Ocean Dynamics (2010) 60:155–170
DOI 10.1007/s10236-009-0255-9

On non-Gaussian SST variability in the Gulf Stream and other strong currents

Philip Sura

Received: 8 May 2009 / Accepted: 26 November 2009 / Published online: 22 December 2009

Abstract This paper examines the physics of observed non-Gaussian sea surface temperature (SST) anomaly variability in the Gulf Stream system in a recently developed stochastic framework. It is first shown from a new high-resolution observational data set that the Gulf Stream system is very clearly visible as a band of negative skewness all the way from Florida, over Cape Hatteras, to the central North Atlantic. To get an idea about the detailed non-Gaussian variability along the Gulf Stream, probability density functions are calculated at several locations. One important observational result of this study is that the non-Gaussian tails of SST variability in the Gulf Stream system follow a power-law distribution. The study then shows that the observed non-Gaussianity is consistent with stochastic advection of SST anomalies in an idealized zonal current. In addition, stochastic advection is compatible with the observed northward eddy heat flux in the Gulf Stream, providing a new dynamical view at the heat balance in strong currents.

Keywords Non-Gaussian sea surface temperature variability · Gulf Stream · Power-law · Stochastic advection

Responsible Editor: Pierre Lermusiaux

P. Sura (✉)
Department of Meteorology, Florida State University,
1017 Academic Way, Tallahassee, FL 32306-4520, USA
e-mail: psura@fsu.edu

1 Introduction

Since the very early days of physical oceanography, the Gulf Stream system plays a central role in the dynamical description of the general circulation of the ocean. One reason for that is, of course, that the North Atlantic has been the major shipping route connecting the Old with the New World for about 500 years now. The Gulf Stream is a warm western boundary current that transports large amounts of heat northward, and consequently is a major part of the global climate system. Therefore, it is important to study and understand the physics behind its temperature fluctuations. Here, we will study the physics of non-Gaussian sea surface temperature (SST) variability in the Gulf Stream and other strong currents in a recently developed stochastic framework.

The study of non-Gaussian variability and dynamics in the ocean and atmosphere became more popular only recently. That is because we now have the required high-resolution data and models available to meaningfully explore non-Gaussian statistics from observations and models. For a review of non-Gaussian statistics in the atmosphere, see, for example, Stephenson et al. (2004) or Sura et al. (2005) and references therein. Here, we focus on oceanic variability, keeping in mind that non-Gaussianity in the atmospheric forcing fields (e.g., Monahan 2006a, b; Gille 2005) may account for some fraction of the oceanic non-Gaussianity.

The research of non-Gaussian oceanic variability includes many relevant subtopics. It is obvious that extreme value (non-Gaussian) statistics of extreme wave heights and coastal sea levels are of utmost importance (e.g., Jha and Winterstein 2000; Coles 2001). More relevant to the present study are recent efforts to quantify

forecast errors in coastal and shelf sea prediction models (e.g., Lermusiaux 2006; Lermusiaux et al. 2006). By performing ensemble simulations of the models, the non-Gaussian distribution of forecast errors can be determined. Assuming ergodicity, the non-Gaussian distribution of the ensemble members equals the local temporal non-Gaussianity. As it turns out, the spatial patterns of skewness and kurtosis of SST and sea surface height forecast errors resemble the spatial fields of non-Gaussian SST variability (e.g. Lermusiaux et al. 2002; Auclair et al. 2003). In particular, Lermusiaux et al. (2002) showed that the skewness of sound speed (which is proportional to SST) variability changes sign at a temperature front for a shelfbreak application off Cape Cod, and Auclair et al. (2003) showed that the skewness of sea surface height variability in the Gulf of Lions is negative (positive) at the northern (southern) edge of a strong alongshore current. These regional applications are very useful to elucidate the effects of very complex and specific local conditions on non-Gaussian ocean variability. Here, we are taking a different approach, neglecting local conditions as much as possible to provide a more general view of non-Gaussian SST variability.

The earliest work on observed non-Gaussian SST variability using Ocean Weather Station data dates back two to three decades (Blaauboer et al. 1982; Müller 1987). More recently, Burgers and Stephenson (1999) analyzed the non-Gaussianity of observed El Niño SSTs. While these studies are important, they did not provide a global view of non-Gaussian SST variability. In a recent paper, Sura and Sardeshmukh (2008) closed this gap by providing a global stochastic–dynamical view of non-Gaussian SST variability. There, the starting point is the observation that daily SST anomalies are highly non-Gaussian all over the globe for observed skewness and kurtosis, respectively. Most importantly, the skewness and kurtosis of SST anomalies are found to be strongly linked at most locations around the globe. The link, or constraint, is that the kurtosis (kurt) is everywhere equal to or larger than one-and-a-half times the squared skewness (skew): $\text{kurt} \geq (3/2)\text{skew}^2$. This constraint is then analytically explained with a simple multiplicative noise model, which can be directly derived from basic mixed-layer dynamics. That means that a simple linear stochastic differential equation (SDE) with multiplicative noise, derived from first principles, captures the overall dynamics of global SST variability remarkably well. In particular, this shows that the observed non-Gaussianity of SST anomalies is due to multiplicative noise rather than to nonlinearities in the deterministic part of the SST equation.

The theory presented in Sura and Sardeshmukh (2008) does not predict the parameters of the stochastic model. In fact, all specific parameters drop out of the general constraint. That is, Sura and Sardeshmukh (2008) provide observational evidence that the dynamics of non-Gaussian SST variability is governed by a linear SDE with linear multiplicative noise. They do not provide specific local parameters based on specific physical processes. These parameters, however, determine the sign and magnitude of the observed skewness and kurtosis at a given location beyond the general constraint. Therefore, the approach taken by Sura and Sardeshmukh (2008) applies the ideas of statistical mechanics to SST variability, neglecting detailed local dynamics as much as possible.

This study looks into the more detailed SST dynamics within a specific flow regime. It has been noted before that strong currents such as the Gulf Stream show a very pronounced negative SST anomaly skewness. Thus, this study takes a closer look at non-Gaussian SST variability in the Gulf Stream system within the stochastic framework provided by Sura and Sardeshmukh (2008). One new observational result is that the non-Gaussian tails of SST variability in the Gulf Stream system follow a power-law distribution, consistent with multiplicative noise dynamics of, so far, unspecified physical origin. This study then shows that the negative SST skewness in the Gulf Stream may originate from stochastic advection of SST anomalies. Stochastic advection is also shown to be consistent with the observed northward eddy heat flux in the Gulf Stream, providing a novel dynamical look at the heat balance in strong currents. While this approach might look like a purely academic endeavor, there are important applications closely related to this study. One is connected to the recent interest in quantifying uncertainties in ocean predictions (e.g., Lermusiaux 2006; Lermusiaux et al. 2006). To parameterize the effect of sub-grid scale variability in ocean models, stochastic parameterizations are becoming more popular. As we will see, this study gives some guidance on the physics of stochastic model components by showing that observed non-Gaussian SST variability is consistent with stochastic advection of SSTs in an idealized zonal current. In addition, the knowledge of the detailed non-Gaussian statistics of an oceanic variable such as SST allows to better quantify sampling uncertainties and errors in observed or modeled data.

In Section 2, the observational results are presented. In particular, it is shown that the negatively skewed non-Gaussianity in the Gulf Stream system is consistent with a power-law. Then, Section 3 provides a theoretical explanation of the negatively skewed power-law

behavior of SST anomalies in strong currents based on stochastic advection. Finally, Section 4 provides a summary and discussion.

2 Observations

Probability density functions (PDFs) are useful diagnostic measures of the dynamics of stochastic systems. In particular, higher moments beyond the variance (that is, deviations from Gaussianity) can shed light on the underlying dynamics. For example, Sura and Sardeshmukh (2008) provided a global view of non-Gaussian SST variability by presenting global maps of skewness and kurtosis of SST anomalies to characterize the overall shape of the PDF. They then used an observed skewness–kurtosis link to elucidate the underlying dynamics of global non-Gaussian SST variability.

2.1 Data

In this study, we use the same data as in Sura and Sardeshmukh (2008) to study the non-Gaussian SST variability in the Gulf Stream system. To briefly recapitulate, this high-resolution observational data set compiled and provided by Reynolds et al. (2007) consists of a blended analysis of daily SST fields based on infrared satellite data from the Advanced Very High Resolution Radiometer (AVHRR) and in situ data from ships and buoys. The analysis was performed using optimum interpolation with a separate step to correct satellite biases relative to the in situ data. The in situ data were obtained from the International Comprehensive Ocean—Atmosphere Data Set (ICOADS; http://icoads.noaa.gov/). This daily SST analysis is available on a 0.25-degree latitude/longitude grid from January 1985 to the present. A more detailed description of the data set and analysis procedure can be found in Reynolds et al. (2007). SST anomalies were calculated by subtracting the daily climatology and linear trend from the full daily values. Then, the extended summer (May–October) and extended winter (November–April) seasons are analyzed.

It should be noted that Reynolds et al. (2007) also provided a more advanced data set combining AVHRR SSTs with the Advanced Microwave Scanning Radiometer (AMSR) data. The advantage of AMSR SSTs is that they have a near all-weather coverage; in contrast, AVHRR SSTs are only available in cloud-free conditions. That is, the combined AVHRR–AMSR data set is less biased by missing data. So why did we chose the AVHRR-based data set rather than the AVHRR–AMSR product? The answer is that to reliably estimate higher moments such as skewness and kurtosis (discussed below) requires timeseries as long as possible. AMSR data, however, only became available in June 2002. That is, the main rationale for using the AVHRR SSTs is the length of the record: almost 18 years more than the AVHRR–AMSR product. We are, however, aware of the potential biases of the AVHRR SSTs and discuss them accordingly in the next subsection, and also in the summary and discussion.

2.2 Non-Gaussian SST variability in the Gulf Stream system

First skewness (third moment) and kurtosis (fourth moment) are used to characterize the overall shape of a PDF. If the standard deviation of an anomaly x with zero mean ($\overline{x} = 0$) is denoted by σ, the skewness (skew) and kurtosis (kurt) become

$$\text{skew} \equiv \frac{\overline{x^3}}{\sigma^3} \quad , \quad \text{kurt} \equiv \frac{\overline{x^4}}{\sigma^4} - 3 \quad , \tag{1}$$

where the overbar denotes a time or, assuming ergodicity, ensemble average. In the following discussion, we can employ the notion of a time average in most of the cases. However, if we want to highlight the use of an ensemble average, we will use angle brackets $\langle ... \rangle$ (for example, in Appendix A, where we discuss some elements of SDEs). Skewness is a measure of asymmetry of a PDF. For unimodal PDFs, the following interpretations are valid. If the left tail is heavier (more pronounced) than the right tail, the PDF has negative skewness. If the reverse is true, it has positive skewness. If the PDF is symmetric, it has zero skewness. Kurtosis (or more accurately, "excess kurtosis," since we subtract the kurtosis of 3 for a Gaussian distribution) measures the excess probability (fatness) in the tails, where excess is defined in relation to a Gaussian distribution. In general, moments such as skewness and kurtosis are measures to obtain a first-order characterization of the overall shape of a PDF. In principle, the knowledge of all moments is almost equivalent to that of the full PDF (see, e.g., Sornette 2006). That is, skewness and kurtosis are shown first, before moving on to present the detailed shape of the PDFs.

The skewness of SST anomalies in the North Atlantic in the extended summer (upper panel) and winter (lower panel) is shown in Fig. 1. The related kurtosis is shown in Fig. 2. One of the most striking features in the skewness fields is that the Gulf Stream (and its extension, the North Atlantic Current) is very clearly visible as a band of negative skewness all the way

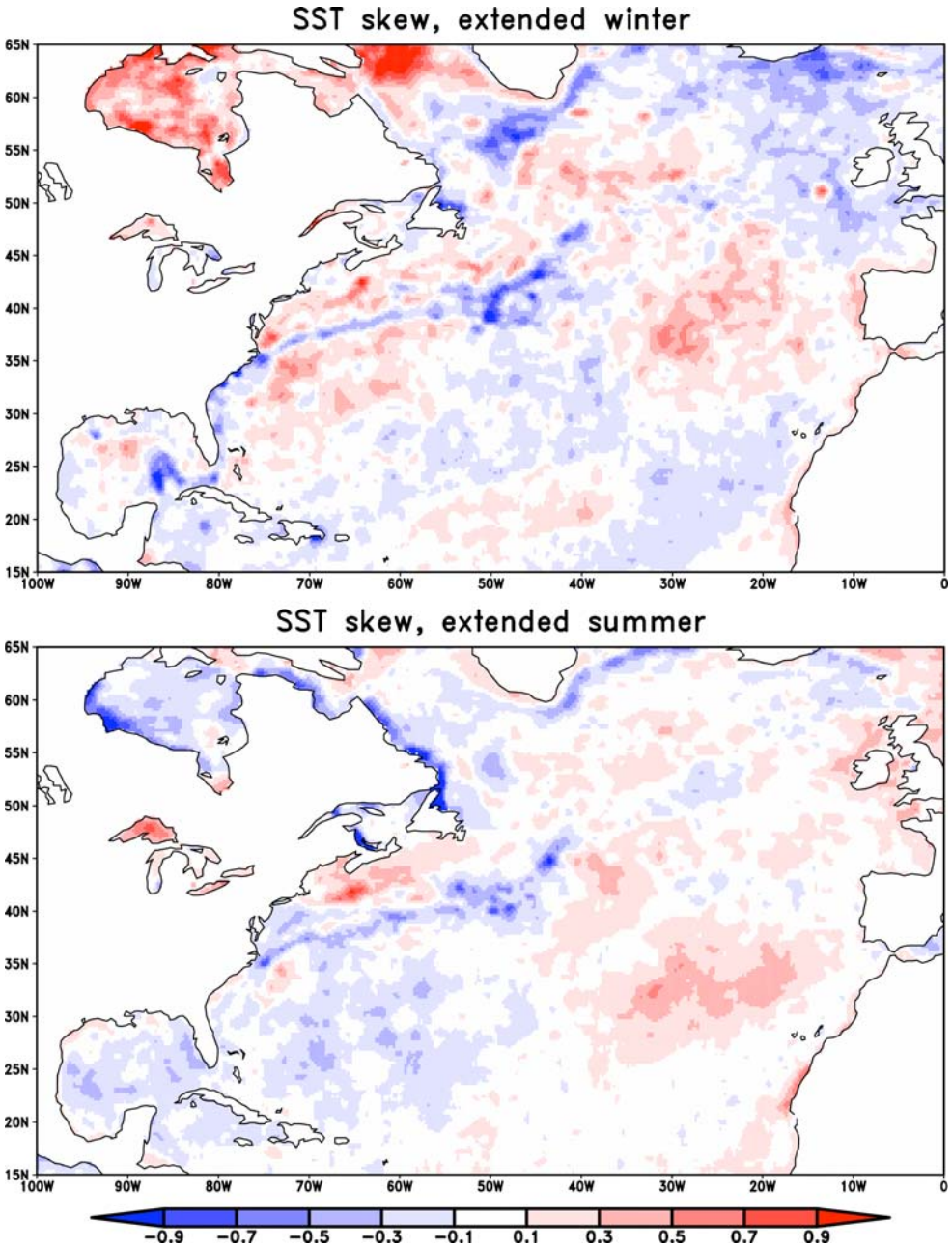

Fig. 1 Skewness of SST anomalies for (*top*) extended summer and (*bottom*) winter

from Florida, over Cape Hatteras, to the central North Atlantic. It is also noteworthy that there is very little seasonal dependence in the non-Gaussianity of the Gulf Stream system. That is, full-year Gulf Stream PDFs will be studied subsequently to reduce the uncertainty of the power-law statistics discussed below. Other features such as the Loop Current in the Gulf of Mexico and the East Greenland Current are also clearly visible by negative skewness in at least one of the extended seasons. However, in this paper, the focus is on the year-round non-Gaussianity of the Gulf Stream system.

Is it possible that the non-Gaussianity in the Gulf Stream system is due to biases of the AVHRR SSTs? In particular, could the non-Gaussianity in the Gulf Stream system be biased due to cloud cover? While we cannot completely rule out the possibility that the tails of the PDFs are somehow influenced by cloud cover, it seems very unlikely that a well defined and spatially coherent structure like the SST skewness of the Gulf Stream system (Fig. 1) is induced by a potential cloud cover bias. That is, there is little physical reason to believe that the coherent skewness pattern *strictly* following the Gulf Stream can be due to a more

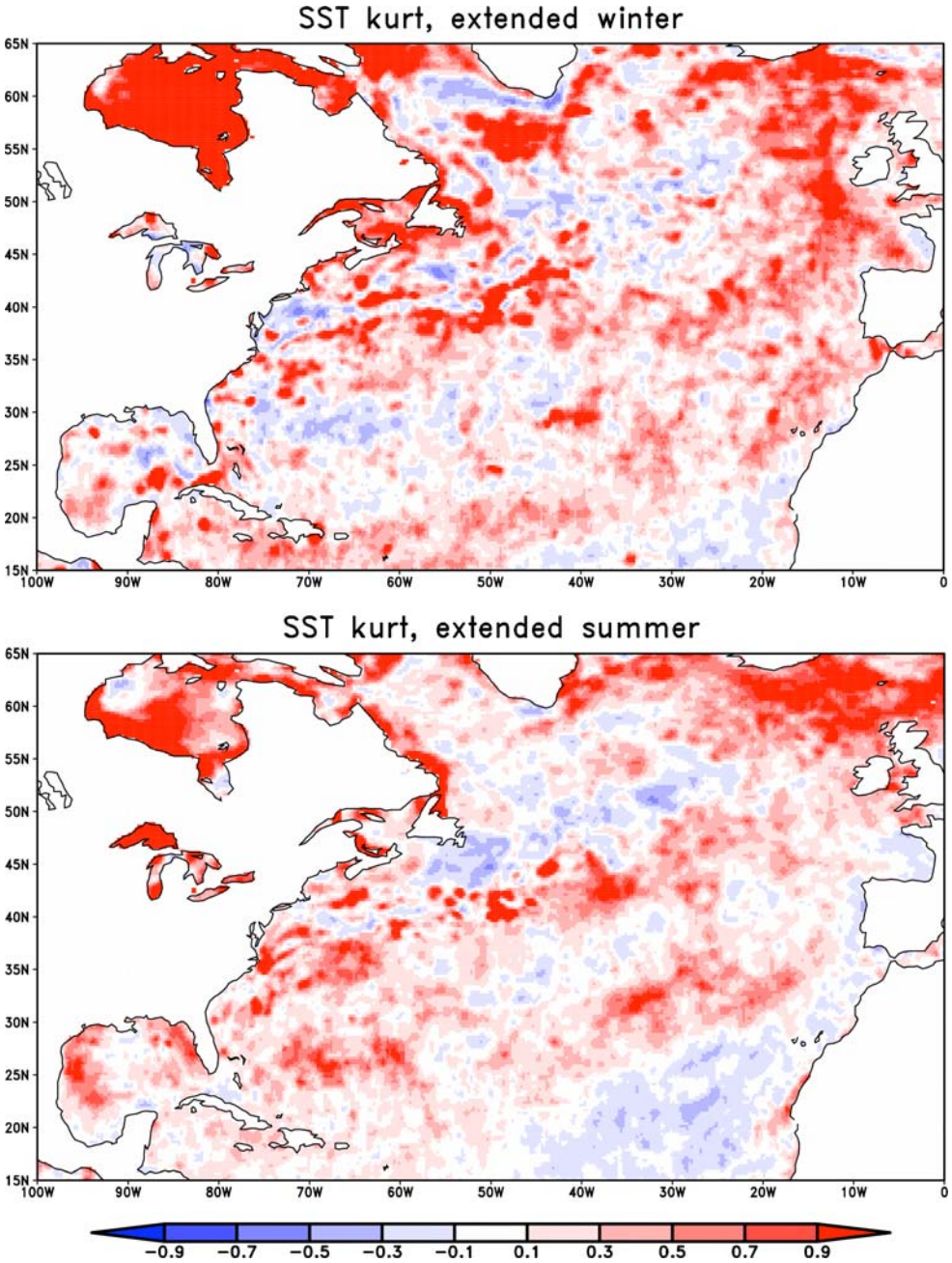

Fig. 2 Kurtosis of SST anomalies for (*top*) extended summer and (*bottom*) winter

large-scale atmospheric cloudiness pattern. In fact, even on a global scale, the observed AVHRR SST skewness and kurtosis patterns (Sura and Sardeshmukh 2008) are congruent with known physical properties of the upper ocean such as currents and fronts. Therefore, we are confident that the moments and PDFs of the AVHRR SSTs are indeed reliable and not due to cloud biases in the data set.

To get an idea about the detailed shape of the PDF, we picked several locations along the Gulf Stream and North Atlantic Current and calculated the full PDFs there. As expected from the minor seasonal dependence in SST skewness and kurtosis of the Gulf Stream system, the PDFs of extended winter and summer seasons have the same non-Gaussian structure. Thus, for the sake of clarity and considerable statistical significance, only the PDFs of full-year SST anomalies are presented. First, a parametric method is used to calculate the PDFs of observed SST anomalies, in which the parameters of a "skew-*t*" distribution are determined by a maximum likelihood estimate. The skew-*t* distribution, a skewed and kurtotic alternative to the normal distribution, is used because it is capable of adapting very closely to skewed and heavy-tailed

data (Azzalini and Capitanio 2003; Jones and Faddy 2003; Azzalini 2005; R Development Core Team 2004). Two representative PDFs are presented in Fig. 3. For the examples, two locations with strong skewness were chosen. In Fig. 3a, the PDF of full-year SST anomalies off Cape Hatteras (35° N, 75° W) is shown, whereas, in Fig. 3b, the PDF of full-year SST anomalies at a location in the North Atlantic Current (40° N, 50° W) are presented. The skew-*t* distribution is denoted by the thick solid black line, a simple histogram by the thin red line. The corresponding Gaussian distribution is also included in both figures as a dashed line. Note the logarithmic scale on the ordinates.

To examine the non-Gaussian structure of SST variability in the Gulf Stream system in even more detail, the raw PDFs (histograms) are also presented on a log–log scale to highlight the scaling properties of the data (Figs. 4 and 5). For example, a highly non-Gaussian power-law tail (that is, a PDF $P(x) \propto x^{-\alpha}$ with the exponent α) will appear as a straight line on a log–log plot. As shown below, PDFs of anomalous SST variability in the Gulf Stream system indeed obey a power-law, as tested by a Monte-Carlo approach using Kolmogorov–Smirnov statistics. Note that the power-law characteristics are estimated from the histogram (and not the skew-*t* fit) to avoid nested fits (fitting a

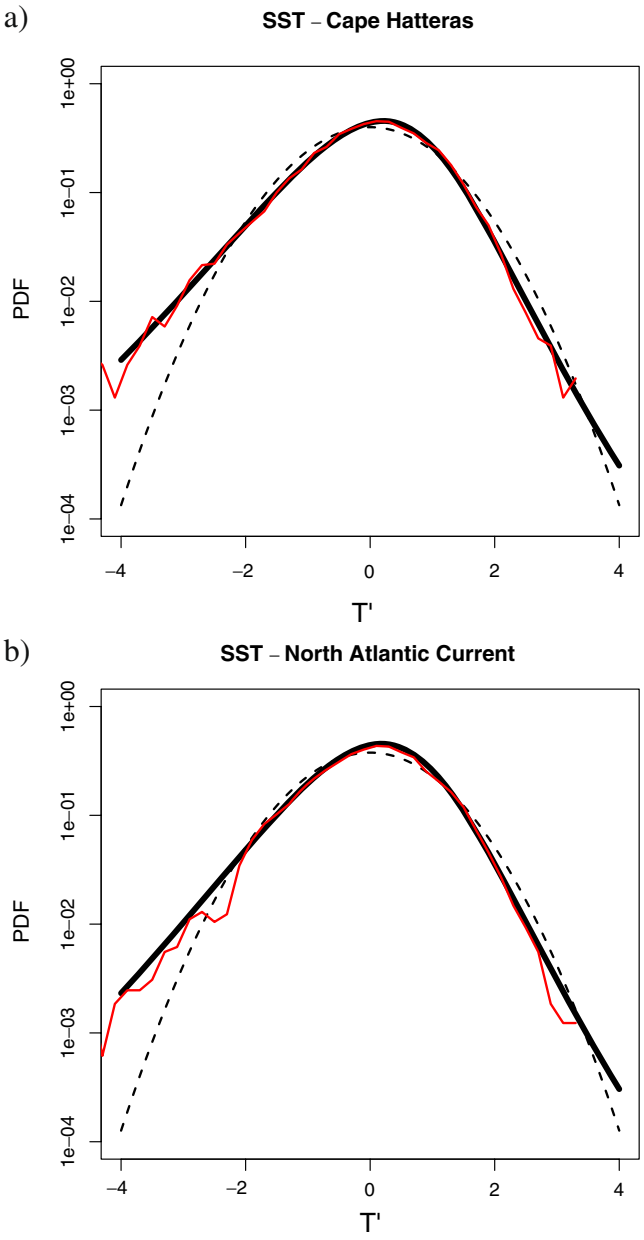

Fig. 3 **a** PDF of full-year SST anomalies off Cape Hatteras (35° N, 75° W). **b** PDF of full year SST anomalies at a location in the North Atlantic Current (40° N, 50° W). In both plots, the skew-*t* distribution is denoted by the *thick solid black line*, a simple histogram by the *thin red line*. The corresponding Gaussian distribution is included as a *dashed line*. Note the logarithmic scale on the ordinates

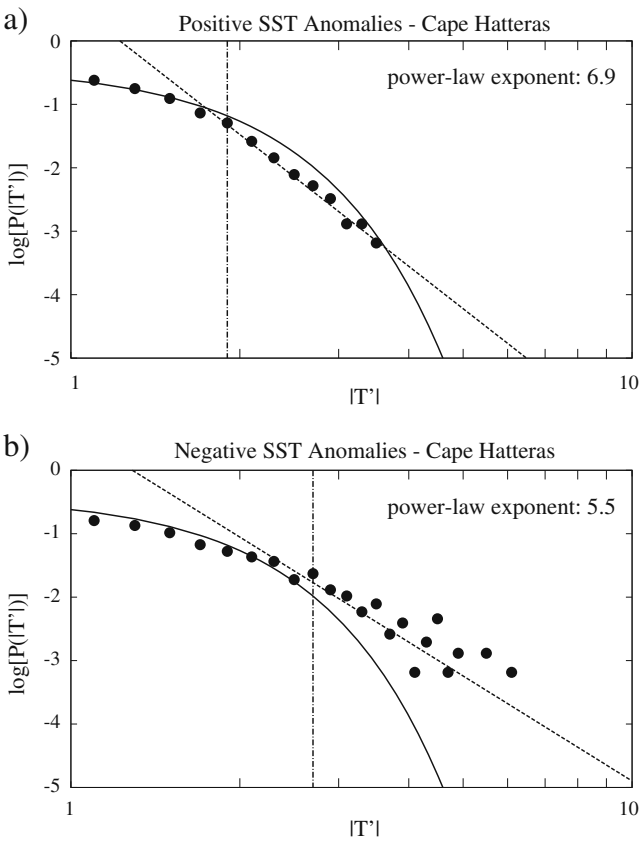

Fig. 4 Log-log scale PDFs (*circles*) of full-year **a** positive and **b** negative SST anomalies off Cape Hatteras (35° N, 75° W). The *straight dashed lines* are maximum likelihood estimates of the power-law behavior and the *solid lines* denote Gaussian distributions. The lower bounds of the power-law scaling $|T'|_{\min}$ are shown by the *dotted-dashed vertical lines*

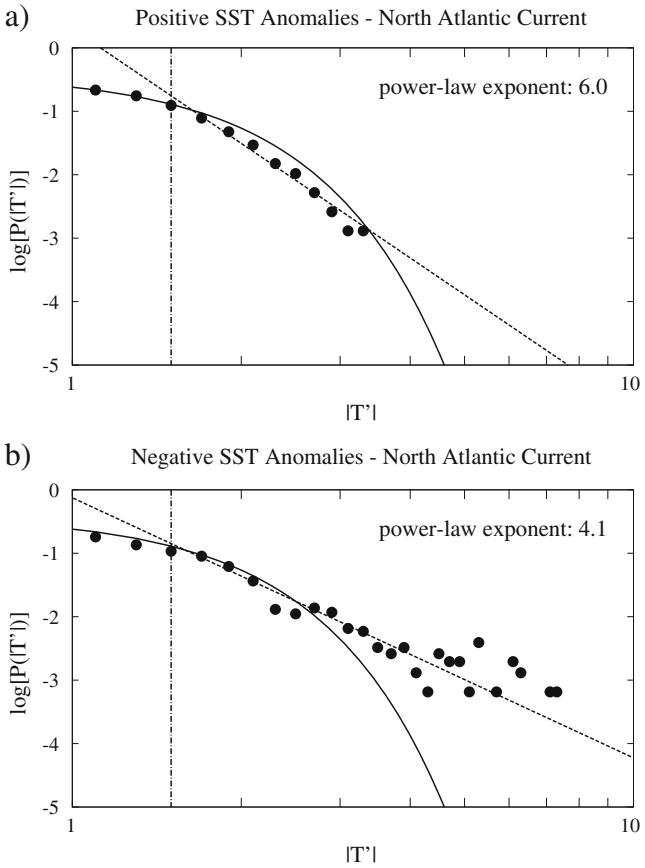

Fig. 5 Log–log scale PDFs (*circles*) of full-year **a** positive and **b** negative SST anomalies in the North Atlantic Current (40° N, 50° W). The *straight dashed lines* are maximum likelihood estimates of the power-law behavior and the *solid lines* denote Gaussian distributions. The lower bounds of the power-law scaling $|T'|_{\min}$ are shown by the *dotted-dashed vertical lines*

power-law to a skew-*t* fit). Because of the log scale on the *x*-axis, absolute values $|T'|$ are shown. In addition, because the emphasis is on extreme events, the center of the PDFs (± one standard deviation) are not shown. The PDFs of positive/negative SST anomalies at Cape Hatteras are shown in Fig. 4a and b, and the PDFs of positive/negative SST anomalies in the North Atlantic Current are shown in Fig. 5a and b. In all plots, the solid line denotes a Gaussian distribution. Note the distinct heavy-tail power-law behavior for negative anomalies giving rise to negative skewness. However, even the positive anomalies follow a power-law (weaker than Gaussian, though). The straight dashed lines are maximum likelihood estimates of the power-law behavior (given by the power-law exponent α also included in each plot) above a systematically estimated lower bound $|T'|_{\min}$. The lower bounds $|T'|_{\min}$ are shown by the dotted–dashed vertical lines. The procedure to estimate the power-law parameters and the related goodness-of-fit test are described in the following paragraph. The numerical values of α, $|T'|_{\min}$, and the results of a goodness-of-fit test (p values) are summarized in Table 1. In fact, in all cases, the power-law hypothesis is statistically significant (cannot be rejected) as tested by Kolmogorov–Smirnov statistics.

As shown in many publications (e.g., Sornette 2006; Newman 2005; Clauset et al. 2009), the best-fit (that is, maximum likelihood) power-law exponent α for a general positive timeseries x_i of length n is

$$\alpha = 1 + n \left[\sum_{i=1}^{n} \ln \left(\frac{x_i}{x_{\min}} \right) \right]^{-1} , \tag{2}$$

where $x_{\min}$ is the lower bound of the power-law behavior. Note that the general description in this section requires x_i and $x_{\min}$ to be positive and, therefore, not directly applicable to T'. Hence, Eq. 2 has to be applied separately to the absolute values of SST anomalies $|T'|$ for positive and negative perturbations. The approximate standard error ϵ of α can be derived from the width of the likelihood maximum as

$$\epsilon = \frac{\alpha - 1}{\sqrt{N_{\text{in}}}} , \tag{3}$$

where N_{in} is the effective number of independent observations. That is, the interval $\alpha \pm 1.96\epsilon$ contains the true power-law exponent with approximately 95% certainty. In our case, N_{in} is estimated by dividing the total number of observations n by the observed decorrelation timescale of SST anomalies. $x_{\min}$ is chosen to make the PDF of the observed data and the best-fit power-law model as similar as possible above $x_{\min}$ (Clauset et al. 2009). As a measure for quantifying the distance between two PDFs, the Kolmogorov–Smirnov statistic $\mathcal{D}$ (e.g., Wilks 2006; Press et al. 1992) is used (there are more sophisticated test statistics available, but as the non-Gaussianity is rather strong in our example, a basic test is appropriate here). $\mathcal{D}$ is simply the maximum

Table 1 Best-fit power-law and goodness-of-fit test parameters: power-law exponent α, lower power-law bound $|T'|_{\min}$, and p value of Kolmogorov–Smirnov statistics

Location	α	$T_{\min}$	p
Cape Hatteras (35° N, 75° W)			
Positive anomalies	6.9 ± 0.4	1.9	0.13
Negative anomalies	5.5 ± 0.3	2.7	0.87
North Atlantic Current (40° N, 50° W)			
Positive anomalies	6.0 ± 0.3	1.7	0.12
Negative anomalies	4.1 ± 0.2	1.5	0.73

distance between the cumulative density function (CDF) of the data and the best-fit model:

$$\mathcal{D} = \max_{x \geq x_{\min}} |S(x) - P(x)| \ , \tag{4}$$

where $S(x)$ is the CDF of the data above $x_{\min}$, and $P(x)$ is the CDF of the best-fit power-law model above $x_{\min}$. The best estimate of $x_{\min}$ is the value that minimizes $\mathcal{D}$. It is also straightforward to use Kolmogorov–Smirnov statistics to test if the observed data are consistent with a power-law distribution, that is, to perform a goodness-of-fit test. Here, this is done using a Monte-Carlo approach. First, the power-law exponent α and the cutoff $x_{\min}$ is calculated from the data. The goodness-of-fit to the data is quantified by the Kolmogorov–Smirnov distance $\mathcal{D}$. Then, using the observed parameters, a large number (here 10,000) of synthetic power-law timeseries of length N_{in} are generated and the power-law exponent α, lower-bound $x_{\min}$, and Kolmogorov–Smirnov distance $\mathcal{D}$ (now from its know CDF) are estimated for each of them employing the techniques described above. Finally, the fraction of the Kolmogorov–Smirnov statistics for the synthetic data sets whose value exceed the statistics of the real data yields the so called p value. Only if the p value is sufficiently small can a power-law distribution be ruled out. Note that we use Kolmogorov–Smirnov statistics and the related p value to reject (rather than confirm) a hypothesis for the observed data; hence, high p values, not low, are considered good. Here, a power-law is ruled out if $p \leq 0.1$. That is, if $p \leq 0.1$, there is a 10% or less probability to draw a power-law sample merely by chance that agrees that poorly with power-law statistics, and the hypothesis that the data are drawn from a power-law is rejected. See Newman (2005) and Clauset et al. (2009) for more details on how to estimate power-law distributions, including goodness-of-fit tests, from empirical data.

To summarize, the non-Gaussian structure with a heavy negative power-law tail, responsible for the negative skewness, can be clearly seen at both locations. Even the positive SST anomalies show a weak power-law behavior. The PDFs at other locations along the Gulf Stream and North Atlantic Current have the same structure as both the shown PDFs. This non-Gaussianity is, therefore, representative for the non-Gaussian SST variability in the Gulf Stream system and, perhaps, other strong currents. In particular, the observed skewness along the Gulf Stream and the North Atlantic Current is indicative of a specific physical process confined to the strong flow. In the next section, we will present a theory that explains the pronounced negative skewness and the power-law behavior in strong currents.

3 Theory

In the previous section, we have seen that the Gulf Stream system shows a very pronounced negative skewness of SST anomalies. In this section, we present a theory that explains this interesting feature of SST variability in strong currents. To develop the theory of non-Gaussian SST variability in strong currents, we have to first briefly rehash the closely related study of global non-Gaussian SST variability (see Sura and Sardeshmukh 2008, for details).

3.1 A univariate theory of global non-Gaussian SST variability

The starting point for the global view of non-Gaussian SST variability is the heat budget equation for T_o, defined as an average temperature over the mixed-layer depth h (e.g., Frankignoul 1985):

$$\frac{\partial T_o}{\partial t} = -\mathbf{v}_o \cdot \nabla T_o + \frac{Q}{qCh} - \frac{w_e}{h}\left(T_o - T_o^b\right) + \kappa \nabla^2 T_o \equiv F \ , \tag{5}$$

where the rate of change of T_o is governed by advection through ocean currents, surface heat fluxes, vertical entrainment, and horizontal mixing. Here, $\mathbf{v}_o$ is the horizontal velocity in the mixed-layer, Q is the heat flux through the sea surface, q and C are the density and heat capacity of sea water, w_e and T_o^b are the vertical velocity and temperature just below the mixed-layer, and κ is the horizontal mixing coefficient. F denotes the sum of all terms. Note that all of these terms are directly or indirectly affected by atmospheric quantities like winds and temperatures. For small temperature anomalies, T_o', a Taylor expansion of the heat flux F with respect to $T_o = \overline{T}_o + T_o'$ yields

$$\frac{\partial T_o'}{\partial t} = \frac{\partial F}{\partial T_o} T_o' + F' + R' \quad , \tag{6}$$

where it is assumed that the evolution of the mean temperature $\overline{T}_o$ is balanced by the mean heat flux $\overline{F}$ (using $F = \overline{F} + F'$), and that the derivative is evaluated at $\overline{T}_o$. We also added a residual term R' to represent higher-order terms of the Taylor expansion and other processes not included in the mixed-layer Eq. 5. Because the heat flux anomaly F' is, to a large degree, due to rapid transient atmospheric fluctuations, it may be represented as temporal noise $\eta_{F'}$. For the sake of

simplicity (and the lack of detailed knowledge), the residual R' is dealt with in the same way and represented as another (uncorrelated to $\eta_{F'}$) noise term $\eta_{R'}$. The derivative $\partial F/\partial T_o$, on the other hand, is usually represented by a constant parameter $-\lambda$. That is, the effect of atmospheric forcing on SST anomalies T_o' is often represented by a simple stochastic driving of the oceanic mixed layer,

$$\frac{\partial T_o'}{\partial t} = -\lambda T_o' + \eta \ , \tag{7}$$

where λ is a rate coefficient representing the damping of the slowly evolving mixed-layer heat anomaly, and $\eta = \eta_{F'} + \eta_{R'}$ is Gaussian white-noise representing the heat flux due to rapidly varying weather fluctuations and other residual processes. Equation 7 is, of course, the classic stochastic model of SST variability (Frankignoul and Hasselmann 1977) in particular, and low-frequency climate variability in general (Hasselmann 1976).

The stochastic view in Eq. 7 implies that T_o' has a Gaussian PDF. Indeed, temporally (monthly, seasonally, or even yearly) or spatially (e.g., several degrees) averaged SST anomalies are nearly Gaussian. We expect this partly from the Central Limit Theorem (e.g., Gardiner 2004; Paul and Baschnagel 1999) to the extent that it is applicable to time-averaged quantities. However, we have already seen in Figs. 1–5 that, on daily scales, observed SSTs are significantly non-Gaussian. In particular, we showed in Sura and Sardeshmukh (2008) that there is a striking parabolic relationship between skewness and kurtosis, kurt $\geq (3/2)\text{skew}^2$. Sura and Sardeshmukh (2008) used a multiplicative noise model to explain this observed constraint. The derivation is sketched in the following.

We already noticed (Eq. 6) that a Taylor expansion of the heat flux with respect to $T_o = \overline{T}_o + T_o'$ yields

$$\frac{\partial T_o'}{\partial t} = \frac{\partial \overline{F}}{\partial T_o} T_o' + \frac{\partial F'}{\partial T_o} T_o' + F' + R' \quad , \tag{8}$$

where, in contrast to Eq. 6, we just replaced the full heat flux F with its mean and perturbation, $F = \overline{F} + F'$. As before in Eq. 6, the derivatives are evaluated at $\overline{T}_o$. This, at first glance, trivial replacement is done to highlight the fact that the derivative $\partial F/\partial T_o$ actually consists of two terms: the constant term $\partial \overline{F}/\partial T_o$ and the rapidly varying term $\partial F'/\partial T_o$. As a relevant example closely related to Frankignoul and Hasselmann (1977), consider the form of a common bulk heat flux parameterization $F \equiv F_Q = \beta(T_a - T_o)|\mathbf{U}|$, which depends on the SST T_o, air temperature T_a, wind speed $|\mathbf{U}|$, and a constant β. We set $\overline{F}_Q = \beta(\overline{T_a} - \overline{T_o})\overline{|\mathbf{U}|}$ and $F_Q' = \beta(\overline{T_a} - \overline{T_o})|\mathbf{U}|'$, allowing for rapidly varying wind speed fluctuations $|\mathbf{U}|'$. Note that the mean SST $\overline{T_o}$ is still allowed to vary very slowly. We then obtain $\partial \overline{F}_Q/\partial T_o = -\beta\overline{|\mathbf{U}|}$ and $\partial F_Q'/\partial T_o = -\beta|\mathbf{U}|'$, or, equivalently, $\partial F_Q/\partial T_o = -\beta|\mathbf{U}| = -\beta(\overline{|\mathbf{U}|} + |\mathbf{U}|')$. It is the constant term that justifies the introduction of the constant feedback parameter $-\lambda$ in the general formulation Eq. 7. However, the rapidly varying term $\partial F'/\partial T_o$ cannot be neglected as is done in many studies (e.g., in Frankignoul and Hasselmann 1977). Why has this term been neglected in previous studies?

The main reason is that, in older, coarse (in space and time) resolution data sets, SST variability is close to being Gaussian, mainly because of the Central Limit Theorem. Then, the classic stochastic model of SST variability is sufficient to describe the Gaussian statistics of the data. Only recently, the new high-resolution data sets became available, allowing for a more detailed analysis of non-Gaussian statistics [the exception being investigators using Ocean Weather Station data, such as Müller 1987; Blaauboer et al. 1982]. That means we are using the same very general assumption (slow–fast timescale separation) as in the classic papers, but include non-Gaussian effects through the inclusion of the so far neglected term $\partial F'/\partial T_o$. In fact, the success of Sura and Sardeshmukh (2008) suggests that the timescale separation assumption is indeed valid for non-Gaussian high-resolution SST dynamics, the same way it is for Gaussian coarse-resolution SSTs. So how to model the rapidly varying derivative $\partial F'/\partial T_o$?

The easiest way to model this term is to linearize it. That is, we assume that the derivative $\partial F'/\partial T_o$ is proportional to the rapidly varying forcing F' itself: $\partial F'/\partial T_o \approx -\phi F'$, where ϕ is a locally constant parameter. Because ϕ can be positive or negative, the minus sign is arbitrary and just introduced for the sake of convenience. As shown next, within the framework of small temperature anomalies T_o', this approximation means that F' depends weakly linearly on T_o and, therefore, imposes no major constraint on F' giving the arbitrary parameter ϕ. The equation $\partial F'/\partial T_o = -\phi F'$, of course, means that $F' = \eta(t)\exp(-\phi T_o)$, with the time-dependent, arbitrary amplitude white-noise term $\eta(t)$. For small temperature anomalies, the exponential function can be very well linearized as $F' = \eta(t)\exp[-\phi(\overline{T}_o + T_o')] \approx \eta(t)\exp(-\phi\overline{T}_o)(1 - \phi T_o')$. Plugging this F' back into our starting Eq. 8, we obtain exactly the same functional form we started with, namely, correlated additive and linear multiplicative noise terms. That means that we made a physically plausible and consistent approximation by using $\partial F'/\partial T_o = -\phi F'$.

To summarize, we approximate the rapidly varying derivative $\partial F'/\partial T_o$ as a constant times F', $\partial F'/\partial T_o = -\phi F'$, to get the following Stratonovich SDE for SST anomalies T'_o:

$$\frac{\partial T'_o}{\partial t} = -\lambda_{\text{eff}} T'_o - \phi F' T'_o + F' + R' + \phi \overline{F' T'_o} \quad , \qquad (9)$$

with the locally constant parameters $-\lambda_{\text{eff}}$ and $-\phi$ and the rapidly varying (approximately Gaussian white-noise) forcing terms F' and R'. That is, we approximated F' and R' as independent, zero mean Gaussian white-noise processes with amplitudes $\sigma_{F'}$ and $\sigma_{R'}$: $\overline{F'(t)F'(t')} = (\sigma_{F'})^2\delta(t-t')$ and $\overline{R'(t)R'(t')} = (\sigma_{R'})^2\delta(t-t')$. Recall that, in the Stratonovich calculus, we have to include the noise-induced drift in the effective drift: $-\lambda_{\text{eff}} = -\lambda + (1/2)(\phi\sigma_{F'})^2$. The noise-induced drift appears in Stratonovich systems because then the time mean of the multiplicative noise term, $-\phi\overline{F'T'_o}$, is not zero. This is also why we have to introduce the additional mean forcing $\phi\overline{F'T'_o}$ to guarantee that the time mean of T'_o is zero, $\overline{T'_o} = 0$.

At this point, it is helpful to recall the scaling properties of the noise terms in *continuous* SDEs such as Eq. 9 to avoid confusion about dimensions. This scaling, of course, comes from the fact that SDEs are properly defined only in the integral sense. A brief summary of how to interpret SDEs is given in Appendix A [for a detailed discussion, see, e.g., Gardiner 2004, Kloeden and Platen 1992, Paul and Baschnagel 1999, or any other textbook on SDEs].

One basic and very important result is that the variance of white-noise (the so called Wiener process W) scales with the time increment dt. Casually, we often say the that Wiener process effectively scales with $\sqrt{dt}$ in contrast to the dt scaling of the determinist terms, when we actually mean the standard deviation of the Wiener process (see below and Appendix A). In a nutshell, it is this $\sqrt{dt}$ scaling that is responsible for so called *stochastic calculus*, in contrast to the dt scaling of the ordinary *deterministic calculus*. For example, if we integrate over white-noise F' with amplitude $\sigma_{F'}$ (that is, $F' = \sigma_{F'}\eta$ with the normalized, zero mean, and unit standard deviation, white-noise η), we obtain $\int F'dt = \sigma_{F'}\int \eta\, dt = \sigma_{F'}\int dW$, where $dW \equiv \eta\, dt$ is the increment of a Wiener process. However, as shown in Appendix A, dW has the dimensional variance dt: $\langle dW^2\rangle = dt\ (= \overline{dW^2})$. That is, the Wiener process itself scales with $\sqrt{dt}$. That, in turn means that $\sigma_{F'}$ does not have the same units as F' itself, but has to be scaled by $\sqrt{dt}$.

The multiplicative noise system Eq. 9 has one important property of interest here. In general, Eq. 9 will produce non-Gaussian statistics. As shown in Sura and Sardeshmukh (2008), the second, third, and fourth moments of the SDE Eq. 9 are (with $T'_o \equiv x$)

$$\overline{x^2} = \frac{\left(\sigma^2_{F'} + \sigma^2_{R'}\right)}{\left[2\lambda_{\text{eff}} - (\phi\sigma_{F'})^2\right]} \ ,$$

$$\overline{x^3} = \frac{-2\phi\sigma^2_{F'}\overline{x^2}}{\left[\lambda_{\text{eff}} - (\phi\sigma_{F'})^2\right]} \ ,$$

$$\overline{x^4} = \frac{\left[-3\phi\sigma^2_{F'}\overline{x^3} + (3/2)\left(\sigma^2_{F'} + \sigma^2_{R'}\right)\overline{x^2}\right]}{\left[\lambda_{\text{eff}} - (3/2)(\phi\sigma_{F'})^2\right]} \ . \qquad (10)$$

These moment equations result in the general relation between skewness and kurtosis, kurt $\geq (3/2)\text{skew}^2$, in almost perfect agreement with observations (Sura and Sardeshmukh 2008). In the context of this study, it is important to note that T'_o is positively (negatively) skewed if ϕ is negative (positive). This follows from the fact that, for the fourth moment to exist, there is an upper limit for the strength of the multiplicative noise: $(\phi\sigma_{F'})^2 < (2/3)\lambda_{\text{eff}}$. In such conditions, the term $[\lambda_{\text{eff}} - (\phi\sigma_{F'})^2]$ is always positive, and the sign of the skewness is solely determined by the sign of ϕ. Put differently, the skewness is positive if the additive and multiplicative noises are positively correlated (that is, have the same sign), and negative if the noises are correlated negatively (that is, have opposite signs). This behavior is closely connected to the following other properties (derived in Appendix B) we will use later. First, $\overline{F'T'^2_o}\phi \leq 0$. That is, if T'_o is positively skewed ($\phi < 0$), the third-order moment $\overline{F'T'^2_o}$ is also positive (and vice versa). Second, in the pure additive noise case $\phi = 0$, we have $\overline{F'T'^2_o} = 0$ (see Appendix B). Third, the noises and T'_o are positively correlated, $\overline{F'T'_o} > 0$ and $\overline{R'T'_o} > 0$.

Besides the skewness and kurtosis link of the multiplicative noise system, Eq. 9, there is one more important non-Gaussian property of interest here. As shown in Sardeshmukh and Sura (2009), Eq. 9 will produce power-law tails. This can be easily seen from the Fokker–Planck equation for the stationary PDF p of SST anomalies $T'_o \equiv x$ governed by Eq. 9:

$$\frac{d}{dx}(\lambda_{\text{eff}}\, x\, p) + \frac{1}{2}\frac{d^2}{dx^2} \times \left[\left(\sigma^2_{F'} + \sigma^2_{R'} + \phi^2\sigma^2_{F'}x^2 - 2\phi\,\sigma^2_{F'}\,x\right)p\right] = 0 \ , \qquad (11)$$

or, after one integration,

$$\lambda_{\text{eff}}\, x\, p + \frac{1}{2}\frac{d}{dx}\left[\left(\sigma^2_{F'} + \sigma^2_{R'} + \phi^2\sigma^2_{F'}x^2 - 2\phi\,\sigma^2_{F'}\,x\right)p\right] = const. \equiv 0 \ . \qquad (12)$$

Note that, in general, the constant has to be zero to fulfill the boundary conditions of vanishing probability at $x = \pm\infty$. For large x and small noise amplitudes, this equation simplifies to

$$\lambda_{\text{eff}}\, x\, p + \frac{1}{2}\frac{d}{dx}\left(\phi^2\, \sigma_{F'}^2\, x^2\, p\right) = 0\,. \tag{13}$$

Using the power-law ansatz $p \propto |x|^{-\alpha}$, one easily finds that p is a solution if

$$\alpha = 2\left[\frac{\lambda_{\text{eff}}}{\phi^2\, \sigma_{F'}^2} + 1\right]. \tag{14}$$

That is, Eq. 9 will produce power-law tails in agreement with observations in the Gulf Stream system (Figs. 4 and 5). Note that there exists a closed solution $p(x)$ of the full Fokker–Planck Eq. 11 (see, e.g., van Kampen 1981; Müller 1987; Sardeshmukh and Sura 2009) which has, of course, the same power-law tails. For the sake of simplicity, we omitted the full solution here to merely present the simple derivation of the power-law exponent.

Equation 14 tells us to expect heavier tails (smaller α) for a weaker damping (or stronger multiplicative noise forcing $\phi^2\, \sigma_{F'}^2$) and vice versa. It should be noted, though, that the limit of very small $\phi^2\, \sigma_{F'}^2$ (resulting in $\alpha \to \infty$) is not allowed within the approximation we made: $\phi^2\, \sigma_{F'}^2 \to 0$ counteracts the approximation of large x. The limit $\phi^2\, \sigma_{F'}^2 \to 0$ simply means that the multiplicative noise is negligible, and the PDF $p(x)$ becomes Gaussian. It has already been shown that the power-law exponent α has values between about four and seven in the Gulf Stream system. Interestingly, negative anomalies exhibit heavier tails (smaller exponents α) than positives ones. In this stochastic framework, this points to a slightly weaker damping λ_{eff} (or stronger multiplicative noise forcing $\phi^2\, \sigma_{F'}^2$) of negative anomalies. It is interesting that this is consistent with observational studies showing a positive local (on scales smaller than 1,000 km) correlation between SSTs and sea surface winds (Chelton et al. 2004; Small et al. 2005; Xie 2004; O'Neill et al. 2003). Negative SST anomalies mean weaker mean winds and, therefore, a weaker damping. The study of this phenomenon in the presence of a strong current is part of our ongoing research.

Equation 14 also points to an interesting way to estimate the strength of the multiplicative noise in Eq. 9. So far, we have not come up with a stable and reliable method to estimate all parameters of the multiplicative noise SDE Eq. 9 from relatively short records in order to directly compare modeled and observed power-law exponents and PDFs. In general, it is non-trivial to estimate coefficients of SDEs from limited data (e.g., Kloeden and Platen 1992; Sura and Barsugli 2002). One method we are exploring is to use the observational estimates of λ_{eff} and the moments $\overline{x^2}$, $\overline{x^3}$, and $\overline{x^4}$ to solve the nonlinear set of Eq. 10 for the remaining noise parameters $\sigma_{F'}^2$, $\sigma_{R'}^2$, and ϕ. Preliminary results show a satisfactory agreement of the power-law exponent α calculated from Eq. 14 with the direct estimate from data using Eq. 2. However, the method is fraught with the potential for error and uncertainty due to the nonlinearity of the equations involved. Therefore, we might use the stable estimate of α from Eq. 2 and the easy-to-determine effective damping parameter λ_{eff} to solve Eq. 14 for $\phi^2\, \sigma_{F'}^2$. The remaining additive noise could then be determined from Eq. 10. However, to explore and discuss estimation techniques for SDEs is beyond the scope of this paper, but also part of our current research.

3.2 Non-Gaussian SST variability in strong currents

The general theory discussed above does not specify the value, magnitude, or even the sign of the parameter ϕ. Therefore, the general idea outlined above does allow for positive and negative skewness, depending on the sign of ϕ. In order to closer determine ϕ, it is necessary to take a more detailed look at the terms included in the overall heat flux F. So far, we only used the fact that the timescale of the forcing F is faster than the SST response. Is is already known that, away from strong currents, the heat flux $Q/\rho CH$ through the sea surface is the dominating term over most parts of the ocean. Aspects of the stochastic nature of this term have already been discussed in Sura et al. (2006). For example, while the heat flux $Q/\rho CH$ is not entirely random on a daily scale (large-scale pressure systems may have a decorrelation timescale of up to a week), it can be well modeled as a random process on a daily scale because the "slow" ocean does not care about the spectral form of the "fast" atmospheric forcing below its decorrelation scale of at least about 2 or 3 weeks. That is, as long there is some spectral gap between the ocean and the atmosphere, a stochastic forcing of the ocean is justified (see also Frankignoul and Hasselmann 1977; Hasselmann 1976). However, most relevant here is the fact that the heat flux $Q/\rho CH$ will induce positive skewness, that is, a negative ϕ. This can be easily seen from the common form of the bulk formula for the flux $Q \propto (T_a - T_o)|\mathbf{U}|$, which depends on the SST T_o, air temperature T_a, and wind speed $|\mathbf{U}|$. That is, the evolution equation for SST driven solely by fluxes through the sea surface is

$$\frac{\partial T_o}{\partial t} = \beta(T_a - T_o)|\mathbf{U}| \equiv F_Q\ , \tag{15}$$

where all constants are included in β. Therefore, $\partial F_Q/\partial T_o = -\beta|\mathbf{U}| = -\beta(\overline{|\mathbf{U}|} + |\mathbf{U}|')$ and $F'_Q = \beta(\overline{T_a} - \overline{T_o})|\mathbf{U}|'$. Because the annual mean temperature difference $\overline{T_a} - \overline{T_o}$ is negative (the ocean is warmer than the atmosphere) over large swaths of the ocean (with the exception of regions with very cold currents, such as areas south of 40°, to the west off California, and the cold tongue of the equatorial Pacific), the additive and multiplicative noise terms due to the wind speed anomaly $|\mathbf{U}|'$ are mostly positively correlated (that is, have the same sign), resulting in positively skewed SST anomalies T'_o. That is, SST anomalies driven by the heat flux Q are mostly positively skewed due to the mean atmosphere–ocean temperature difference with a warmer ocean combined with the structure of the bulk heat flux parameterization. It has been shown that this relation also holds for locally coupled air–sea interaction models (Sura and Sardeshmukh 2009). To summarize, the heat flux Q cannot explain the negative skewness observed in strong currents such as the Gulf Stream. Thus, we have to look for another mechanism. The key candidate in a strong current is, of course, the advection term $-\mathbf{v}_o \cdot \nabla T_o$. Therefore, let us study the dynamics of this term.

The corresponding evolution equation reads

$$\frac{\partial T_o}{\partial t} + \mathbf{v}_o \cdot \nabla T_o = D \ , \tag{16}$$

where D denotes a dissipative process such that $\overline{DT_o} \leq 0$ (for example, the conventional harmonic diffusion has this property). Equation 16 is, of course, the advection equation of a nearly conserved tracer (here SST). An equation for the variance of SST anomalies can be obtained by decomposing the temperature and velocity fields into the usual time-mean and eddy components, and then multiplying the resulting SST anomaly equation by the SST anomaly itself. After averaging, we obtain

$$\frac{1}{2}\frac{\partial}{\partial t}\overline{T_o'^2} + \overline{\mathbf{v}'_o T'_o} \cdot \nabla \overline{T_o} + \frac{1}{2}\overline{\mathbf{v}}_o \cdot \nabla \overline{T_o'^2} + \frac{1}{2}\nabla\left(\overline{\mathbf{v}'_o T_o'^2}\right) = \overline{D' T'_o} \, , \tag{17}$$

where we again may assume for the dissipation $\overline{D' T'_o} \leq 0$ (e.g., Vallis 2006). Very often, the time dependence, the advection by the mean current in a zonal flow, and the third-order term in Eq. 17 are neglected to obtain the balance $\overline{\mathbf{v}'_o T'_o} \cdot \nabla \overline{T_o} \approx \overline{D' T'_o} \leq 0$, stating that, on average, the eddy temperature flux is downgradient in regions of dissipation (e.g., Vallis 2006). In the following, we apply Eq. 17, keeping the third-order term, to an idealized, baroclinically unstable, strictly zonal warm-core current driven by a meridional pressure and temperature (linked through the thermal-wind balance) gradient. The condition of strict zonality ensures that all zonal gradients vanish in Eq. 17. While this is, of course, a simplification of the real world, it serves as a useful approximation to dynamically discuss the advection of SST anomalies along-stream a strong current such as the Gulf Stream. Therefore, with the velocity vector $\mathbf{v}_o = (u_o, v_o)$, the spatial terms become

$$\overline{\mathbf{v}'_o T'_o} \cdot \nabla \overline{T_o} = \left(\overline{u'_o T'_o}, \overline{v'_o T'_o}\right) \cdot \left(0, \frac{\partial \overline{T_o}}{\partial y}\right) = \overline{v'_o T'_o} \, \frac{\partial \overline{T_o}}{\partial y} \ , \tag{18}$$

$$\overline{\mathbf{v}}_o \cdot \nabla \overline{T_o'^2} = (\overline{u}_o, 0) \cdot \left(0, \frac{\partial \overline{T_o'^2}}{\partial y}\right) = 0 \ , \tag{19}$$

and

$$\nabla\left(\overline{\mathbf{v}'_o T_o'^2}\right) = \nabla\left(\overline{u'_o T_o'^2}, \overline{v'_o T_o'^2}\right) = \frac{\partial}{\partial y}\overline{v'_o T_o'^2} \ . \tag{20}$$

Assuming a statistically steady state, the final form of the SST anomaly variance equation in a strictly zonal flow becomes

$$\overline{v'_o T'_o}\,\frac{\partial \overline{T_o}}{\partial y} + \frac{1}{2}\frac{\partial}{\partial y}\overline{v'_o T_o'^2} = \overline{D' T'_o} \ . \tag{21}$$

In the following, we discuss Eq. 21, whereby we also neglect friction ($\overline{D' T'_o} = 0$) to highlight the dynamics of the third-order term in our simplified setting.

To make further progress, we now make a common approximation used to study the statistics of a passive tracer: we consider the problem of *stochastic advection*. That is, we assume that the SST is advected by a random velocity field $\mathbf{v}'_o$. This is the key step to link Eq. 21 to our stochastic SST anomaly Eq. 9. Remember, that we have a very good stochastic model of non-Gaussian SST anomaly variability in the SDE Eq. 9, where we have not specified the detailed physical origin of the multiplicative noise F'. However, from the general form of the advection, Eq. 16, we know that a stochastic velocity field $\mathbf{v}'_o$ will appear as a multiplicative noise term in a general SDE (e.g., Sardeshmukh and Sura 2009; Majda et al. 2008). That is, for the stochastic advection problem, we can set F' proportional to the meridional component v'_o of the random velocity vector $\mathbf{v}'_o$: $F' \propto v'_o$. That is, we obtain the inequalities (see above)

$$\overline{v'_o T_o'^2}\phi \leq 0 \quad \text{and} \quad \overline{v'_o T'_o} > 0 \ , \tag{22}$$

linking the tracer variance Eq. 21 with the SDE Eq. 9. What do we learn from Eq. 22 in connection with Eq. 21? The remainder of this section tries to answer that question. To facilitate the discussion, a schematic

drawing of the underlying dynamical balance is provided in Fig. 6.

Let us start by summarizing what we already know, including our assumptions. We derived Eq. 21 having a strictly zonal, frictionless warm-core current (an idealized version of the Gulf Stream) in mind. The mean temperature is assumed to be warmest in the center (core) of the current; the boundaries are assumed to be colder. That implies, of course, that the gradient $\partial\overline{T_o}/\partial y > 0$ south of the center, and $\partial\overline{T_o}/\partial y < 0$ north of it. This behavior is depicted on the right-hand side of Fig. 6. We also know that, if we observe negative skewness (equivalent to $\phi > 0$) of stochastically advected SST anomalies (as in the Gulf Stream system), we have $\overline{v_o' T_o'^2} < 0$, and vice versa. Furthermore, on the meridional boundaries of the current, the skewness of SST anomalies vanishes, implying $\phi = 0$ and $\overline{v_o' T_o'^2} = 0$. This gives us the sign of the derivative: $\partial\overline{v_o' T_o'^2}/\partial y < 0$ south of the warm core, and $\partial\overline{v_o' T_o'^2}/\partial y > 0$ north of the warm core (also depicted in Fig. 6).

We are now in a position the study the balance

$$\overline{v_o' T_o'}\,\frac{\partial\overline{T_o}}{\partial y} + \frac{1}{2}\frac{\partial}{\partial y}\overline{v_o' T_o'^2} = 0 \tag{23}$$

in more detail:

- South of the center (warm core), we have $\partial\overline{T_o}/\partial y > 0$ and $\partial\overline{v_o' T_o'^2}/\partial y < 0$. To balance the terms, we need a northward (up-gradient) eddy heat flux $\overline{v_o' T_o'} > 0$, consistent with our stochastic advection model.
- North of the center (warm core), we have $\partial\overline{T_o}/\partial y < 0$ and $\partial\overline{v_o' T_o'^2}/\partial y > 0$. Again, to balance the terms, we need a northward (now down-gradient) eddy heat flux $\overline{v_o' T_o'} > 0$, again consistent with our model.

Several interesting features emerge. First, in a stochastic non-Gaussian setting, friction is not explicitly necessary to balance the eddy heat fluxes. The divergence of the third-order correlation $\overline{v_o' T_o'^2}$ is sufficient to balance the northward heat flux in a zonal warm-core current. While this is an idealization of the real world (because randomly advected water parcels carry heat and momentum), it is enlightening to have a model that gives meaningful results without relying on poorly understood frictional effects. Second, the non-Gaussian view is consistent with observations, provided that the stochastic advection model is applicable to tracer transports in the Gulf Stream region. However, as discussed above, we have clear observational evidence that the stochastic view is a very good model of non-Gaussian SST variability, and that the assumptions made above are, therefore, approximately valid. In addition, the *derived* requirement that the eddy heat flux is solely northward ($\overline{v_o' T_o'} > 0$) in the Gulf Stream System is also consistent with observations and models (Stammer 1998; Jayne and Marotzke 2002). Third, by intuition, one could expect the Gulf Stream SSTs to be negatively skewed. In the warm Gulf Stream, the strongest cold anomalies result from rings and meanders that bring cold water southward from regions north of stream. Warm anomalies due to other processes, such as radiative effects and advection from downstream,

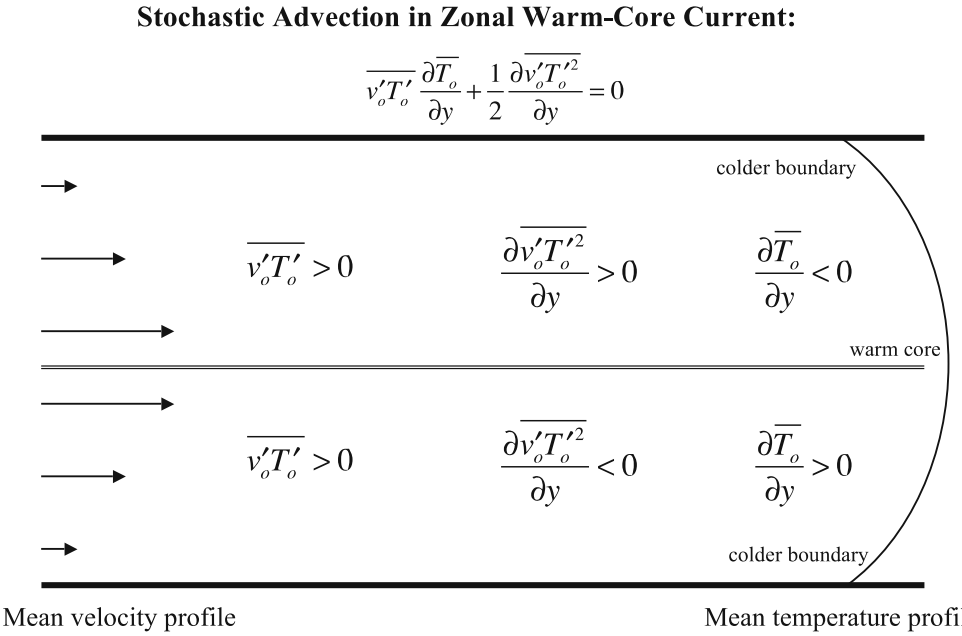

Fig. 6 Schematic drawing of the dynamical balance of the SST variance Eq. 21 in a strictly zonal, frictionless warm-core current. The mean temperature is assumed to be warmest in the center (core) of the current; the boundaries are assumed to be colder. That implies that the gradient $\partial\overline{T_o}/\partial y > 0$ south of the center, and $\partial\overline{T_o}/\partial y < 0$ north of it. The sign of the third-order derivative can also deduced from observations: $\partial\overline{v_o' T_o'^2}/\partial y < 0$ south of the warm core, and $\partial\overline{v_o' T_o'^2}/\partial y > 0$ north of the warm core. This balance then requires a northward eddy heat flux: $\overline{v_o' T_o'} > 0$. See text for more details

might well be weaker and less frequent. This is basically the intuition that the net eddy heat flux is northward in the Gulf Stream. The fact that this comes out of our independent stochastic analysis is a satisfying result.

To summarize, stochastic advection of SST is a good candidate to account for the negative SST anomaly skewness observed in the Gulf Stream system. As we have seen, the evolution equation for SST driven solely by heat fluxes through the seas surface Eq. 15 would give us positively skewed SST anomalies. That is, a combination of SST advection and heat flux forcing could result in negative or positive skewness. It is the relative strength of each process responsible for the net skewness. In very strong currents such as the Gulf Stream, the SST advection is strong enough to significantly affect the mixed-layer heat budget, resulting in the pronounced negative SST skewness observed along-stream strong currents. However, the heat fluxes through the seas surface are still important to provide the damping of SST anomalies without dominating the net skewness.

4 Summary and discussion

This study has looked at non-Gaussian SST variability in the Gulf Stream system. It has been shown from observations that the Gulf Stream (and its extension, the North Atlantic Current) is very clearly visible as a band of negative skewness all the way from Florida, over Cape Hatteras, to the central North Atlantic. To get an idea about the detailed non-Gaussian variability along the Gulf Stream, PDFs were calculated at several locations. One important observational result of this study is that the non-Gaussianity in the Gulf Stream system is consistent with a power-law, as tested by Kolmogorov–Smirnov statistics. The study then provides a theoretical explanation of the negatively skewed power-law behavior of SST anomalies in strong currents.

The model that explains the power-law and the negative skewness of SST anomalies in strong currents is based on a recent stochastic theory of non-Gaussian SST variability by Sura and Sardeshmukh (2008). That theory states that a stochastic approximation, including multiplicative noise, of the general mixed-layer SST budget is an excellent globally applicable model of non-Gaussian SST variability. The first candidate for a stochastic forcing of SST anomalies is, of course, the wind-driven surface heat flux. However, it is known (Sura et al. 2006; Sura and Sardeshmukh 2009) that the surface heat flux alone cannot explain the coherent negative skewness observed in swift currents such as the Gulf Stream. Therefore, the starting hypothesis is that SST advection in the mixed layer significantly contributes to the negatively skewed non-Gaussian SST variability in strong currents. It is then shown that stochastic advection of SSTs is consistent with the observed properties of the Gulf Stream system.

To summarize, SST advection in strong ocean currents very likely contributes to the induction of extreme negative SST anomalies in such regions. The discussed results provide useful insight into the relative importance of local air–sea coupling and oceanic heat transports in the dynamics of SST variability in general, and of extreme SST anomalies in particular. Of course, other processes we neglected so far, such as variable mixed-layer depth, entrainment, and skewed atmospheric wind contributions, may potentially be responsible for negative skewness as well. A detailed analysis of these factors is part of our ongoing research.

Acknowledgements The author thanks Dr. Pierre Lermusiaux and three anonymous reviewers whose comments greatly improved the manuscript. This work was supported by NSF Grant ATM-0552047 "The impact of rapidly-varying heat fluxes on air-sea interaction and climate variability."

Appendix A

Stochastic differential equations in a nutshell

This appendix heuristically reviews a few basic ideas of SDEs used in this paper. More comprehensive treatments may be found in many textbooks (e.g., Gardiner 2004; Horsthemke and Léfèver 1984; Paul and Baschnagel 1999; Kloeden and Platen 1992; van Kampen 1981). Note that we here use the notion of the ensemble average $\langle ... \rangle$ to be mathematically consistent with textbooks on SDEs.

Consider the dynamics of a scalar x governed by the SDE

$$\frac{dx}{dt} = A(x) + B(x)\eta(t) \quad , \tag{24}$$

with the delta-correlated white-noise η: $\langle \eta(t) \rangle = 0$ and $\langle \eta(t)\eta(t') \rangle = \delta(t - t')$. How do we interpret this equation? The underlying problem is that, while the white-noise $\eta(t)$ is defined for every time t, it is not continuous and, therefore, not differentiable. This means that, mathematically speaking, the SDE Eq. 24 is not well defined. An alternative interpretation is called for. The solution is to interpret Eq. 24 in its integral form and, hence, expect the noise term to be

integrable. That is, we integrate Eq. 24 to obtain the integral equation

$$x(t) - x(t_0) = \int_{t_0}^{t} A[x(t')]dt' + \int_{t_0}^{t} B[x(t')]\eta(t')dt' \,. \quad (25)$$

which can be interpreted consistently. In addition, the discontinuous white-noise $\eta(t)$ is usually replaced with its continuous integral, the Wiener process $W(t)$, given by

$$W(t) = \int_0^t \eta(t')dt' \,, \quad (26)$$

or

$$\frac{dW(t)}{dt} = \eta(t) \,. \quad (27)$$

Informally, we may write $dW(t) = \eta(t)dt$, keeping in mind that this expression is rigorously defined only within an integral. From the definition of the Wiener process, its autocovariance can also be easily calculated: $\langle W(t)W(s)\rangle = \min(t, s)$. Using the definition of the Wiener process, the integral Eq. 25 becomes

$$x(t) - x(t_0) = \int_{t_0}^{t} A[x(t')]dt' + \int_{t_0}^{t} B[x(t')]dW(t') \,. \quad (28)$$

To summarize, the SDE Eq. 24 has to be interpreted in its mathematically consistent integral form. As a reminder of the integral definition, the SDE Eq. 24 is often written as

$$dx = A(x)dt + B(x)dW(t) \quad . \quad (29)$$

The rigorous integral expression of the increment of the Wiener process is

$$\begin{aligned} dW(t) &= W(t+dt) - W(t) \\ &= \int_0^{t+dt} \eta(t')dt' - \int_0^t \eta(t')dt' \\ &= \int_t^{t+dt} \eta(t')dt' \,. \end{aligned} \quad (30)$$

Using this expression, we can calculate the autocovariance $\langle dW(t)dW(s)\rangle$ as

$$\begin{aligned} \langle dW(t)dW(s)\rangle &= \left\langle \int_t^{t+dt} \int_s^{s+dt} \eta(t')\eta(s')\, dt'ds' \right\rangle \\ &= \int_t^{t+dt} \int_s^{s+dt} \langle \eta(t')\eta(s')\rangle \, dt'ds' \\ &= \int_t^{t+dt} \int_s^{s+dt} \delta(t'-s')\, dt'ds' \\ &= dt\, \delta_{t,s} \,, \end{aligned} \quad (31)$$

where $\delta_{t,s}$ denotes the Kronecker delta. For $t = s$, we obtain $\langle dW(t)^2\rangle = dt$, stating that the variance of white-noise (the so called Wiener process W) scales with the time increment dt. That is, the standard deviation of the Wiener process effectively scales with $\sqrt{dt}$ in contrast to the dt scaling of the determinist terms. It is this $\sqrt{dt}$ scaling of the stochastic terms that is responsible for so called *stochastic calculus*, in contrast to the dt scaling of the ordinary *deterministic calculus*. Again, the reader should consult one of many textbooks on SDEs for a more comprehensive treatment (e.g., Gardiner 2004; Horsthemke and Léfèver 1984; Paul and Baschnagel 1999; Kloeden and Platen 1992; van Kampen 1981).

Appendix B

More on third-order moments

The starting point is the Stratonovich SDE

$$\frac{\partial T_o'}{\partial t} = -\lambda_{\text{eff}} T_o' - \phi F' T_o' + F' + R' + \phi \overline{F'T_o'} \quad , \quad (32)$$

with the locally constant parameters $-\lambda_{\text{eff}}$ and $-\phi$ and the rapidly varying (approximately Gaussian white-noise) forcing terms F' and R' with amplitudes $\sigma_{F'}$ and $\sigma_{R'}$: $\overline{F'(t)F'(t')} = (\sigma_{F'})^2\delta(t-t')$ and $\overline{R'(t)R'(t')} = (\sigma_{R'})^2\delta(t-t')$. The effective drift $-\lambda_{\text{eff}}$ equals the sum of the deterministic drift and the noise-induced drift: $-\lambda_{\text{eff}} = -\lambda + (1/2)(\phi\sigma_{F'})^2$. Multiplying this equation with T_o', averaging, and assuming steady-state statistics yields

$$0 = -\lambda_{\text{eff}} \overline{T_o'^2} - \phi \overline{F'T_o'^2} + \overline{F'T_o'} + \overline{R'T_o'} \,, \quad (33)$$

or

$$\overline{T_o'^2} = \frac{1}{\lambda_{\text{eff}}} \left(-\phi \overline{F'T_o'^2} + \overline{F'T_o'} + \overline{R'T_o'} \right) \,. \quad (34)$$

The variance $\overline{T_o'^2}$ is known from Eq. 10. Therefore, we obtain

$$\frac{(\sigma_{F'}^2 + \sigma_{R'}^2)}{\left[2\lambda_{\text{eff}} - (\phi\sigma_{F'})^2\right]} = \frac{1}{\lambda_{\text{eff}}} \left(-\phi \overline{F'T_o'^2} + \overline{F'T_o'} + \overline{R'T_o'} \right) \,. \quad (35)$$

We see that for no multiplicative noise ($\phi = 0$) $\overline{F'T_o'} > 0$ and $\overline{R'T_o'} > 0$. From Eq. 35, we can also derive the following properties of $\overline{F'T_o'^2}$:

- Solving Eq. 35 for $\overline{F'T_o'^2}$ and letting $\phi \to 0$, we obtain the limit $\overline{F'T_o'^2} = 0$. In short, we write $\overline{F'T_o'^2} = 0$ for $\phi = 0$.
- From Eq. 35, we can also derive the sign of $\overline{F'T_o'^2}\phi$. We start by setting $\phi = 0$. The remaining terms are all positive. If we now allow for a small perturbation

ϕ away from zero, keeping $\overline{F'T'_o}$ and $\overline{R'T'_o}$ constant, we obtain $\overline{F'T'^2_o}\phi < 0$. Assuming continuous behavior of the moments, the general relations $\overline{F'T'^2_o}\phi \leq 0$ and $\overline{F'T'_o} > 0$ follow.

References

Auclair F, Marsaleix P, DeMey P (2003) Space-time structure and dynamics of the forecast error in a coastal circulation model of the Gulf of Lions. Dyn Atmos Ocean 36:309–346

Azzalini A (2005) R package 'sn': the skew-normal and skew-t distribution (version 0.33). Università di Padova, Italia

Azzalini A, Capitanio A (2003) Distributions generated by perturbations of symmetry with emphasis on a multivariate skew-*t* distribution. J R Stat Soc B 65:367–389

Blaauboer D, Komen GJ, Reiff J (1982) The behaviour of the sea surface temperature (SST) as a response to stochastic latent and sensible heat forcing. Tellus 34:17–28

Burgers G, Stephenson DB (1999) The "normality" of El Niño. Geophys Res Lett 26:1027–1030

Chelton DB, Schlax MG, Freilich MH, Milliff RF (2004) Satellite measurements reveal persistent small-scale features in ocean winds. Science 303:978–983

Clauset A, Shalizi CR, Newman MEJ (2009) Power-law distributions in empirical data. SIAM Rev 51:661–703

Coles S (2001) An introduction to statistical modeling of extreme values. Springer, New York, 208 pp

Frankignoul C (1985) Sea surface temperature anomalies, planetary waves, and air–sea feedback in the middle latitudes. Rev Geophys 23:357–390

Frankignoul C, Hasselmann K (1977) Stochastic climate models. Part II. Application to sea-surface temperature anomalies and thermocline variability. Tellus 29:289–305

Gardiner CW (2004) Handbook of stochastic methods for physics, chemistry and the natural science, 3rd edn. Springer, New York, 415 pp

Gille ST (2005) Statistical characterization of zonal and meridonal ocean wind stress. J Atmos Ocean Technol 22:1353–1372

Hasselmann K (1976) Stochastic climate models. Part I. Theory. Tellus 28:473–484

Horsthemke W, Léfèver R (1984) Noise-induced transitions: theory and applications in physics, chemistry, and biology. Springer, New York, 318 pp

Jayne SR, Marotzke J (2002) The oceanic eddy heat transport. J Phys Oceanogr 32:3328–3345

Jha AK, Winterstein SR (2000) Nonlinear random ocean waves: prediction and comparison with data. In: Proceedings of ETCE/OMAE2000 joint conference

Jones MC, Faddy MJ (2003) A skew extension of the *t*-distribution, with applications. J R Stat Soc B 65:159–174

Kloeden P, Platen E (1992) Numerical solution of stochastic differential equations. Springer, New York, 632 pp

Lermusiaux PFJ (2006) Uncertainty estimation and prediction for interdisciplinary ocean dynamics. J Comput Phys 217:176–199

Lermusiaux PFJ, Chiu C-S, Robinson AR (2002) Modeling uncertainties in the prediction of the acoustic wavefield in a shelfbreak environment. In: Proceedings of the 5th ICTCA, 21–25 May 2001

Lermusiaux PFJ, Chiu C-S, Gawarkiewicz GG, Abbot P, Robinson AR, Miller RN, Haley PJ, Leslie WG, Mjumdar SJ, Pang A, Lekien F (2006) Quantifying uncertainties in ocean predictions. Oceanography 19:92–105

Majda AJ, Franzke C, Khouider B (2008) An applied mathematics perspective on stochastic modelling for climate. Philos Trans R Soc 366:2429–2455

Monahan AH (2006a) The probability distributions of sea surface wind speeds part I: theory and SSM/I observations. J Climate 19:497–520

Monahan AH (2006b) The probability distributions of sea surface wind speeds part II: dataset intercomparison and seasonal variability. J Climate 19:521–534

Müller D (1987) Bispectra of sea-surface temperature anomalies. J Phys Oceanogr 17:26–36

Newman MEJ (2005) Power laws, Pareto distributions and Zipf's law. Contemp Phys 46:323–351

O'Neill LW, Chelton DB, Espensen SK (2003) Observations of SST-induced perturbations of the wind stress field of the Southern Ocean on seasonal time scales. J Climate 16:2340–2354

Paul W, Baschnagel J (1999) Stochastic processes: from physics to finance. Springer, New York, 231 pp

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical recipes in Fortran, the art of scientific computing. Cambridge University Press, Cambridge, 963 pp

R Development Core Team (2004) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0

Reynolds RW, Smith TM, Liu C, Chelton DB, Casey KS, Schlax MG (2007) Daily high-resolution blended analyses for sea surface temperature. J Climate 20:5473–5496

Sardeshmukh PD, Sura P (2009) Reconciling non-Gaussian climate statistics with linear dynamics. J Climate 22:1193–1207

Small RJ, Xie SP, Hafner J (2005) Satellite observations of mesoscale ocean features and copropagating atmospheric surface fields in the tropical belt. J Geophys Res 110:C02021

Sornette D (2006) Critical phenomena in natural sciences. Springer, New York, 528 pp

Stammer D (1998) On eddy characteristics, eddy transports, and mean flow properties. J Phys Oceanogr 28:727–739

Stephenson DB, Hannachi A, O'Neill A (2004) On the existence of multiple climate regimes. Q J R Meteorol Soc 130:583–605

Sura P, Barsugli JJ (2002) A note on estimating drift and diffusion parameters from timeseries. Phys Lett A 305:304–311

Sura P, Sardeshmukh PD (2008) A global view of non-Gaussian SST variability. J Phys Oceanogr 38:639–647

Sura P, Sardeshmukh PD (2009) A global view of air–sea thermal coupling and related non-Gaussian SST variability. Atmos Res 94:140–149. doi:10.1016/j.atmosres.2008.08.008

Sura P, Newman M, Penland C, Sardeshmukh PD (2005) Multiplicative noise and non-Gaussianity: a paradigm for atmospheric regimes? J Atmos Sci 62:1391–1409

Sura P, Newman M, Alexander MA (2006) Daily to decadal sea surface temperature variability driven by state-dependent stochastic heat fluxes. J Phys Oceanogr 36:1940–1958

Vallis GK (2006) Atmospheric and oceanic fluid dynamics. Fundamentals and large-scale circulation. Cambridge University Press, 745 pp

van Kampen NG (1981) Stochastic processes in physics and chemistry. Elsevier North-Holland, Amsterdam, 419 pp

Wilks DS (2006) Statistical methods in the atmospheric sciences, 2nd edn. Academic, London, 627 pp

Xie S-P (2004) Satellite observations of cool ocean-atmosphere interaction. Bull Am Meteorol Soc 85:195–208