decay and the global PDF shape are reasonably well reproduced by the ensemble mean (blue curves in Fig. 7), while 10 individual ensemble members necessarily differ, as expected, due to the stochastic nature of the DAH–MSLM. The PDF of the negative values are, however, underestimated by the DAH–MSLM model, possibly due to external ocean-atmosphere factors that are not included in the present data-driven modelling.

5.2. Prediction of September SIE

To calibrate and assess the predictive skill of DAH–MSLM for submission into the SIE September Outlook (Stroeve *et al.*, 2015) (SIO), retrospective forecast experiments were made consistent with the adopted SIO protocols. Given the short record of available observations and the need to estimate cross-correlation coefficients within the embedding window width M' (see (3.2)), retrospective forecasts were restricted to the 2013–2016 time interval.

First, calendar weeks 35–38 that typically align with September are taken to represent the prediction target. The DAH–MSLM model given by Eqns (MSLM) is trained on the MASIE data set from its beginning in 2006 through the 23rd, 27th and 31st calendar week in 2013–2016, when June, July and August predictions are solicited by SIPN, respectively. In particular, each time when a prediction is made, the seasonal cycle and MASIE anomalies are recalculated,

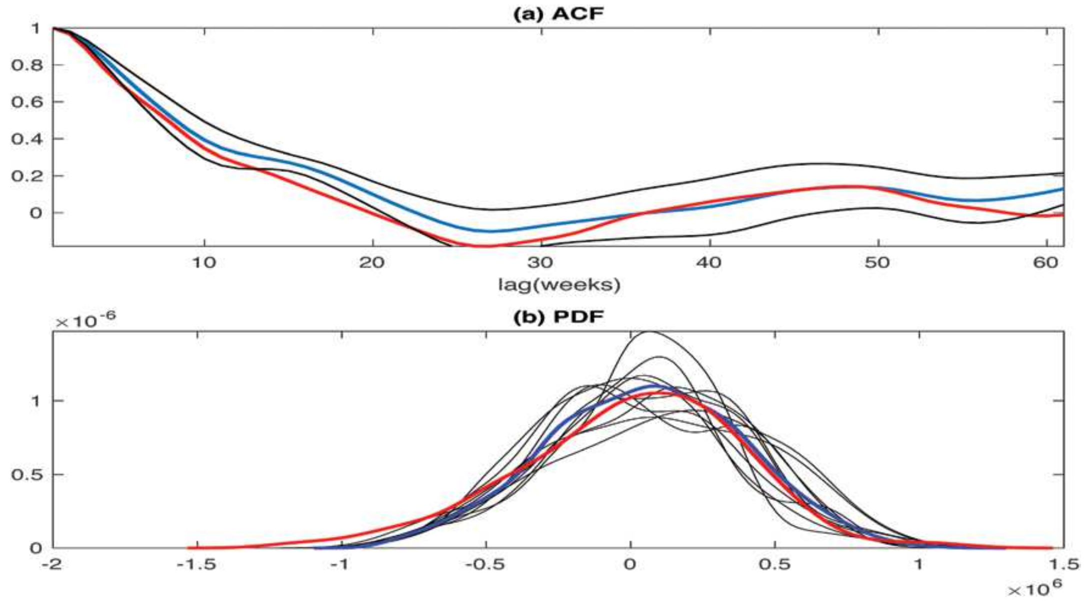

Figure 7. Statistical properties of the total MASIE anomaly as simulated by DAH–MSLM approach: Red indicates observations and blue indicates ensemble mean of 10 stochastic simulations. (a) Autocorrelation function (ACF); here, the black lines indicate the standard deviation of individual ensemble members. (b) Probability density function (PDF); here, the black lines indicate individual ensemble members.

and we performed the DAH decomposition of the latter, by using only data available until the time of the start of the forecast, i.e. our retrospective experiments emulate real-time predictions and are strictly with no look-ahead.

We further optimize M' by assessing the resulting skill of the retrospective forecasts. Our optimal DAH–MSLM according to the stopping criterion of Kondrashov *et al.* (2015) includes only one layer to model the noise $(\epsilon_j^x, \epsilon_j^y)$, as in Eqns (MSLM), and the results presented below were obtained with the optimal embedding window of $M' = 77$ weeks.

A 100-member ensemble of predictions of DAHCs is obtained from the stochastic DAH–MSLM model initialized in the 23rd, 27th and 31st week of each year in the period 2013–2016, and integrated forward for $K = 15$, 11 and 7 weeks ahead, to respond to the June, July and August solicitations, respectively. The predicted DAHCs are then convolved with the corresponding DAHMs to yield a stochastic ensemble of predictions of the HRCs at each frequency.

The HRC ensemble members corresponding to a particular noise realization are (i) added across the frequencies to yield ensemble predictions of the MASIE anomaly in each of the four regions; (ii) the sum of the associated ensemble means over the four regions yields the total MASIE anomaly; and (iii) the latter is added to the seasonal cycle. The last four calendar weeks, i.e. weeks 35–38 of the predictions, are then averaged to obtain September values for SIPN submission. The results for all four regions and for the total MASIE are displayed in Figure 8a–e.

The HRC predictions in several frequency bands are shown in Figures 10–13 for the total MASIE anomaly in each of the 4 years 2013–2016, respectively, and they are plotted for the forecast-starting months of June (green), July

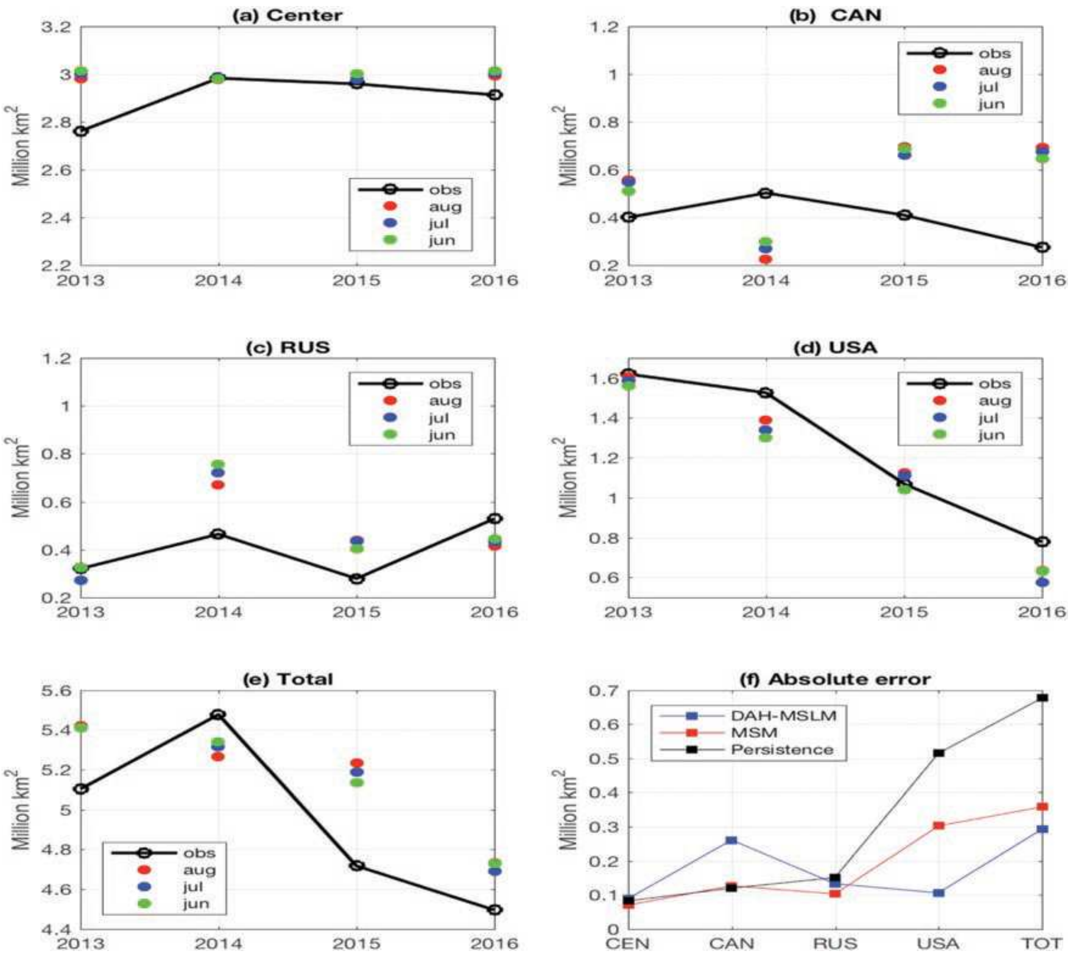

Figure 8. Retrospective DAH–MSLM predictions of observed (obs) September MASIE, according to the SIPN protocol, i.e. in June, July and August for the years 2013–2016. (a–d) Regional outlooks for the Center, CAN, RUS and USA regions, as defined in Section 2, and (e) total MASIE outlooks. (f) Prediction skill of averaged June–August outlooks versus persistence forecast from September of the previous year, as well as MSM forecasts shown in Figure 9.

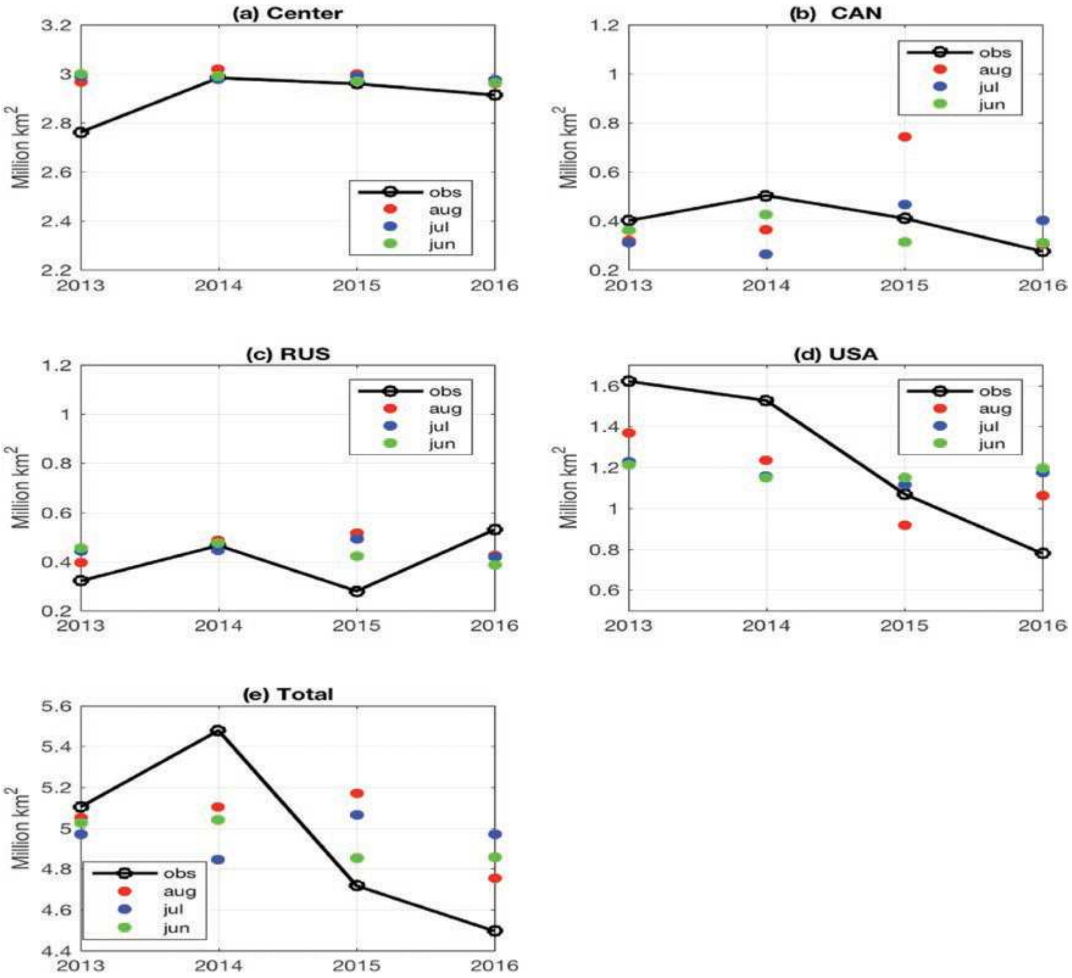

Figure 9. Linear MSM predictions of observed September MASIE, according to the SIPN protocol, i.e. in June, July and August for the years 2013–2016. see text for details.

(blue) and August (red). For most of the frequencies, the predictions of HRCs are fairly consistent and are close to the HRCs obtained from observations (black circles in Figs 10–13). The frequencies for which the HRC predictions are robust tend to change from year to year, but typically lie in the lower-frequency range.

However, for certain frequencies and depending on the month and the year, the predictions may deviate substantially from each other and/or from the observed values. This feature is visible for instance at $f = 0, 0.039, 0.079$ and 0.118 in 2015 (see Fig. 12), $f = 0.105$ in 2014 (see Fig. 11) and $f = 0, 0.118$ in 2013 (see Fig. 10). Such forecast sensitivity may be due to highly transient anomalies that develop during the melt season, possibly caused by external (and non-stationary) factors that are not taken into account in the MASIE data set, and thus not included in our modelling.

Figure 8e demonstrates that, as discussed above, the predicted total September MASIE values from ensemble mean—defined as the average of calendar weeks 35–38 and over the noise realizations—vary little for the June, July and August outlooks. The August forecasts are $\approx 4.7 \cdot 10^6$ km^2, $5.2 \cdot 10^6$ km^2, $5.3 \cdot 10^6$ km^2, and $5.4 \cdot 10^6$ km^2 in 2016, 2015, 2014 and 2013, respectively, while the corresponding observed MASIE values are $4.5 \cdot 10^6$ km^2, $4.72 \cdot 10^6$ km^2, $5.47 \cdot 10^6$ km^2 and $5.10 \cdot 10^6$ km^2. Note that the standard deviation of the DAH–MSLM ensemble predictions is here about $0.1 \cdot 10^6$ km^2, showing thus a narrow forecast model uncertainty for such short-term predictions.

For baseline prediction comparison with the DAH–MSLM results, we have used the following linear stochastic model with memory obtained by MSM framework (Kondrashov *et al.*, 2015):

$$\begin{aligned} \dot{\mathbf{X}} &= \mathbf{A}\mathbf{X} + \mathbf{r}_t^0, \\ \dot{\mathbf{r}}_t^0 &= \mathbf{L}[\mathbf{X}, \mathbf{r}_t^0] + \Sigma \dot{W}_t \,. \end{aligned} \tag{5.1}$$

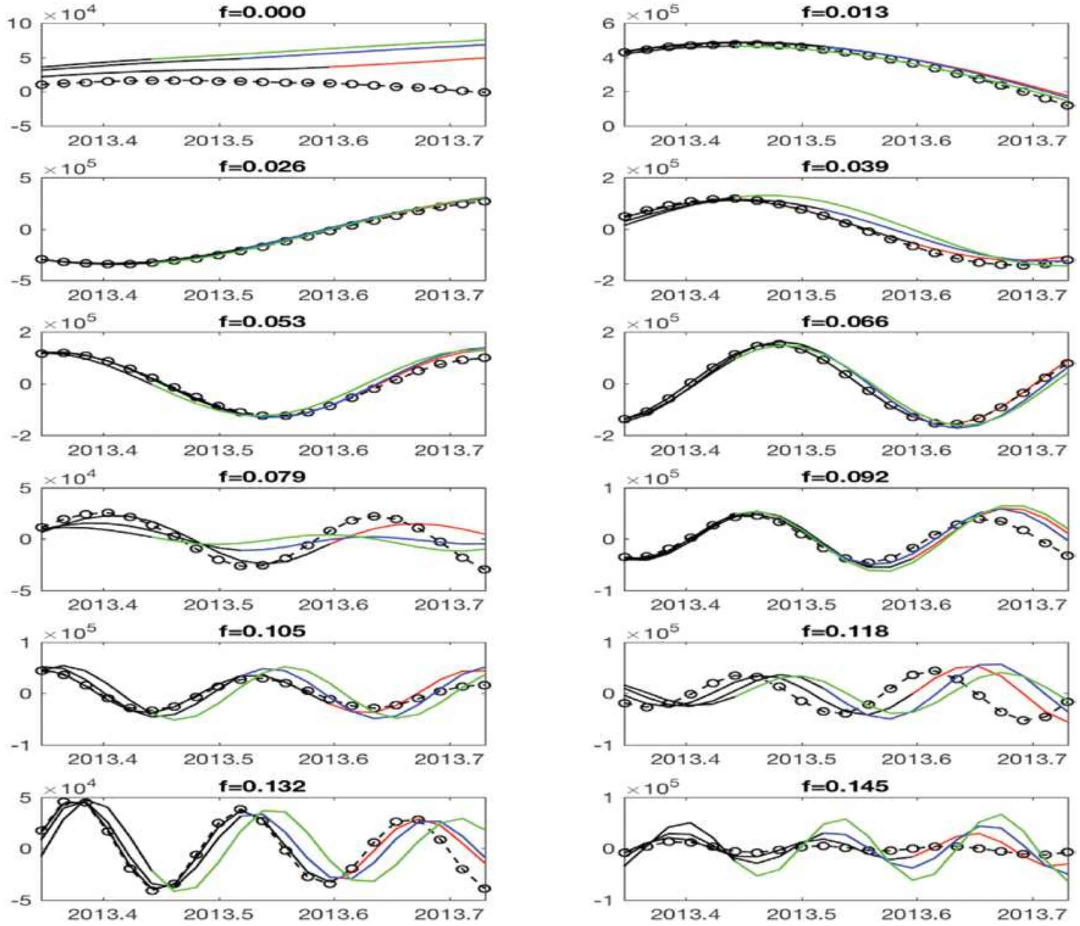

Figure 10. Predicted harmonic reconstructed components (HRCs) for total MASIE anomaly in 2013, i.e. $\sum_{k=1}^{d} R_k(t;f)$: Thin black lines indicate sum over the channels of HRCs given by (3.17) prior to the start of the forecasts and then extended by June (green), July (blue) and August (red) predictions; black circles indicate sum over the channels of HRCs using the data through the end of September 2013 and representing prediction target in the forecast interval.

Here, $\mathbf{X}$ are the anomaly MASIE time series in four Arctic regions, $\mathbf{A}$ and $\mathbf{L}$ are the model coefficients obtained by consecutive regressions on the main and auxiliary model layer, respectively; $\mathbf{r}_t^0$ is a hidden variable aimed at simulating the regression residual of the main model layer, while W_t denotes a four-dimensional Brownian motion and Σ a real, symmetric and positive definite 4×4 matrix. The optimal number of additional model layers—here equal to one—has been determined by the stopping criterion described in Appendix A of Kondrashov *et al.* (2015).

The retrospective forecast experiments for 2013–2016 SIO were repeated using 300-member ensemble of stochastic realizations by MSM model, and corresponding predictions for ensemble mean are shown in Figure 9. In comparison with the DAH–MSLM predictions, the ones obtained by MSM exhibit larger spread between June, July and August outlooks of a given year for CAN and USA sectors, yielding thus larger variations in predictions of total MASIE, compare Figures 8e and 9e. Figure 8f compares average absolute error of DAH–MSLM predictions with the MSM forecasts, as well as with the baseline persistence forecasts using the September value from the previous year. These results show that both persistence and MSM are competitive with DAH–MSLM in RUS and CEN sectors; while being marginally better in CAN sector, MSM predictions show a larger spread than DAH–MSLM. However, the key USA sector that contributes the most to MASIE and exhibits a pronounced downward trend in September 2013–2016 is predicted by DAH–MSLM substantially better than persistence and MSM, resulting thus in overall better prediction for the total MASIE. The mean absolute error of the 2013–2016 retrospective DAH–MSLM forecasts is $0.28 \cdot 10^6\ \mathrm{km}^2$.

The regional difference in skill is difficult to interpret both due to shortness of the data set and complexity of SIE dynamics. On the other hand, CAN sector is lying mostly inland, while other three sectors are largely within the

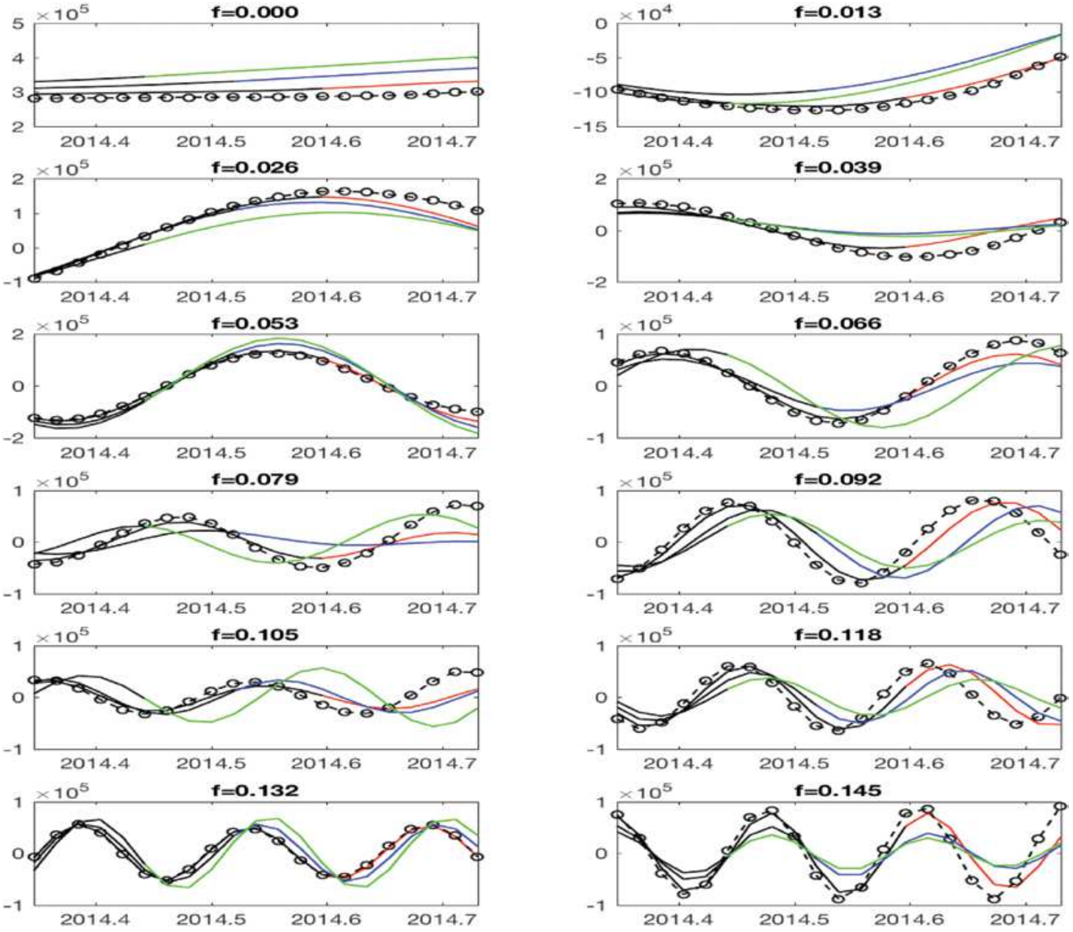

Figure 11. Predicted harmonic reconstructed components (HRCs) for total MASIE anomaly in 2014.

Arctic's open waters. Both larger absolute prediction errors for DAH–MSLM and higher seasonal spread for MSM outlooks in CAN sector than in other sectors, suggest the need to include additional ocean–atmosphere data in order to improve prediction of this inland sector, and it will be addressed in the future work.

6. Discussion

It is instructive to compare the 2016 real-time DAH–MSLM with other model predictions in SIPN plume (Stroeve *et al.*, 2015; Hamilton and Stroeve, 2016) by using post-season SIPN report (https://www.arcus.org/sipn/sea-ice-outlook/2016/post-season). The model-median September estimate by SIPN for August-based predictions across all methods was $\approx 4.4 \cdot 10^6$ km^2, while an interquartile range of ≈ 4.2–$4.7 \cdot 10^6$ km^2 barely includes observed value of $4.72 \cdot 10^6$ km^2 based on SII data set. On the other hand, the DAH–MSLM August prediction for the 2016 SIPN September Outlook was $4.79 \cdot 10^6$ km^2, when calibrated by 2013–2015 retrospective experiments, as described in the previous section. The DAH–MSLM outperformed not only the majority of ≈ 20 statistical models that overall predicted much lower values than observed one but also most physics-based models, even though median and interquartile range of the latter models outlooks were displaying better match with the observations.

We should note that independent analysis of past SIO predictions (Stroeve *et al.*, 2014; Hamilton and Stroeve, 2016) already found that (i) the SIO models typically have difficulty when observations depart from the long-term trend and that (ii) persistence forecast may be competitive in certain years—which is also the case for DAH–MSLM in the CAN sector. On the other hand, an apparent competitive advantage of DAH–MSLM seems to rely on the exhibited good prediction skills for the USA sector, leading to an overall improved forecasting for the whole Arctic.

Our prediction results should also be interpreted in the context of existing range of observational estimates for the observed SIE that reflect different algorithms to derive sea ice concentration from the satellite measurements,

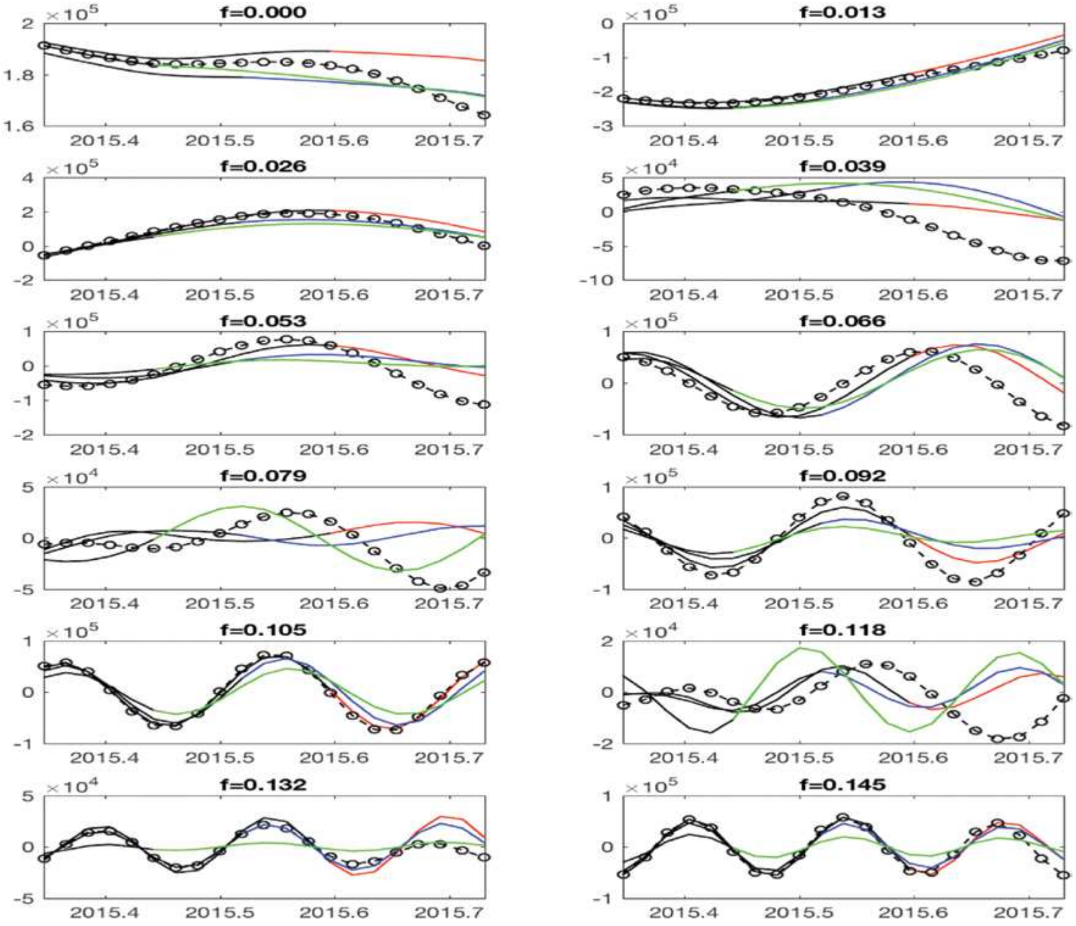

Figure 12. Predicted harmonic reconstructed components (HRCs) for total MASIE anomaly in 2015.

for example MASIE versus SII data set (Cavalieri *et al.*, 1996). The range of the SIE values resulting from different algorithms provides an indication of observational uncertainty that can be as large as $0.4 \cdot 10^6$ km^2. With the exception of 2015 when error in CAN sector dominates, MASIE-based DAH–MSLM retrospective forecasts are very close—within $0.1 \cdot 10^6$ km^2 to observed SII index in 2013, 2014 and 2016, thus confirming the potential value of such predictions for SIO. Preliminary results of 2017 SIO submissions show that the corresponding DAH–MSLM forecasts were also fairly accurate. In particular, the DAH–MSLM model predicted September SIE values of $4.67 \cdot 10^6$ km^2 and $4.5 \cdot 10^6$ km^2 in June and August, respectively, while observed SIE reached an annual 2017 minimum of $4.64 \cdot 10^6$ km^2 on September 13; see http://nsidc.org/arcticseaicenews/2017/09/arctic-sea-ice-at-minimum-extent-2/.

To summarize, we developed a dynamical statistical model to forecast September MASIE by adopting the recent data-driven DAH–MSLM approach of (Chekroun and Kondrashov, 2017) in a predictive context. By using the joint data set of MASIE anomalies of four key Arctic regions, we computed DAH modes that account for regional ‘interactions’ and allow for the extraction of DAH coefficients that can be efficiently modelled within the class of MSLMs. These stochastic models, as particular type of MSM models (Kondrashov *et al.*, 2015), allow for providing non-Markovian data-driven emulators that are here able to account for the memory effects intrinsic to SIE dynamics.

In turn, these features allow us to successfully model and predict complicated regional ‘interactions’ by a set of MSLMs coupled per frequency only by the same noise realization. The resulting modelling approach is computationally efficient, taking e.g. 15 seconds of CPU time on a 4-core 2.9 GHz Intel Core i7 MacBook Pro laptop to compute monthly prediction for SIO. The successful application of the DAH–MSLM approach, despite an intrinsic difficulty of the relatively short observational record used here, makes it very appealing for other geoscience applications.

Future extensions of this study will focus on key Arctic subregions of the USA such as Alaska. The SIE prediction approach presented in this article may also benefit from other SIE data sets as well as information about

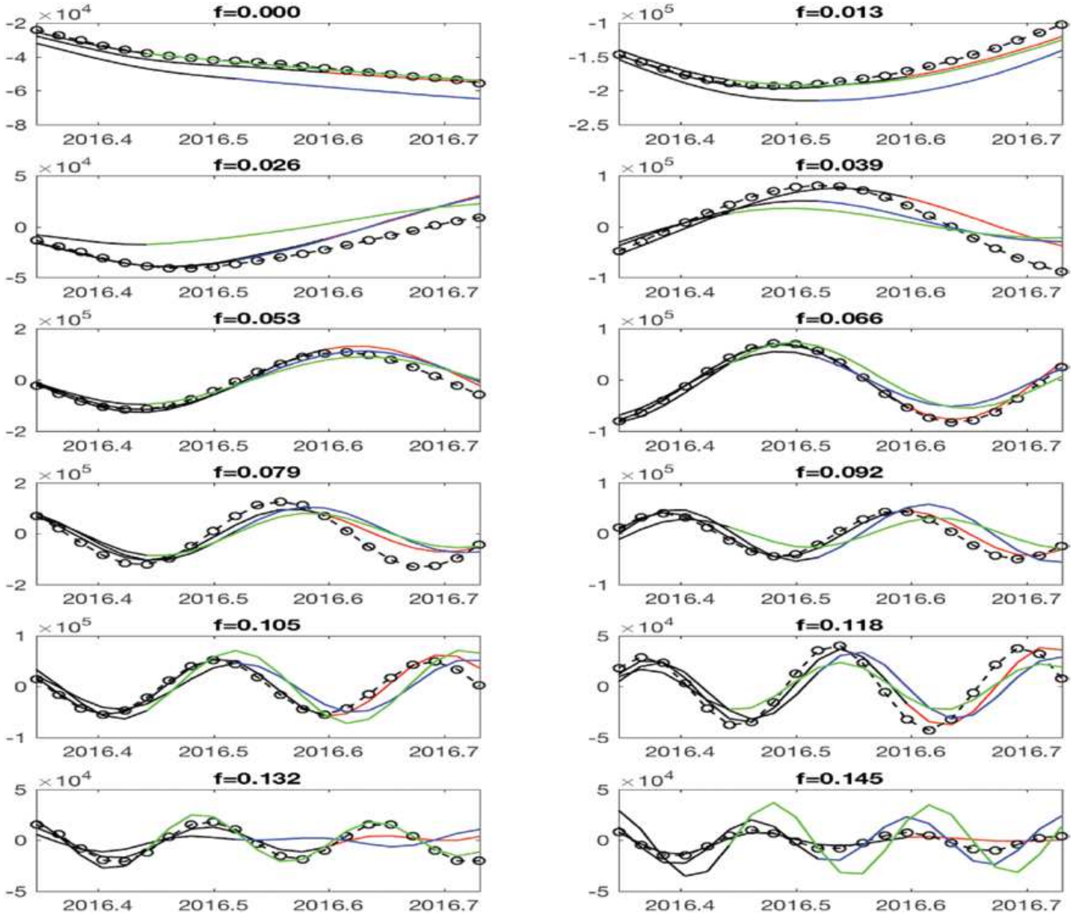

Figure 13. Predicted harmonic reconstructed components (HRCs) for total MASIE anomaly in 2016.

ocean-atmosphere conditions such as available from meteorological and oceanographic data, that may allow, in particular, to improve prediction for the inland CAN sector. DAH–MSLM approach has been already successfully applied to stochastic modelling of sea ice concentration data (Kondrashov *et al.*, 2018). Future work should also pay greater attention to the timing of retreat and advance of the Arctic sea ice (Sigmond *et al.*, 2016).

Declaration

Funding: Multidisciplinary University Research Initiative (MURI) of the Office of Naval Research (N00014-16-1-2073); National Science Foundation (OCE-1243175, OCE-1658357 and DMS-1616981).
Ethical approval: none.
Conflict of interest: none.

Acknowledgements

The authors are grateful for the constructive reviewers' comments that help improve the presentation of the results. The MASIE data used for this study are available at http://nsidc.org/data/masie/.

References

Barnston AG, Tippett MK, Heureux ML, Li S, DeWitt DG. Skill of real-time seasonal ENSO model predictions during 2002–2011: is our capability improving? *Bull Amer Meteorol Soc* 2012; 93: 631–51.

Battisti DS, Hirst AC. Interannual variability in a tropical atmosphere-ocean model: influence of the basic state, ocean geometry and nonlinearity. *J Atmos Sci* 1989; 46: 1687–712.

Bhattacharya K, Ghil M, Vulis I. Internal variability of an energy-balance model with delayed albedo effects. *J Atmos Sci* 1982; 39: 1747–73.

Blanchard-Wrigglesworth E, Cullather RI, Wang W, Zhang J, Bitz CM. Model forecast skill and sensitivity to initial conditions in the seasonal sea ice outlook. *Geophys Res Lett* 2015; 42: 8042–48.

Bushuk M, Giannakis D, Majda AJ. Arctic sea ice reemergence: the role of large-scale oceanic and atmospheric variability. *J Clim* 2015; 28: 5477–509.

Cavalieri D, Parkinson C, Gloersen P, Zwally HJ. *Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data, 1979–2015*. Technical report. NASA National Snow and Ice Data Center, Distributed Active Archive Center, 1996 (updated yearly).

Chekroun MD, Ghil M, Liu H, Wang S. Low-dimensional Galerkin approximations of nonlinear delay differential equations. *Disc Cont Dyn Sys A* 2016; 36: 4133–77.

Chekroun MD, Ghil M, Neelin JD. Pullback attractor crisis in a delay differential ENSO model. In: Tsonis A (ed.). *Advances in Nonlinear Geosciences*. 2018. Springer, 1–33.

Chekroun MD, Kondrashov D. Data-adaptive harmonic spectra and multilayer Stuart-Landau models. *Chaos* 2017; 27: 093110.

Chekroun MD, Kondrashov D, Ghil M. Predicting stochastic systems by noise sampling, and application to the El Niño-Southern oscillation. *Proc Natl Acad Sci USA* 2011a; 108: 11766–71.

Chekroun MD, Liu H, McWilliams JC. The emergence of fast oscillations in a reduced primitive equation model and its implications for closure theories. *Comput Fluids* 2017; 151: 3–22.

Chekroun MD, Liu H, Wang S. *Stochastic Parameterizing Manifolds and Non-Markovian Reduced Equations: Stochastic Manifolds for Nonlinear* SPDEs II. New York, NY: Springer Briefs in Mathematics, Springer, 2015.

Chekroun MD, Neelin JD, Kondrashov D, McWilliams JC, Ghil M. Rough parameter dependence in climate models: The role of Ruelle-Pollicott resonances. *Proc Natl Acad Sci USA* 2014; 111: 1684–1690.

Chekroun MD, Simonnet E, Ghil M. Stochastic climate dynamics: random attractors and time-dependent invariant measures. *Phys D* 2011b; 240: 1685–700.

Chen C, Cane MA, Henderson N *et al*. Diversity, nonlinearity, seasonality and memory effect in ENSO simulation and prediction using empirical model reduction. *J Clim* 2016; 29: 1809–30.

Chorin AJ, Hald OH. *Stochastic Tools in Mathematics and Science*. No. 147 in Surveys and Tutorials in the Applied Mathematical Sciences. New York, NY: Springer, 2006.

Chorin AJ, Hald OH, Kupferman R. Optimal prediction with memory. *Phys D* 2002; 166: 239–57.

Comiso JC. *Bootstrap Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS, Version 2 [Northern Hemisphere Daily Data]*. Digital media. Boulder, CO: NASA National Snow and Ice Data Center, Distributed Active Archive Center, 2014.

Crommelin D, Vanden-Eijnden E. Subgrid-scale parameterization with conditional Markov chains. *J Atmos Sci* 2008; 65: 2661–75.

Day J, Hawkins E, Tietsche S. Will Arctic sea ice thickness initialization improve seasonal forecast skill? *Geophys Res Lett* 2014; 41: 7566–75.

Engel KJ, Nagel R. *A Short Course on Operator Semigroups*. Springer Science & Business Media, Springer New York, 2006.

Fetterer F, Savoie M, Helfrich S, Clemente-Colón P. *Multisensor Analyzed Sea Ice Extent - Northern Hemisphere*. Technical report. Boulder, CO: National Snow and Ice Data Center, 2010.

Franzke CLE, O'Kane TJ, Berner J, Williams PD, Lucarini V. Stochastic climate theory and modeling. *Wiley Interdiscipl Rev Clim Change* 2015; 6: 63–78.

Ghil M, Allen MR, Dettinger MD *et al*. Advanced spectral methods for climatic time series. *Rev Geophys* 2002; 40: 1003.

Giannakis D, Majda AJ. Nonlinear Laplacian spectral analysis for time series with intermittency and low-frequency variability. *Proc Natl Acad Sci USA* 2012; 109: 2222–7.

Hamilton LC, Stroeve J. 400 predictions: the SEARCH Sea Ice Outlook 2008–2015. *Polar Geogr* 2016; 39: 274–87.

Harlim J, Mahdi A, Majda A. An ensemble Kalman filter for statistical estimation of physics constrained nonlinear regression models. *J Comput Phys* 2014; 257 (part A): 782–812.

Jin FF, Neelin JD, Ghil M. El Niño on the devil's staircase: annual subharmonic steps to chaos. *Science* 1994; 274: 70–2.

Jung T, Gordon GD, Bauer P et al. Advancing polar prediction capabilities on daily to seasonal time scales. *Bull Am Meteorol Soc* 2016; 97: 1631–47.

Kondrashov D, Berloff P. Stochastic modeling of decadal variability in ocean gyres. *Geophys Res Lett* 2015; 42: 1543–53.

Kondrashov D, Chekroun MD, Ghil M. Data-driven non-Markovian closure models. *Phys D* 2015; 297: 33–55.

Kondrashov D, Chekroun MD, Robertson AW, Ghil M. Low-order stochastic model and "past-noise forecasting" of the Madden-Julian oscillation. *Geophys Res Lett* 2013; 40: 5305–10.

Kondrashov D, Chekroun MD, Yuan X, Ghil M. Data-adaptive harmonic decomposition and stochastic modeling of Arctic sea ice. In: Tsonis A (ed.). *Advances in Nonlinear Geosciences*. Springer, 2018, 179–205.

Kondrashov D, Kravtsov S, Robertson AW, Ghil M. A hierarchy of data-based ENSO models. *J Clim* 2005; 18: 4425–44.

Kravtsov S, Kondrashov D, Ghil M. Multi-level regression modeling of nonlinear processes: derivation and applications to climatic variability. *J Clim* 2005; 18: 4404–24.

Kravtsov S, Kondrashov D, Ghil M. Empirical model reduction and the modeling hierarchy in climate dynamics and the geosciences. In: Palmer TN, Williams P (eds). *Stochastic Physics and Climate Modeling*. Cambridge University Press, Cambridge, 2009, pp. 35–72.

Lu F, Lin KK, Chorin AJ. Data-based stochastic model reduction for the Kuramoto–Sivashinsky equation. *Phys D* 2017a; 340: 46–57.

Lu F, Tu X, Chorin AJ. Accounting for model error from unresolved scales in ensemble Kalman filters by stochastic parametrization. *Mon Weather Rev* 2017b; 2017; 145: 3709–3723.

Majda AJ, Harlim J. Physics constrained nonlinear regression models for time series. *Nonlinearity* 2013; 26: 201–17.

Meier WN, Fetterer F, Stewart JS, Helfrich S. How do sea-ice concentrations from operational data compare with passive microwave estimates? Implications for improved model evaluations and forecasting. *Ann Glaciol* 2015; 56: 332–40.

Neelin JD, Battisti DS, Hirst AC et al. ENSO theory. *J Geophys Res* 1998; 103: 14261–90.

Newman M. An empirical benchmark for decadal forecasts of global surface temperature anomalies. *J Clim* 2013; 26: 5260–9.

Penland C. A stochastic model of IndoPacific sea surface temperature anomalies. *Phys D* 1996; 98: 534–58.

Roques L, Chekroun MD, Cristofol M, Soubeyrand S, Ghil M. Parameter estimation for energy balance models with memory. *Proc R Soc A* 2014; 470: 20140349.

Selivanov A, Lehnert J, Dahms T *et al*. Adaptive synchronization in delay-coupled networks of Stuart-Landau oscillators. *Phys Rev E* 2012; 85: 016201.

Sigmond M, Reader MC, Flato GM, Merryfield WJ, Tivy A. Skillful seasonal forecasts of Arctic sea ice retreat and advance dates in a dynamical forecast system. *Geophys Res Lett* 2016; 43: 12457–65.

Stroeve J, Blanchard-Wrigglesworth E, Guemas V *et al*. Improving predictions of Arctic sea ice extent. *Eos* 2015; 96: 11–5.

Stroeve J, Hamilton LC, Bitz CM, Blanchard-Wrigglesworth E. Predicting September sea ice: ensemble skill of the SEARCH Sea Ice Outlook 2008–2013. *Geophys Res Lett* 2014; 41: 2411–8.

Strounine K, Kravtsov S, Kondrashov D, Ghil M. Reduced models of atmospheric low-frequency variability: parameter estimation and comparative performance. *Phys D* 2010; 239: 145–66.

Suarez MJ, Schopf PS. A delayed action oscillator for ENSO. *J Atmos Sci* 1988; 45: 3283–7.

Tziperman E, Stone L, Cane MA, Jarosh H. El Niño chaos: overlapping of resonances between the seasonal cycle and the Pacific ocean-atmosphere oscillator. *Science* 1994; 264: 72–4.

Venkataramani SC, Venkataramani RC, Restrepo JM. Dimension reduction for systems with slow relaxation. *J Stat Phys* 2017; 167: 892–933.

Wang L, Yuan X, Ting M, Li C. Predicting summer Arctic sea ice concentration intraseasonal variability using a vector autoregressive model. *J Clim* 2016; 29: 1529–43.

Wilks DS. Effects of stochastic parametrizations in the Lorenz'96 system. *Q J R Meteorol Soc* 2005; 131: 389–407.

Wouters J, Lucarini V. Multi-level dynamical systems: connecting the Ruelle response theory and the Mori-Zwanzig approach. *J Stat Phys* 2013; 151: 850–60.

Zakharova A, Loos S, Siebert J *et al*. Controlling chimera patterns in networks: interplay of structure, noise, and delay in control of self-organizing nonlinear systems. In: Schöll PHE, Klapp SHL (eds). *Control of Self-organizing Nonlinear Systems*. Berlin, Germany: Springer, 2016, pp. 35–72.