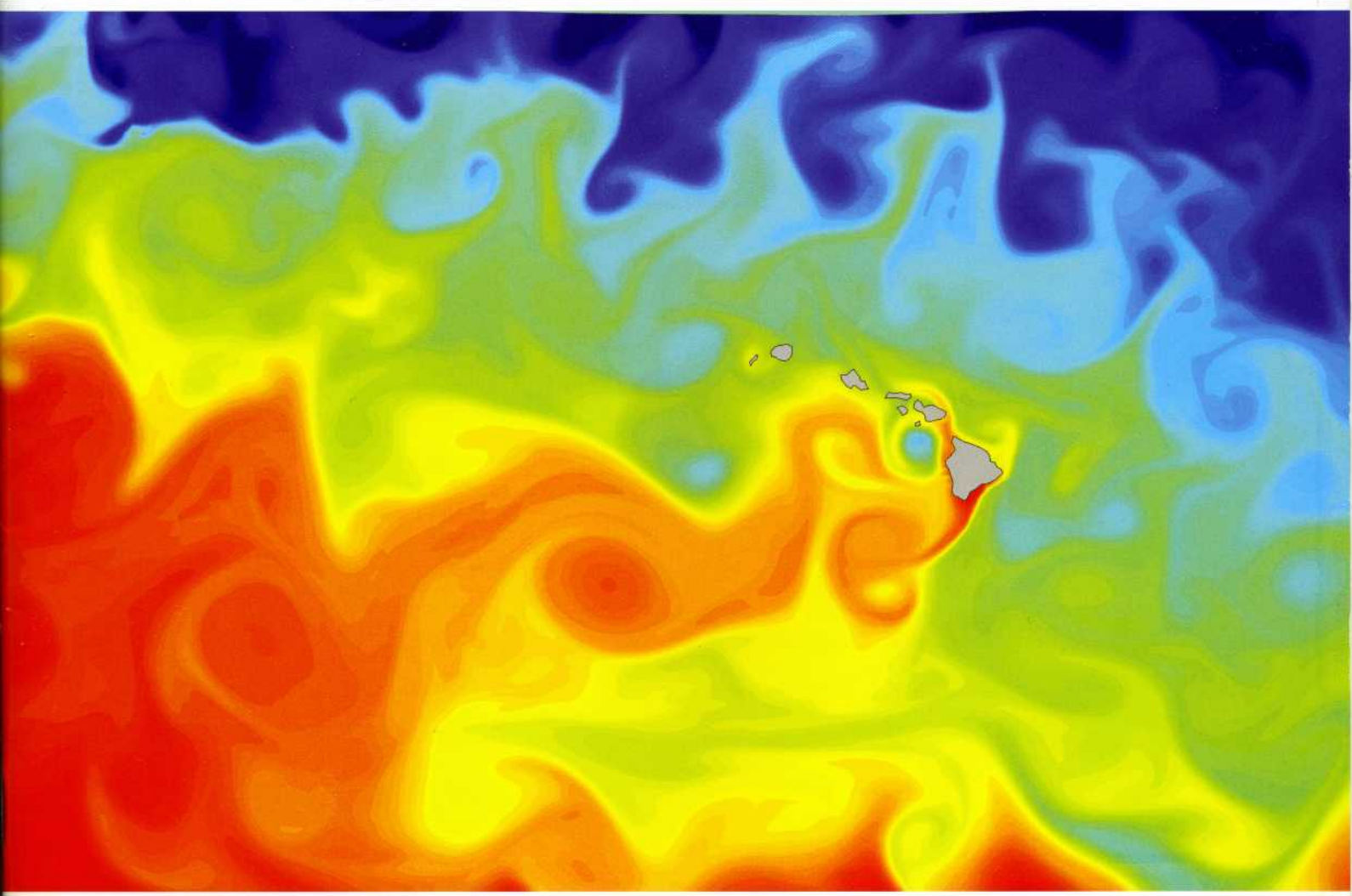
Journal of Operational Oceanography

Volume 5 No.2
August 2012
ISSN 1755-876X

Proceedings of The Institute of Marine Engineering, Science and Technology



Publications
IMAREST

Contents

Journal of Operational Oceanography

Proceedings of The Institute of Marine Engineering, Science and Technology Vol. 5 No.2 2012

Publications

IMAREST

Published by The Institute of Marine Engineering, Science and Technology
Aldgate House ▪ 33 Aldgate High Street ▪ London ▪ EC3N 1EN ▪ UK
Tel: +44(0)20 7382 2600 ▪ Fax: +44(0)20 7382 2670
www.imarest.org

Registered Charity No 212992
Founded 1889. Incorporated by Royal Charter 1933. Licensed body of the Engineering Council (UK) and the Science Council

Europe ▪ ANSPAC ▪ North East Asia ▪ South East Asia ▪ Americas ▪ Middle East

Application of a vectorial-algebraic method for investigation of spatial-temporal variability of sea ice drift and validation of model calculations in the Arctic Ocean

VA Volkov, NE Ivanov† and DM Demchev*†*
**Nansen International Environmental and Remote Sensing Center (NIERSC), St Petersburg, Russia*
†Arctic and Antarctic Research Institute (AARI), St Petersburg, Russia

This paper considers the theoretical bases of the vectorial-algebraic method developed and applied in Russia for the time-series analysis of vector values – wind, currents and ice drift – and presents examples of the analysis of measurement data. The vectorial-algebraic approach allows the initial information to be significantly compressed and adequately describe the vector time-series of full-scale and model data restricted by a set of statistical characteristics in the invariant form. The methodology was used in the MyOcean Project (FP7).

AUTHORS' BIOGRAPHIES

Dr Vladimir A Volkov is an Oceanographer who graduated from St Petersburg State University in 1970 and gained his PhD in 1984 from the Arctic and Antarctic Research Institute, St Petersburg. In 1970 he began his academic career in AARI as a young scientist, and from 1984 to 2000 he was head of the sea dynamics laboratory. From 2000 he is Scientific Secretary/Research Director – Head of MetOcean Group in the 'Nansen-Center' Foundation, St Petersburg.

Nikolay Ye Ivanov is an Oceanographer who graduated from Russian State Hydrometeorological University, St Petersburg in 1980, and worked in the 'Roshydromet' system, in particular, in the St Petersburg department of the State Oceanographic Institute. From 2006 he has been a scientist at AARI. His area of scientific activity includes methodology of the full-scale analysis of metocean data, climatology and atmosphere-ocean interaction problems.

Denis M Demchev is a marine engineer who graduated from Russian State Hydrometeorological University, St Petersburg in 2007. He is currently leading engineer/PhD student at AARI and young scientist (part-time) in the 'Nansen-Center' Foundation.

INTRODUCTION

Ice drift is an important element of the oceanographic regime. The spatial structure of the drift field and its temporal variability reflect the processes occurring both in the atmosphere and the hydrosphere. Drift variability at prolonged time intervals is inseparably connected with global climatic changes and changes in total and partial sea ice concentration in the Arctic Ocean (AO).

As studies at high latitudes of the Arctic developed, ice drift was estimated from indirect indications – visual observations, 'bottle mail', etc, and on the basis of direct measurements. The latter included, in particular, instrumental visual observations in the coastal zone, ice drift determination by means of astronomical observations, current meters deployed from drifting ice floes in deep water layers of low mobility, and by using drifting radio-markers and other methods. These data allowed only general descriptions of the ice drift features, since they were not made over a long enough time, or had poor accuracy or low spatial resolution.

Nowadays, direct drift measurements are carried out at a higher technical level than before, by means of buoys equipped with a high-precision satellite system for coordinate determination, and with a higher time frequency than before. The indirect methods for drift estimates, based on processing sensor data from spacecraft, for example AMSR-E onboard the AQUA satellite measuring brightness temperatures, have a high precision for determining drift parameters, and an enormous spatial coverage providing sufficiently high temporal and, especially, spatial resolution. The parameters of these data allow the drift to be investigated within a wide range of temporal and spatial scales.

It is obvious that the most accurate are the drift measurements by modern drifting buoys, which can be considered as 'reference' buoys for validation of both the model calculations and the drift data, based on processing the satellite images. A dataset of drifting buoys, including data from buoys deployed at manned drifting stations and onboard ships, has been generated from 1978 onwards. The operation times of such buoys can be from several months to several years, and the measurement frequency is up to several times a day. The number of buoys operating at present in the Arctic is several dozen (around 50 as of 8 February 2011).[(a)]

The vector drift fields derived by processing the AMSR-E sensor data are presented at regular grid points with a step of 32.5km over the entire Arctic Ocean (AO) area, where sea ice was present in the wintertime. These data, with a two-day time interval prepared by the IFREMER Institute (France), are available for the period 1 October 2002 through to the present time and can be accessed via the Internet.[(b)] Estimates of the accuracy of such fields calculated by the developers show that this dataset allows a sufficiently adequate assessment of ice dynamics in the Arctic Basin.[(c)]

There is also a different version of the global climate dataset of the ice drift fields – Pathfinder[(d)] – which was prepared using satellite images in the visible and IR regions (AVHRR, SMMR, SSMI), and buoy data (IAPB), that were interpolated to a regular grid with a step of 25km. According to the studies of the Southern Ocean area[1] this dataset can be used for estimating the ice drift at a climatic scale.

The drift data are also obtained from model calculations and predictions. These data require validation – comparison with data of direct measurements or calculation results based on satellite observation data. The model and prognostic drift fields are prepared on a regular basis. One example of such data is that of the prognostic TOPAZ system (Nansen-Center, Bergen, Norway)[(e)] that provides a daily forecast for three days ahead.

At present, researchers have at their disposal a vast amount of data on the ice drift in the Arctic Ocean (AO). Analysis of this data requires physically justified methods, which then allows the information to be significantly compressed and adequately describe the regularities of natural processes with a restricted set of parameters.

The authors faced the problem of validation and analysis of a large amount of information on the vector fields in the MyOcean Project (FP7) and attempted to adapt the so-called vectorial-algebraic approach[2] to the project objectives.

Since the ice drift (similar to currents and wind) is often considered as a vector probabilistic process, and is described by a set of two-dimensional vectors with a specific time step, the choice of the mathematical vector model is the basic problem in selecting a methodology for analysing the ice drift or other vector series.

The applied studies often use a vector presentation in the form of components – projections to the orthogonal axes – as a complex number, and in the form of the motion vector, which has a module (vector magnitude) and a direction, which is added to similar vectors by the parallelogram rule. Since the drift (or a current/wind) is a progressive motion of ice (water, air) masses, an approach to the analysis of such processes through a displacement vector, and the analysis method based on the vector algebra laws, turns out to be most valid. It is been shown[3] that using a component-wise, complex-valued presentation of vectors, or applying a method of rotational components for the analysis of vector processes, it is possible to obtain an illusionary, non-adequate understanding about the investigated natural process. Only the vectorial-algebraic method yields physically-substantiated conclusions in the analysis of full-scale data, because it strictly takes into account the vector character of the process and presents the results of time series analysis by a compact set of scalar characteristics in the invariant form, ie, not dependent on the choice of the coordinate system.

It should be noted that the bases of the vectorial-algebraic approach to the analysis of vector series were developed quite a long time ago, but so far this method is used infrequently. This method has, however, the status of a standard method in the framework of the Russian Uniform Information System on the Global Ocean – ESIMO.[(f)] This method appears little known outside Russia, therefore this paper is an attempt to make it more accessible and understandable for specialists in the area of oceanography and meteorology, the more so because it is still being improved by a group of scientists from St Petersburg.

The description of the methodology of the vectorial-algebraic probabilistic analysis is illustrated by the examples of processing the time series of the ice drift speeds in the Arctic Ocean, obtained from satellite images,[(b),(d)] for the points with fixed coordinates and averaged for a day and a calendar month. The test calculations were performed for the series of mean daily drift values at three points for the period 4 October 2007 to 30 April 2008 (the series length is about half a year) and for six points with mean monthly data from January 1979 to December 2006 (the series length is 28 years); locations of these points is given on the maps in Fig 1Aa and Fig 2a.

METHODOLOGY OF THE ANALYSIS AND DISCUSSION OF THE EXAMPLES OF APPLICATION

Data analysis is performed using models of a vector random value, a vector stationary random process, a mathematical expectation trend and a system of two connected vector random values.

Since wind, sea currents and the ice drift present a directed mass transfer, their speeds are physical vectors. Therefore, the Euclidean vector $\vec{V}$ with Cartesian projections V_x, V_y, has been assumed as a speed model – a mathematical object

characterised by module $|\vec{V}|$ and direction φ, determined as an addition by the rule of parallelogram, rules of transformation of vector components at transfer to a new coordinate system and three types of vector multiplying – scalar, vector (oblique) and tensor.[4] The vectorial-algebraic method based on these provisions[2] allows data to be considered in terms of a vector random value, a vector random process, a system of connected vector random values and a vector spatial-temporal random field. Since the characteristics of vectors are multidimensional (vectors and tensors), not only are the calculation procedures discussed, but also the methods of digital and graphical presentation of the results.

Model of a vector random value

In the model of a vector random value (RV), the drift speed values $\vec{V}$ at the fixed points are represented by statistical estimates of the distribution of probabilities $f(\bullet)$ and its moments.[2,5] A complete probabilistic characteristic of stochastic vector $\vec{V}$ as an RV is a probability distribution in a polar coordinate system (by module and direction) and corresponding marginal distributions of probabilities. The corresponding recurrences $P(V,\varphi)$ are determined as probabilities of values $\vec{V}$ being in the prescribed gradation $f(\bullet)$:

$$\begin{aligned} P(V,\varphi) &= f\{V_1 \le V < V_2, \varphi_1 \le \varphi, \varphi_2\}, \\ P(V) &= f\{V_1 \le V < V_2\}, \\ P(\varphi) &= f\{\varphi_1 \le \varphi < \varphi_2\} \end{aligned} \qquad (1)$$

These characteristics are traditionally represented by a table of two-dimensional and marginal recurrences by rhumbs and module gradations (the calm recurrence is presented as an independent gradation). Such representation, being quite useful due to large information content, can be used only for a very restricted set of spatial points in the typical seasons because of a large multidimensionality. So, by identifying 16 gradations by direction and 10 gradations by the drift speed module for each point and for each season, a table is obtained of two-dimensional recurrence of $f(V,\varphi)$ from 160 (16×10) cells and two tables of marginal recurrences of $f(\varphi)$, $f(V)$, from 16 and 10 cells, respectively. In the graphical form, these characteristics are represented by the field of $f(V,\varphi)$ and the plots of $f(\varphi)$, $f(V)$ – Fig 1B.

Fig 1B demonstrates the principal difference of distributions at points 1 and 3, in comparison with point 2. At point 2, drift absence and slow drift is observed – the calm recurrence is about 50%, and accumulated recurrence of a speed mogul up to 5cm/s and 10cm/s is 75% and 95%, respectively; the distribution by rhumbs does not have pronounced modes. At points 1 and 3, the distribution is characterised by the pronounced modes. The drift with the speed modules of 5–15cm/s has the maximum recurrence – in total about 70%, directed to W-NW-N. The main distribution mode is well identified in the field of two-dimensional recurrence of $P(V,\varphi)$; some difference of points 1 and 3 is manifested in that, at point 3, the distribution is more centred in the vicinity of the main mode and more wide in the area of high speeds. The distribution by rhumbs of the mean and maximum speed module also has clear maximums, which are consistent with the recurrence maximum of $P(\varphi)$, but do not fully coincide with them. The derived ice drift characteristics at each of the points are in good agreement with the existing knowledge of the drift structure and character in these regions.

For mapping it is better to use the recurrence roses, but in this case the number of rhumbs (direction gradations) should be 4 or 8, and the number of module gradations should not be too large. The chart of recurrence roses for points 1–3 is given in Fig 1Ab; a sharp decrease of the rose size corresponds to frequent calm situations (point 2 particularly). The calm recurrence for each point is given by numbers in the figures.

As in mapping the fields, prescribed at a large number of points, a small scale is used, so recurrence roses illustrate best of all the prevailing speed directions (rhumbs), which are characterised by the most rapid drift, but an analysis of the highest speeds is difficult. For a more detailed description of the drift speeds, see the quantile rose in Fig 1Ac. For each rhumb φ, a separate sampling is made. The speed module quantile V_p of order p is the root of equation $F(V) = p$, where $F(V)$ – exceedance probability, ie, V_p – a value inverse to the accumulated probability. For constructing the quantile rose (Fig 1Ac) the quantiles V_{min}, V_{25}, V_{50}, V_{75}, V_{max} are laid on the rays corresponding to rhumbs φ and are combined by the enveloping lines. Two external envelopes, V_{75} and V_{max}, show an area of 25% of the highest speeds. A comparison of the charts of quantile roses with the recurrence roses, in Fig 1Ab and c, shows that at points 1 and 3, the drift directed to the N and NW, respectively, has the highest drift. These directions are consistent with the rhumbs of the maximum recurrence, but do not fully coincide with them. So, the recurrence roses and the quantile roses are not interchanging, but on the contrary present mutually complementary characteristics.

For information compression, a concept of the vectorial-algebraic approach[2] is used. To describe the distributions of probabilities, the distribution moments are usually used, being restricted most often by the first initial moment – mathematical expectation, and the second central moment – dispersion. The mathematical expectation is determined as vector $\vec{m}_{\vec{V}}$, its projections being equal to mathematical expectations (m_{Vx}, m_{Vy}) of projections of vector $\vec{V}$:

$$\vec{m}_{\vec{V}} = M\{V_X \vec{e}_X + V_Y \vec{e}_Y\} \qquad (2)$$

where $\vec{e}_X$, $\vec{e}_Y$ are orts (unit vectors) of axes OX, OY.

This determination coincides with the determination of mean speed in the component-wise method, in which the speed model is the matrix-line (m_{Vx}, m_{Vy}), or a corresponding matrix-column.

The determination of dispersion $D_{\vec{V}}$ is introduced on the basis of a more general determination of the correlation function $K_{\vec{V}}(\tau)$ which, together with distribution (1), yields an exhaustive characteristic of $\vec{V}(t)$ as a vector random process. The correlation function in the stationary approximation $K_{\vec{V}}(\tau)$ is determined as a mathematical expectation M of tensor product $\otimes$ of centred vectors $\vec{V}^0 = \vec{V} - \vec{m}_{\vec{V}}$, displaced over the time interval τ:

$$K_{\vec{V}}(\tau) = M\{\vec{V}^0(t) \otimes \vec{V}^0(t+\tau)\} \qquad (3)$$

Function $K_{\vec{V}}(\tau)$ characterises the interrelation of the directed speed changes at the moments of time shift by

interval, and gives a quantitative measure of their intensity and orientation in the prescribed coordinate system.

For flat vectors $\vec{V}$ at the fixed time shift, $K_{\vec{V}}(\tau)$ is the tensor of the second rank, the elements of which are represented by co-variations of the corresponding projections. This tensor can be expanded to a sum of symmetric and skew-symmetric tensors:

$$K_{\vec{V}}(\tau)=\begin{pmatrix} K_{V_xV_x}(\tau) & K_{V_xV_y}(\tau) \\ K_{V_yV_x}(\tau) & K_{V_yV_y}(\tau) \end{pmatrix}=C(\tau)+G(\tau)$$
$$=\begin{pmatrix} \lambda_1(\tau) & 0 \\ 0 & \lambda_2(\tau) \end{pmatrix}+0.5G(\tau)\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad (4)$$

While in the component-wise method, the matrix with an independent value of each element is a co-variation, the tensor is one common object, which is invariant relative to transformation of the initial coordinate system, and the matrix in (4) is only a form of its writing. The components of tensor $K_{\vec{V}}(\tau)$ depend, generally, on the choice of the coordinate system; however, invariants exist – scalar functionals of components that do not depend on the change of the coordinate system. Invariants $\lambda_{1,2}(\tau)$ of the symmetrical part C are the values by each fixed τ:

$$\lambda_{1,2}=0.5\left\{I_1K_{V_xV_{xy}}+K_{V_yV_y}\pm\sqrt{\left(K_{V_xV_x}K_{V_x}-K_{V_yV_y}K_{V_y}\right)^2+2K_{V_xV_y}^2}\right\} \quad (5)$$

and the tensor C itself is reduced to a symmetrical form by the initial coordinate system turning by an angle:

$$\alpha=0.5arctg\frac{2K_{V_xV_y}}{K_{V_x}-K_{V_y}}\pm180^{\circ} \quad (6)$$

The skew-symmetrical tensor invariant:

$$G=K_{V_xV_y}-K_{V_yV_x} \quad (7)$$

is called the rotation indicator, and the sum of $\lambda_{1,2}$ is called a linear invariant I_1.

$$I_1=\lambda_1+\lambda_2 \quad (8)$$

So, formulas of invariants (equations (4–8)) contain all elements of tensor $K_{\vec{V}}(\tau)$ and each invariant itself characterises different variability properties.

Dispersion is a co-variation function value at zero time shift $D_{\vec{V}} \equiv K_{\vec{V}}(0)$. Since the reciprocal co-variations of projections are in this case equal to $K_{VxVy}(0) = K_{VyVx}(0)$, the dispersion tensor $D_{\vec{V}}$, noted above (equation (6)), itself is reduced to a symmetrical form by the initial coordinate system turning at an angle:

$$D_{\vec{V}}=\begin{pmatrix} D_{V_x} & K_{V_xV_y} \\ K_{V_yV_x} & D_{V_y} \end{pmatrix}=\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad (9)$$

And its linear invariant I_1 determines the general intensity of variability $\vec{V}$ due to changes of both the module and the direction.

The root-mean-square deviation (RMSD) tensor $\sigma_{\vec{V}}$ is determined as:

$$\sigma_{\vec{V}}=\sqrt{D_{\vec{V}}}=\begin{pmatrix} \lambda_1^{(\sigma)} & 0 \\ 0 & \lambda_2^{(\sigma)} \end{pmatrix} \quad (10)$$

Its invariants are the square roots of the corresponding invariants of tensor $D_{\vec{V}}$ – $\lambda_{1,2}^{(\sigma)}=\lambda_{1,2}^{(D)}$. So, for tensor $\sigma_{\vec{V}}$ the designations $\lambda_{1,2}$ are used without the upper index. For characterisation of the total variability (cumulative effect of module and direction variability $\vec{V}$), the $\sqrt{I_1^{(D)}}$ of tensor $D_{\vec{V}}$ is used, designating it for homogeneity as $I_1^{(0.5)}$.

The vector variation coefficient (with mathematical expectation vector module in the denominator) will be:

$$v=\frac{\sqrt{I_1^{(D)}}}{\vec{m}_{\vec{V}}} \quad (11)$$

If $v > 1$ – the motion is unstable and vice-versa. Also used is the number:

$$\gamma=\frac{D_V}{I_1^{(D)}} \quad (12)$$

Since I_1 is the total dispersion and D_V the speed module dispersion, then addition of λ to 1 (number $1 - \lambda$) determines the contribution of variability of direction $\vec{V}$ to the total dispersion.

Geometrically, $\sigma_{\vec{V}}$ is an ellipse with semi-axes $\lambda_{1,2}$ (equation (5)), oriented in the maximum variability direction α (6). This also underlines the advantage of the vectorial-algebraic method as compared with the component-wise method, in which the maximum variability direction can be determined only in the framework of the *a-priori* hypothesis on the form of distribution of projections (for example, a two-dimensional normal distribution). Graphically (both for the individual points, and on the charts) $\sigma_{\vec{V}}$ is represented by an ellipse and/or by ellipse axes, and it is useful to match it with the mean speed vector $m_{\vec{V}}$. The ellipse form is characterised by invariant:

$$\chi=\frac{\lambda_2}{\lambda_1} \quad (13)$$

– indicator of elongation. At $\chi = 1$, the ellipse transforms to a circle, which signifies an equal variability intensity in all

No. of point	Mean speed		Maximum speed		$I_1^{(0.5)}$	Tensor $\sigma_{\vec{V}}$				v	$1-\gamma$
	$m_{\vec{V}}$	φ	V_{max}	φ		λ_1	λ_2	χ	$\alpha \pm 180$		
	cm/s	degree	cm/s		cm/s	cm/s	cm/s	-	degree	-	-
1	4.4	318	33.1	9	9.0	7.3	5.3	0.73	41	2.9	0.68
2	0.5	157	19.2	212	5.1	4.5	2.4	0.53	197	13.3	0.37
3	5.4	289	42.2	310	11.6	9.8	6.3	0.64	342	3.0	0.61

Table 1: Average and maximum mean daily drift speed vectors and estimates of invariants of dispersion and RMSD tensors

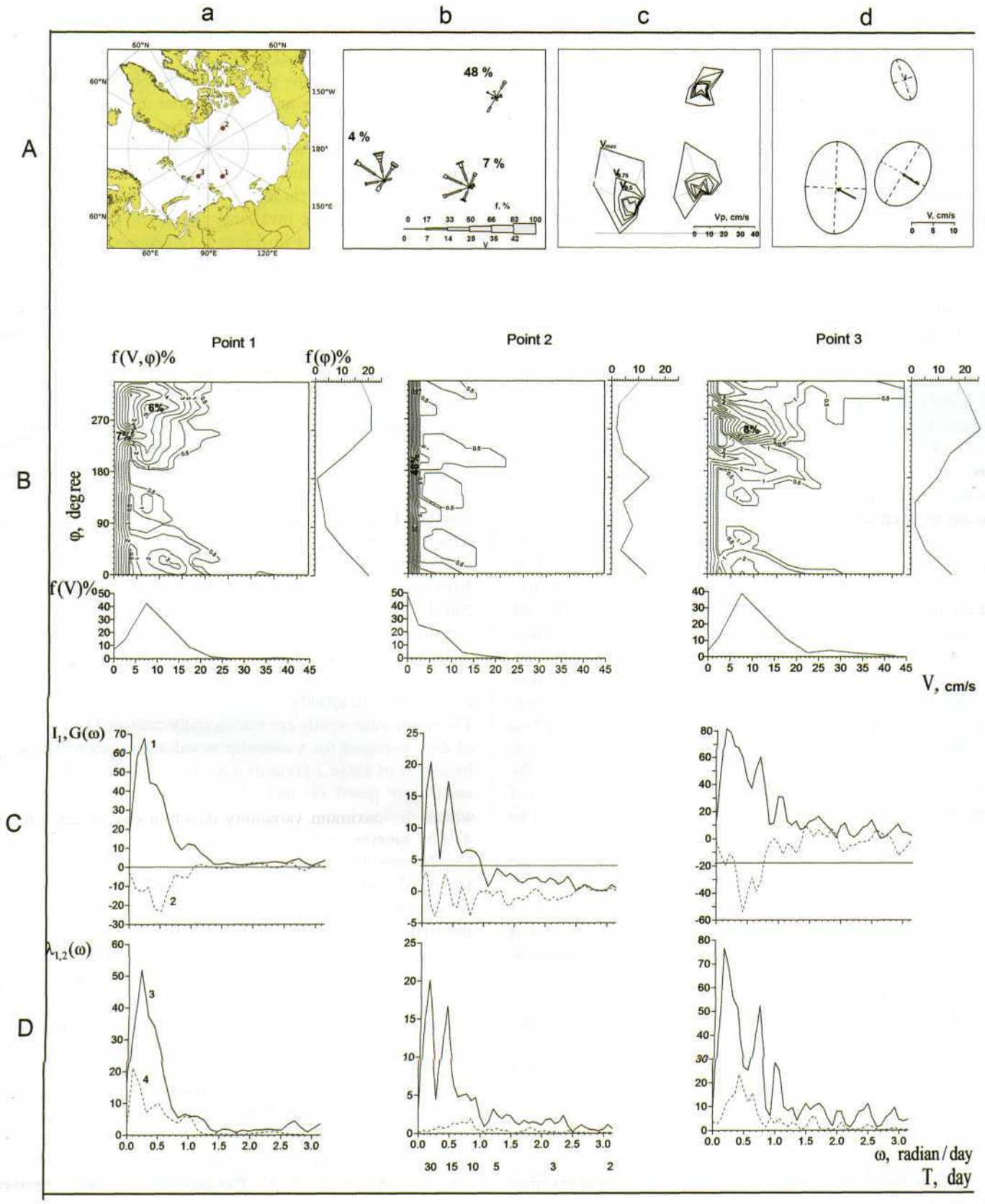


Fig 1: Probabilistic characteristics of variability of mean daily drift at points 1, 2 and 3
A – (l-r): (a) Chart of location of points, (b) recurrence roses – numbers designate the drift absence recurrence, (c) quantile roses, (d) mean speed vectors and RMSD ellipses
B – Plots of two-dimensional and marginal recurrences
C – Spectral tensor invariants I_1 (solid line 1) and G (dashed line 2)
D – Invariants λ_1 (solid line 3) and λ_2 (dashed line 4)
(At the bottom the periods T are given, corresponding to the circular frequency ω)

directions, and at $\chi = 0$, the ellipse degenerates to a straight line segment, meaning reverse changes of $\vec{V}$ or changes of only the module at a constant direction.

Such representation, which is an embodiment of the vectorial-algebraic approach, provides for very significant information compression. Instead of lengthy recurrence tables, the characteristics of variability of the time series of the vector process can be represented by only 11 numbers (Table 1) – module and direction of the average and maximum speed, invariant $I_1^{(0.5)}$, tensor $D_{\vec{V}}$, parameters of the RMSD ellipse λ_1, λ_2, χ, α, variability coefficients v and relative dispersion $1 - \gamma$, determined by the drift direction changes. Since for RMSD ellipses, the opposite directions $\alpha \pm 180$ are equivalent, a value of α, most close to the mean transfer direction, was chosen for ellipse $\sigma_{\vec{V}}$. For the maximum compression of information and mapping vector parameters can be restricted top five: $\vec{m}_{\vec{V}}$ (2 numbers) and RMSD ellipse (3 numbers). Vectors $\vec{m}_{\vec{V}}$ and ellipses $\sigma_{\vec{V}}$ are graphically matched on the map in Fig 1Ad.

Table 1 also demonstrates significant differences in the drift variability at points 1, 3, and at point 2, which was noted in the analysis of probability distributions (Fig 1Ab and Ac). At points 1 and 3, the average speed vector is directed to the NW and is about 4.5–5.0cm/s, and the maximum speeds of 33 and 42cm/s are consistent by direction with $\vec{m}_{\vec{V}}$. Orientation α of ellipse $\sigma_{\vec{V}}$ is predominantly meridional. Large values of elongation parameter χ 0.65–0.75 and non-coincidence of α and φ indicate an important role for direction variability. In fact, parameter $1 - \gamma$ comprises 0.6–0.7, ie, the direction changes explain more than 50% of dispersion. The motion is unstable, since $v = 3$. At point 2, the average speed is negligibly small, the maximum speed (about 20cm/s) is 1.5–2 times smaller than at points 1, 3, the dispersion as compared with points 1, 3 is smaller by approximately four times, the ellipse is more elongated ($\chi = 0.5$), the relative dispersion of the direction variability decreases to 0.35, and instability is an order larger than at points 1, 3.

Vector random stationary process

One more main characteristic (along with $\vec{m}_{\vec{V}}$ and $K_{\vec{V}}(\tau)$) of $\vec{V}(t)$ as a stationary (in a wide meaning) random process is its spectral density, which is determined by the correlation method as:[2]

$$S_{\vec{V}}(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K_{\vec{V}}(\tau) \exp(-i\omega\tau) d\tau \qquad (14)$$

which is represented at fixed frequency ω by the second rank tensor:

$$S_{\vec{V}}(\omega) = \begin{pmatrix} \lambda_1(\omega) & 0 \\ 0 & \lambda_2(\omega) \end{pmatrix} + 0.5G(\omega) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad (15)$$

At the fixed frequency ω the non-negative linear invariant $I_1(\omega)$ of tensor $S_{\vec{V}}(\omega)$ characterises a total variability (due to the module and direction), and the alternating-sign rotation indicator $|G(\omega)| < I_1(\omega)$ shows the prevailing rotation direction – clockwise $G(\omega) > 0$ or counter-clockwise $G(\omega) < 0$. The situation $G(\omega) = 0$ is possible both at the absence of the direction change and at equal rotation intensity in the opposite directions. A necessary and sufficient condition of rotation absence is $\lambda_2(\omega) = 0$ at $\lambda_1(\omega) = I_1(\omega)$. The spectra of invariants $S_{\vec{V}}(\omega)$ for three points are more convenient to present graphically on two plots: on the first the lines $I_1(\omega)$ and $G(\omega)$ are matched (Fig 1C), and on the second – lines $\lambda_{1,2}(\omega)$ (Fig 1D). At the prescribed frequency, tensor $S_{\vec{V}}(\omega)$ can be represented by an ellipse similar to $\sigma_{\vec{V}}$. When determining the ellipse orientation (equation (6)) and elongation (13) at small values of $\lambda_{1,2}(\omega)$, calculation uncertainty can occur, so it is recommended to take into account invariants $\chi(\omega)$ and $\alpha(\omega)$ only for the energy carrying zones $S_{\vec{V}}(\omega)$.

Spectra of tensor invariants (equation (15)) are given in Fig 1C. The plot of invariant $I_1(\omega)$ shows first of all the spectral structure commonality at all three points – almost all energy is concentrated in the low-frequency spectrum region and the intensity of fluctuations with periods $T < 5$ days is small. The main energy active zones are confined to the frequency with periods of about 30–40 days, about 15 days (one of the possible causes is a monthly and half-monthly lunar tide) and 5–10 days (synoptic variability). The spatial differences are manifested in $I_1(\omega)$ values, decreased at point 2 as compared to points 1 and 3. These differences are even more evident by invariants $G(\omega)$, $\lambda_{1,2}(\omega)$. On the plot $G(\omega)$ for points 1, 3, dominance of the counter-clockwise rotation is noted in the low-frequency region, and a corresponding increase of $\lambda_2(\omega)$ indicates additionally an increase of the relative direction variability dispersion in this frequency region. Calculations showed that the orientation of ellipses $\alpha(\omega)$ in the indicated energy carrying zones is approximately the same as for ellipse $\sigma_{\vec{V}}$ in Table 1 and in Fig 1Ad.

Multi-year variability

The multi-year trends are traditionally analysed on the basis of data averaged for a calendar month or a season. The left-hand part of Table 2 presents a module and direction of mean multi-year speed $\vec{m}_{\vec{V}}$ and ellipse parameters of root-mean square deviation $\sigma_{\vec{V}}$, of mean monthly drift speed for points 4–9 by months, with the same denominations as in Table 1. Fig 2a presents matched vectors $\vec{m}_{\vec{V}}$ and ellipses $\sigma_{\vec{V}}$. Data for April and September, the months with the maximum development of ice cover in the Arctic Basin, were used for the analysis.

Fig 2b and c, and Table 2, indicate significant seasonal differences and spatial non-uniformity of the drift speeds. The annual variations for all points characterise an increase in the cold season of both the average transfer rate and the variability intensity. Besides, the RMSD ellipses in the area of Fram Strait are strongly elongated in April, compared to September. Here (points 8, 9), the average speed module values and the size of RMSD ellipses are the largest and the direction φ of vector $\vec{m}_{\vec{V}}$ and orientation α of ellipse $\sigma_{\vec{V}}$ approximately coincide, indicating a predominantly reverse character of variability $\vec{V}$. The latter can be well observed from the value of relative drift direction dispersion $1 - \gamma$ in Table 2, which is about 0.05–0.15 at stations 8–9, and is much higher at the other stations. The seasonal and spatial contrasts are well revealed by the variability coefficient. At stations 8, 9, the motion is stable in September or almost stable in April ($v \leq 1$), while at the other stations $v >> 1$ the instability is increased in September as compared to April; the maximum instability is noted in some months at station 5 (see Table 2).

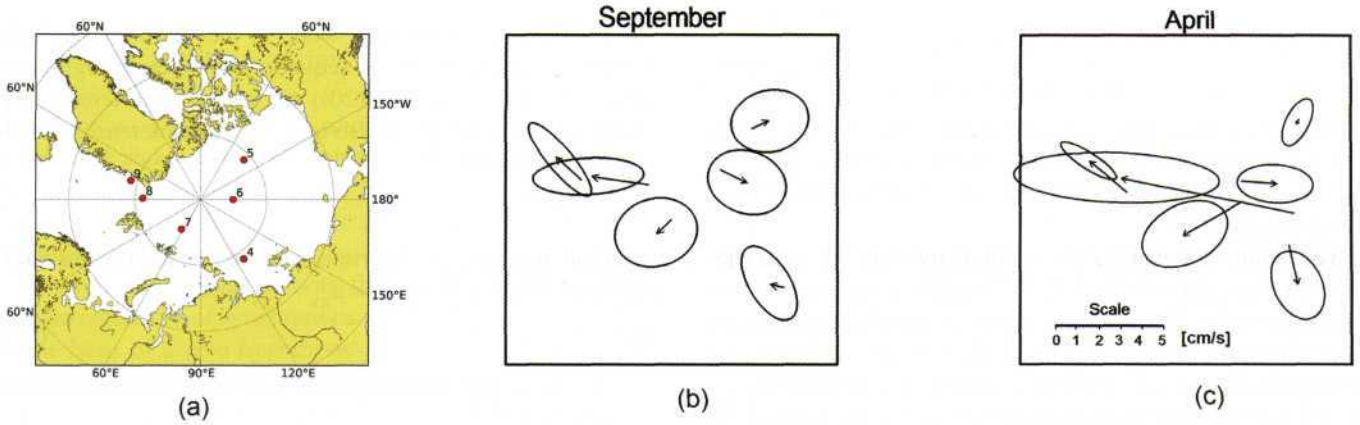


Fig 2: (a) Probabilistic characteristics of variability of mean monthly drift speed at points 4–9, (b) in September, (c) in April, for the period from 1979 to 2006

No. of point	Month	Mean speed		$I_1^{(0.5)}$	Tensor $\sigma_{\vec{V}}$				$1-\gamma$	v
		$m_{\vec{V}}$	φ		λ_1	λ_2	χ	$\alpha \pm 180$		
-	-	cm/s	degree	cm/s	cm/s	cm/s	-	degree	-	-
4	4	1.7	208	1.9	1.6	1.1	0.67	14	0.49	1.62
	9	0.6	318	2.1	1.9	0.9	0.49	9	0.43	4.48
5	4	0.2	332	1.3	1.1	0.5	0.46	-23	0.72	9.96
	9	0.8	197	2.2	1.7	1.4	0.80	26	0.77	4.03
6	4	1.5	186	1.9	1.7	0.8	0.50	3	0.46	1.67
	9	1.4	205	2.2	1.8	1.4	0.79	21	0.76	2.34
7	4	2.9	201	2.4	2.0	1.3	0.65	23	0.42	1.15
	9	0.8	177	2.4	1.8	1.5	0.82	34	0.68	4.06
8	4	7.9	191	4.7	4.5	1.1	0.25	2	0.13	0.71
	9	2.5	186	2.6	2.5	0.8	0.34	-3	0.09	1.33
9	4	2.0	210	1.5	1.4	0.4	0.25	23	0.11	0.90
	9	1.5	211	2.1	2.1	0.5	0.26	40	0.06	1.74

Table 2: Mean multi-year drift speed vectors and estimates of invariants of dispersion and RMSD tensors based on mean monthly data in April and September from 1979 to 2006

Further analysis could be conducted with extraction of the linear trend of the random vector process $\vec{V}(t)$, determined as:[6]

$$\vec{V}(t) = \vec{m}_{\vec{V}} + \vec{a}t + \vec{\varepsilon} \qquad (16)$$

where $\vec{a}(a_x, a_y)$ is the vector with Cartesian components α_x, α_y, which are inclinations of the trends of projections calculated by means of linear regression model[7] $V_{x,y}(t) = m_{Vx,y} + \alpha_{x,y}t + \varepsilon_{x,y}(t)$. The trend parameters are the module and direction of the vector analogue of angular coefficient $\vec{a}$ and invariants of tensor $\sigma_{\vec{\varepsilon}}$ of anomalies $\vec{\varepsilon}$ relative to the trend.

The trend contribution to the total dispersion is determined by addition to 1 (or to 100%) of the relation $\delta = I_1^{(\delta)}/I_1^{(v)}$ of the linear invariant of tensor $D_{(\vec{\varepsilon})}$ of anomalies $\vec{\varepsilon}$ to the corresponding invariant of initial dispersion $D_{(\vec{V})}$.

System of connected vector random values, combined and conventional distribution of probabilities

When dealing with vector systems $\{\vec{V}, \vec{U}\}$, their properties are not completed just with the properties of individual components, but also include interdependencies, such as speeds of wind and ice drift, drift speeds from full-scale and model data (a validation case) or the drift speeds at several points. Consider $\{\vec{V}, \vec{U}\}$ as a model of connecting vector random values, the explicit characterisation of which is prescribed by a combined distribution of probabilities, distribution of probabilities of the system components and a family of conventional distributions.[5] If the distributions $f_{\vec{V}}(V,\varphi)$ $f_{\vec{U}}$, (U, Ψ) (equation (1)) of system components $\{\vec{V}, \vec{U}\}$ are two-dimensional, the combined distribution is four-dimensional:

$$f_{\vec{V},\vec{U}}(V,\varphi,U,\psi) = P\{V_1 \le V < V_2, \quad \varphi_1 \le \varphi < \varphi_2, \quad U_1 \le U < U_2, \quad \psi_1 \le \psi < \psi_2\}, \qquad (17)$$

Although in a particular case, where $\vec{V}$ and $\vec{U}$ are values of the same name (eg, ice drift from full-scale and model data) and their distributions are determined by equal gradations, the results of equations (1) and (17) formally coincide.

The conventional distribution of probabilities is determined as a recurrence of $\vec{V}$, provided that the module and direction of $\vec{U}$ are within the prescribed ranges:

$$f_{\vec{V}\vec{U}}(V,\varphi) = P\{V_1 \le V < V_2, \quad \varphi_1 \le \varphi < \varphi_2, \quad U_1 \le U < U_2, \quad \psi_1 \le \psi < \psi_2\} \tag{18}$$

So, even with a rather coarse data grouping (for example, 8 direction rhumbs and 5 module gradations), the family of conventional distributions consists of 40 (5*8) two-dimensional distributions.

The multi-dimensionality of characteristics (equations (1), (17), (18)) of system $\{\vec{V}, \vec{U}\}$ makes it difficult to use them practically. It is recommended[5] to use the moments of conventional distribution. The conventional mathematical expectation $\vec{m}_{\vec{V}|\vec{U}}$ (U, Ψ) and conventional RMSD $\sigma_{\vec{V}|\vec{U}}$ (U, Ψ) are the values of mean speed vector (equation (2)) and the RMSD ellipse (equation (10)) with invariants (5–6), (8), (11–13) of vector $\vec{V}$ at different values of $\vec{U}$. An indication of dependence is their non-constancy $\vec{m}_{\vec{V}|\vec{U}}(\bullet) \neq const$, $\sigma_{\vec{V}|\vec{U}}(\bullet) \neq const$ and non-coincidence with the common mean value and RMSD over the entire sampling $\vec{m}_{\vec{V}|\vec{U}}(\bullet) \neq \vec{m}_{\vec{V}}$, $\sigma_{\vec{V}|\vec{U}}(\bullet) \neq \sigma_{\vec{V}}$. By the form of function $\sigma_{\vec{V}|\vec{U}}(\bullet)$, one can judge the form of dependence between $\vec{V}$ and $\vec{U}$. Since in this case it is necessary to analyse multi-dimensional objects, the question whether a linear approximation is permissible is important. However, in a number of cases, the hypothesis of linearity can be assumed *a-priori*. In the case for points 1, 3 in Fig 1 and points 7, 8, 9 in Fig 2, this assumption is justified by the small distances between them and by similarity of the variability characteristics.

Linear regression and correlation

The covariation of the system of connecting vector random values is represented by a tensor of cross-variance:

$$D_{\vec{V}\vec{U}} = \begin{pmatrix} K_{V_xU_x} & K_{V_xU_y} \\ K_{V_yU_x} & K_{V_yU_y} \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} + 0.5G \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{19}$$

The kinematic treatment of its invariants is similar to that of tensor $K_{\vec{V}}(\tau = 0)$.

Others[8] have determined the linear regression of abstract algebraic vectors by equation:

$$\vec{V}_i = A_{\vec{V}\vec{U}}\vec{U}_i + \vec{B}_{\vec{V}\vec{U}} + \vec{\varepsilon}_i \equiv \hat{\vec{V}}_i + \vec{\varepsilon}_i \tag{20}$$

where $A_{\vec{V}\vec{U}} = D_{\vec{V}\vec{U}}D_{\vec{V}}^{-1}$ is the regression tensor, and $\vec{B}_{\vec{V},\vec{U}} = \vec{m}_{\vec{V}} - A_{\vec{V}\vec{U}}\vec{m}_{\vec{U}}$ the vector (analogues of coefficient a and a free term b of regression of scalar values $Y = aX = b$). The vector of unexplainable residual predicted by regression (equation (20)) is $\vec{\varepsilon}_i = \vec{V}_i - \hat{\vec{V}}_i$. It has been shown[5] that equation (20) is also applicable to the case of Euclidian vectors.

In the component-wise model of vectors the regression has been determined[9] by a system of equations:

$$\begin{aligned} V_X &= a_XU_X + b_XU_Y + C_X + \varepsilon_X \\ V_Y &= a_YU_Y + b_YU_Y + C_Y + \varepsilon_Y \end{aligned} \tag{21}$$

the numerical solution of which is consistent with equation (20).

The linear regression closeness (20) is determined by correlation, which should take into account a combined effect of cross-changes of modules and directions of vectors $\vec{V}$ and $\vec{U}$. The component-wise regression model (equation (21)), in which correlation is represented by a matrix-line of coefficients of multiple linear regression $\{R_{Vx;UxUy}\ R_{Vy;UxUy}\}$, complicates the interpretation, as a combined interpretation of the elements of this matrix is impossible without the additional assumptions. Since in the vectorial-algebraic method, covariation $D_{\vec{V}\vec{U}}$ (equation (19)) and regression coefficient $A_{\vec{V}\vec{U}}$ (equation (20)) are second rank tensors, the vector correlation should also be a second rank tensor, which has been determined:[8]

$$R^2_{\vec{V}\vec{U}} = A_{\vec{V}\vec{U}}A_{\vec{U}\vec{V}} \tag{22}$$

– vector analogue of determination of the correlation coefficient of scalar values $r_{XY} = a_{XY} \cdot a_{XY}$, and in[5]:

$$\tilde{R}_{\vec{V}\vec{U}} = D_{\vec{V}\vec{U}}D_{\vec{V}}^{-0.5}D_{\vec{U}}^{-0.5} \tag{23}$$

– vector analogue of determination of correlation coefficient $r_{xy} = D_{xy}D_x^{-0.5}D_y^{-0.5}$.

For linear independent vectors the tensors $R_{\vec{V}\vec{U}}$, $\tilde{R}_{\vec{V}\vec{U}}$ are represented by zero matrixes and, in the case of deterministic linear dependence, by single matrixes. However, in the general case of stochastic dependence, their kinematic interpretation is difficult due to multi-dimensionality (equations (22–23)). That is why a system of simplified correlation indicators has been proposed[5] and described,[10] which are determined through invariants of the covariance tensor $D_{\vec{V}|\vec{U}}$ and dispersion tensors $D_{\vec{V}}$, $D_{\vec{U}}$:

$$r_{\uparrow\downarrow} = \frac{I_1^{(\vec{V}\vec{U})}}{\sqrt{I_1^{(\vec{V})}I_1^{(\vec{U})}}}, \quad r_{\perp} = \frac{G_1^{(\vec{V}\vec{U})}}{\sqrt{I_1^{(\vec{V})}I_1^{(\vec{U})}}}, \quad \mu = \sqrt{r^2_{\uparrow\downarrow} + r^2_{\perp}} \tag{24}$$

represented by indicators of the collinear $r_{\downarrow\uparrow}$, orthogonal $r_{\perp}$ correlation and the indicator of vector correlation μ. For linearly independent vectors $\mu = 0$ at deterministic linear dependence, the indicators $r_{\downarrow\uparrow}$, $r_{\perp}$ attain values from the interval of ± 1 which provide fulfilment of equality $\mu = 1$, and at the stochastic dependence, fulfilment of inequality $\mu < 1$. If the collinear vector components $\vec{V}$ and $\vec{U}$ are mainly uniaxial, then $r_{\downarrow\uparrow} > 0$ and vice-versa. If the second vector $\vec{U}$ is turned predominantly to the right relative to the first vector $\vec{V}$, then $r_{\perp} > 0$ and vice-versa. If the angle between vectors $\vec{V}$ and $\vec{U}$ is equal to 45°, then $r_{\downarrow\uparrow} = r_{\perp}$ and $r_{\perp}$ by the module. If the angle is smaller than 45°, then the module $r_{\downarrow\uparrow}$ is larger than the module $r_{\perp}$ and vice-versa. In the case of coincidence of $\vec{V}$ and $\vec{U}$, $r_{\downarrow\uparrow} = \mu = 1$, $r_{\perp} = 0$.

The correlation indicators of mean monthly speeds at points 8–7 and 8–9 comprise:

Points	$r_{\downarrow\uparrow}$	$r_{\perp}$	μ
8–7	0.50	–0.15	0.52
8–9	0.59	–0.03	0.59

The dependence exists as μ is meaningfully larger than 0 and is stochastic, as μ is meaningfully smaller than 1. It is almost entirely determined by connectivity of uniaxial collinear components. For points 8–9, it is slightly closer than for points 8–7, as $r_{\downarrow\uparrow}$ comprises 0.52 and 0.59. The negative correlation of orthogonal components for points 8–7 is at the boundary of the statistical significance level and there is no orthogonal correlation for points 8–9.

At the availability of initial data at grid points or at the points scattered over the space it is very convenient to map the invariant characteristics and perform an analysis of the fields and regionalisation, depending on the character of variability of the vector value.

Comparison of model calculations with full-scale data in the terms of estimates of correlation and regression

The correlation indicators $r_{\downarrow\uparrow}$ and $r_{\perp}$ can be used for a full quantitative and qualitative comparison of model calculations $\vec{U}(\vec{r}, t)$ with measurement data $\vec{V}(\vec{r}, t)$. To find out how adequately the model simulates the character of temporal variability by argument t and spatial variability by argument $\vec{r}$, it is reasonable to begin with the express-analysis of correlation indicators of synchronous values at the chosen points on the basis of a zero hypothesis H_0: 'the data of full-scale measurements and model calculations coincide $\vec{U}(\vec{r}, t) = \vec{V}(\vec{r}, t)$'. According to this hypothesis, the system elements of correlation indicators (equation (24)) should attain the following values:

$$\mu = 1, r_{\uparrow\downarrow} = +1, r_{\perp} = 0 \qquad (25)$$

The condition (equation (25)) is necessary but it is not sufficient. One of the causes of this effect can be connected with the fact that covariation is a central moment of the distribution of probabilities, ie, it determines closeness of dependence between the values centred to the mathematical expectation. For example, if the model simulates ideally the vector direction $\varphi_{\vec{U}} = \varphi_{\vec{V}}$ and overestimates (or underestimates) the module by a constant value $U = V \pm const$, then the condition (25) will be fulfilled all the same. Therefore at validity of H_0 in addition to (25) the following condition should be fulfilled: in the regression (equation (20)) the regression coefficient $A_{\vec{V}|\vec{U}}$ is represented by a single matrix

$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and the free term $B_{\vec{V}|\vec{U}}$ and residual $\vec{\varepsilon}_i$ are zero vectors.

As an example illustrating the possibilities of using the vectorial-algebraic approach for validation (comparison of vector fields), consider the spatial distributions of the correlation indexes $r_{\downarrow\uparrow}$, $r_{\perp}$ and μ, giving information on consistency of the drift fields. Compared here are the drift fields based on IFREMER data for two different months of one season – December 1982 and January 1983 –by spatial structure rather than by the model data with observation data. This is done in order to demonstrate only the principal technical possibilities of the vectorial-algebraic approach.

As shown in Fig 3, presenting information on the initial data in the form of fields of mean monthly values (Fig 3a,b), the fields are quite similar. For the initial time series of the drift at some grid points with an interval of 125km the indexes $r_{\downarrow\uparrow}$, and $r_{\perp}$ were calculated and charts were constructed (Figs 3c,d,e). In this case, the average initial fields depicted in Fig 3a and b, and the degree of consistency of temporal and spatial variability of the drift fields in December 1982 and January 1983 are compared, while the index values at each field point are calculated from the time series at each specific point.

As seen from the charts, the relation between the fields exists in five regions (Fig 3a), where the value of index μ exceeds 0.5. The largest correlation of temporal variability is observed in the north of the East-Siberian Sea, near the shores of the Canadian Arctic archipelago, in the area of the Amundsen Basin, in Fram Strait and in Baffin Bay, and it is mainly determined by consistent unidirectional collinear fluctuations of the drift vector (Figs 3a and b). The contribution due to the orthogonal changes (coordinated direction changes of the drift vectors) is of little significance in most regions except for the Fram Strait region, where the orthogonal index is quite large and has positive values, indicating that the drift direction from December to January was deviating clockwise, ie, being more pressed against the shores of Greenland. A comparison of the drift fields differing by time means that the drift fields in the zones identified had a more stable character in December 1982 and January 1983.

Summarising this brief analysis it can be assumed that, in the case of comparing the model fields with the observation data, the consistency of the fields can, on average, be higher, but it will be possible to delineate the coastal or open ocean regions where the model simulates the spatial structure of the fields and their temporal variability, better or worse, and to introduce corrections to the model parameters.

CONCLUSION

This article demonstrates the advantages of the vectorial-algebraic approach for the analysis of vector series where a Euclidean vector is used as a model, and the vector temporal processes are described by a small set of scalar invariant characteristics, derived from the dispersion and the spectral tensors. The method allows vector processes to be adequately described, the fields of these values to be constructed (if the corresponding data are available), an inter-comparison to be performed of the uniform series or the fields with a different time interval, or model and observed values, and the relationship of the fields of wind or currents to be described. The method has large potential for describing the temporal variability of the vector series and the fields at different scales, for zonation of regions by means of different criteria of temporal variability or by the characteristics of relationship of the vector time series and the fields.

ACKNOWLEDGEMENTS

The authors are grateful to Professor VA Rozhkov and Dr Yu P Klevantsov for their valuable advice given in the process of manuscript preparation. The work was carried out in the framework of WP5 of the MyOcean FP7 Project (2009–2011).

REFERENCES

(a) http://iabp.apl.washington.edu

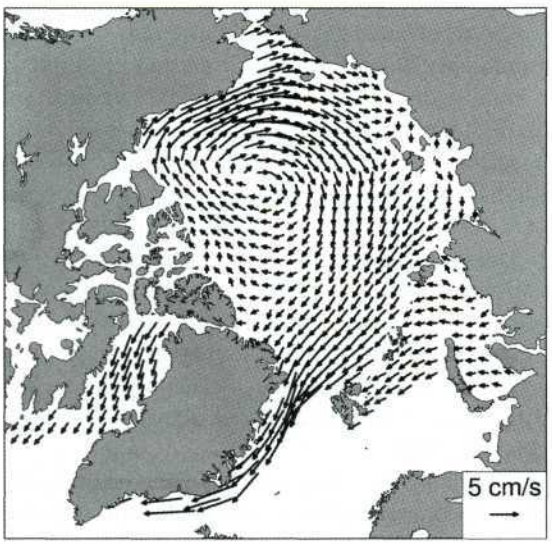
(b) http://cersat.ifremer.fr/data/discovery/by_product_type/gridded_products/psi_amsr_drift

(c) http://cersat.ifremer.fr/science/sea_ice/validation_of_arctic_sea_ice_drift_with_iabp_buoys

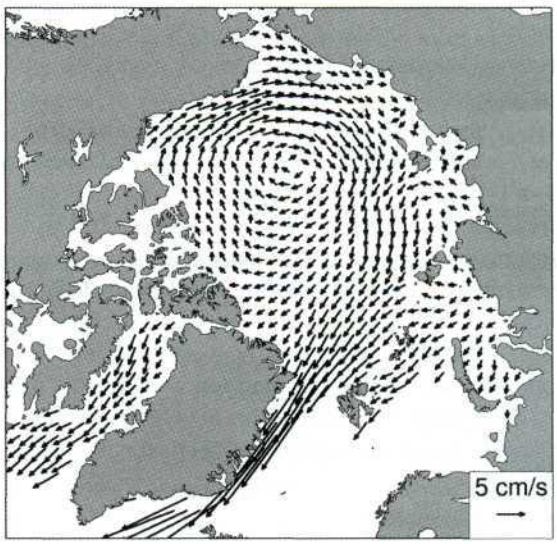
(d) .Fowler C. 2003, updated 2007. *Polar pathfinder daily 25km EASE-Grid sea ice motion vectors*. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media

(e) http://topaz.nersc.no

(f) www.esimo.ru

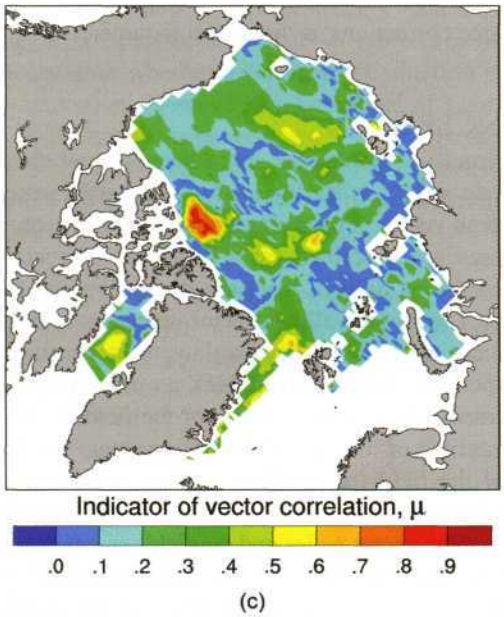


(a)



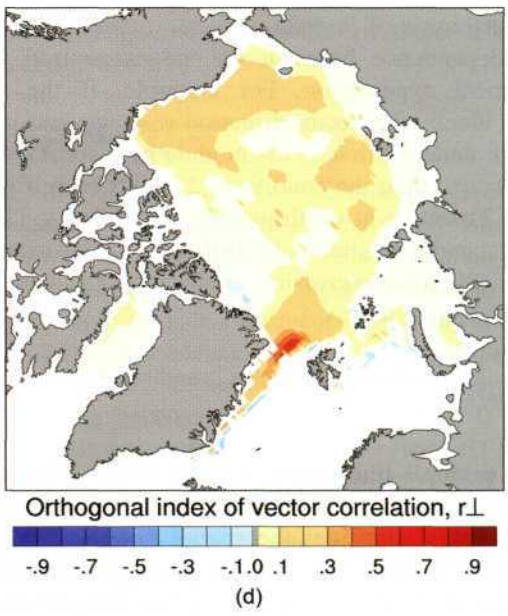
(b)

1982-1983 (Dec-Jan), μ



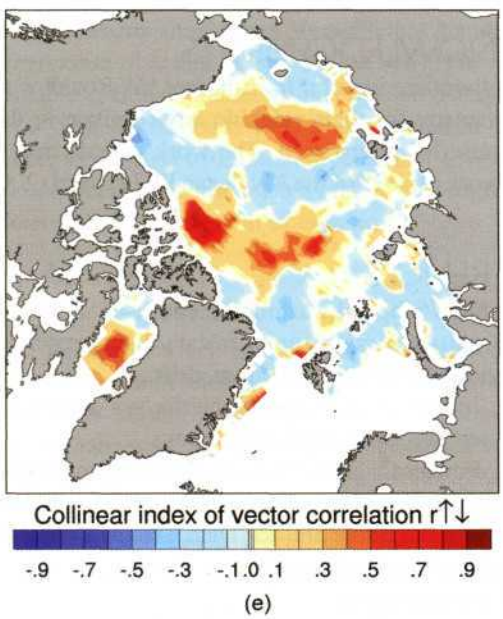
(c)

1982-1983 (Dec-Jan), r⊥



(d)

1982-1983 (Dec-Jan), r↑↓



(e)

Fig 3: Mean monthly ice drift fields for (a) December 1982 and (b) January 1983; and the fields of correlation indexes: (c) total correlation μ, (d) collinear $r_{\uparrow\downarrow}$ and (e) orthogonal $r_{\perp}$

1. Schwegmann S, Haas Ch, Fowler Ch and Gerdes R. 2011. *A comparison of satellite-derived sea-ice motion with drifting-buoy data in the Weddell Sea, Antarctica.* Annals of Glaciology **52(57)**:103–110.

2. Belyshev AP, Klevantsov Yu P and Rozhkov VA. 1983. Probabilistic analysis of sea currents. L., Gidrometeoizdat. 264pp.

3. Belyshev AP, Klevantsov Yu P and Rozhkov VA. 1981. *On illusions and reality in the methods of analysis of sea currents.* Proc. Of the State Oceanographic Institute **157**: 3–19.

4. Kochin NE. 1961. Vector analysis and beginning of tensor calculation. M., Publishing House of the USSR Academy of Science. 425 pp.

5. Klevantsov Yu P, Mikulinskaya SM and Rozhkov VA. 1996. *On the analysis of sea current speed vectors.* Meteorology and Hydrology **9**: 96–105.

6. Bokov VN, Bukhanovsky AV, Ivanov NE, Klevantsov Yu P and Rozhkov VA. 2001. *Spatial-temporal variability of the wind field at temperate latitudes of the Northern Hemisphere.* Izvestoya of RAN. Series Physics of the Atmosphere and the Ocean **37, No. 2**: 170–181.

7. Dreiper N and Smith G. 1986. Applied regression analysis. M., Finances and statistics, v.1. 366 pp.

8. Obukhov AM. 1945. *Theory of vector correlation.* Scientific Notes of Moscow State University **45**: 56–71.

9. Gudkovich ZM. 1965. *Correlation method of processing of observations of currents*, 1965. Problems of the Arctic and the Antarctic **21**: 56–60.

10. Ivanov NE. 2004. On *the characteristics of correlation of wind speeds, sea currents and ice drift.* Meteorology and Hydrology. **8**: 61–72.