Math Geosci
https://doi.org/10.1007/s11004-021-09990-6

Stochastic Modeling of Stratospheric Temperature

Mari Dahl Eggen[1] · **Kristina Rognlien Dahl**[1] · **Sven Peter Näsholm**[2,3] · **Steffen Mæland**[2]

Received: 26 May 2021 / Accepted: 1 December 2021

Abstract This study suggests a stochastic model for time series of daily zonal (circumpolar) mean stratospheric temperature at a given pressure level. It can be seen as an extension of previous studies which have developed stochastic models for surface temperatures. The proposed model is a combination of a deterministic seasonality function and a Lévy-driven multidimensional Ornstein–Uhlenbeck process, which is a mean-reverting stochastic process. More specifically, the deseasonalized temperature model is an order 4 continuous-time autoregressive model, meaning that the stratospheric temperature is modeled to be directly dependent on the temperature over four preceding days, while the model's longer-range memory stems from its recursive nature. This study is based on temperature data from the European Centre for Medium-Range Weather Forecasts ERA-Interim reanalysis model product. The residuals of the autoregressive model are well represented by normal inverse Gaussian-distributed random variables scaled with a time-dependent volatility function. A monthly variability in speed of mean reversion of stratospheric temperature is found, hence suggesting a generalization of the fourth-order continuous-time autoregressive model. A stochastic stratospheric temperature model, as proposed in this paper, can be used in geophysical analyses to improve the understanding of stratospheric dynamics. In particular, such characterizations of stratospheric temperature may be a step towards greater insight in modeling and prediction of large-scale middle atmospheric events, such as sudden stratospheric warming. Through stratosphere–troposphere coupling, the strato-

✉ Mari Dahl Eggen
marideg@math.uio.no

1 Department of Mathematics, University of Oslo, P.O. Box 1053, Blindern 0316 Oslo, Norway

2 NORSAR, Gunnar Randers vei 15, 2027 Kjeller, Norway

3 Department of Informatics, University of Oslo, P.O. Box 1080, Blindern 0316, Oslo, Norway

Published online: 06 January 2022

sphere is hence a source of extended tropospheric predictability at weekly to monthly timescales, which is of great importance in several societal and industry sectors.

Keywords Stratospheric temperature · Stochastic modeling · CARMA processes · Time series analysis

1 Introduction

A thorough understanding of surface weather dynamics is crucial in a wide range of industry and societal sectors. Whether planning marine operations, flights or farming, or managing energy assets, the weather is a key aspect to consider. However, because higher atmospheric layers can couple to levels closer to the surface, in order to understand weather, understanding the dynamics at higher altitudes of the atmosphere is important. The Earth's atmosphere has a layered structure, where each layer has layer-specific properties [see Baldwin et al. (2019) and the references therein for an historical overview]. Closest to the surface lays the troposphere, reaching up to around 15-km altitude. Above, up to around 50 km, lays the stratosphere, which is the atmospheric layer of interest in this paper. These two layers interact, and the dynamics in the stratosphere can couple to the troposphere to affect dynamics and predictability at the surface; see for example Butler et al. (2019) and Baldwin and Dunkerton (2001). Hence, better probing, modeling and understanding of stratospheric dynamics has the potential to enhance numerical surface weather prediction, in particular at weekly to monthly timescales.

In the current paper, a novel stochastic model for stratospheric temperature is proposed. The stochastic approach is similar to what has been applied in previous tropospheric temperature and wind dynamics modeling studies (e.g., Benth and Šaltytė Benth 2005, 2009; Benth et al. 2008, 2014; Benth and Taib 2013).

Temperature tends to revert back to its mean over time (Benth et al. 2008). This feature is reflected in what is called the speed of mean reversion, and is captured by autoregressive (AR) processes. AR processes are discrete time stochastic processes having a direct transformation relation with continuous-time autoregressive (CAR) processes (Brockwell 2004; Benth et al. 2008; Benth and Taib 2013). This transformation relation allows us to introduce a continuous-time mathematical model framework based on empirical derivations and analyses. Periodic behavior is modeled separately from the CAR process, and so is a long-term trend in the stratospheric temperature. The reason for the inclusion of a long-term trend in stochastic models for tropospheric temperature is that it is well known from climate research that there is a long-term warming of the troposphere (e.g., Jones and Wigley 1990; Sévellec and Drijfhout 2018; Intergovernmental Panel on Climate Change 2014). This effect is captured in the surface temperature modeling of Benth et al. (2008). Also, in Benth et al. (2008), cyclic functions are included through truncated Fourier series which can represent the periodic movements of tropospheric temperature. Similarly, several studies (e.g., Cnossen et al. 2015; Danilov and Konstantinova 2020; Steiner et al. 2020) have shown that there is a long-term decreasing trend in stratospheric temperature, and that there are several cyclic (seasonal) patterns (Fu et al. 2010; McCormack and Hood 1996).

The current empirical analysis and stochastic model study is performed on temperatures as represented in the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim atmospheric reanalysis model product (Berrisford et al. 2011; Dee et al. 2011). Full-year daily zonal mean stratospheric temperature data over 60°N and 10 hPa, from 1979 to 2018, are considered. Similarly as in Benth et al. (2008), a seasonality function in the form of a truncated Fourier series is fitted to the stratospheric temperature data to find deseasonalized temperature. Then, an AR model is fitted to the deseasonalized data and thereby subtracted to find the residuals. A search for seasonal heteroscedasticity (variability of variance over time) in the residuals is performed, and such heteroscedasticity is found. Based on this, a time-dependent volatility function, $\sigma(t)$, is defined. The $\sigma(t)$-scaled residuals are proven, with statistical significance, to be normal inverse Gaussian (NIG)-distributed random variables. Hence, the results of the data analysis suggest using a Lévy-driven CAR process with a time-varying volatility function to model stratospheric temperature. The residuals contain small memory effects, indicating that it might be reasonable to also consider a stochastic volatility function. This is beyond the scope of this paper.

An individual stability analysis of the speed of mean reversion over time is also performed, suggesting that the assumption about constant speed of mean reversion is not fulfilled. The result is twofold: the speed of mean reversion shows large variability from month to month, and it is varying with a seasonal pattern. Similarly, Zapranis and Alexandridis (2008) proved that the speed of mean reversion for tropospheric temperature is strongly time-dependent, obtaining a series of daily values of speed of mean reversion through neural networks. However, in contrast to the current paper, they did not observe seasonal patterns. In Benth and Khedher (2015), Ornstein–Uhlenbeck (OU) dynamics are generalized to allow for a stochastic speed of mean reversion, which can incorporate deterministic time dependence as well. However, Benth and Khedher (2015) consider the special case when the Lévy process is a Brownian motion. This is less general than the CAR process proposed for stratospheric temperature modeling in the current paper. Instead of using the aforementioned approaches to include time dependence in the speed of mean reversion of the CAR process, a simpler, approximate approach is suggested: a time-dependent step function with 12 levels is introduced, where the levels represent the months of the year. In this way, both the monthly variability and the seasonal behavior are adjusted for. The procedure of fitting a CAR process to the stratospheric temperature data is repeated for the extended CAR process with time dependence in speed of mean reversion. The inclusion of time dependence does not change the outcome of $\sigma(t)$-scaled residuals: these are still NIG-distributed random variables.

The structure of the paper is as follows: In Sect. 2, a mathematical framework of the stochastic model for stratospheric temperature is proposed. In Sect. 3, a non-Gaussian CAR process with constant speed of mean reversion and time-dependent volatility function is introduced and proposed to model stratospheric temperature. The methodology for fitting the model to the stratospheric temperature data is explained. Furthermore, it is shown that the most important features of stratospheric temperature dynamics are explainable through the proposed model. Then, in Sect. 4 a stability analysis of speed of mean reversion is performed, revealing that the proposed CAR

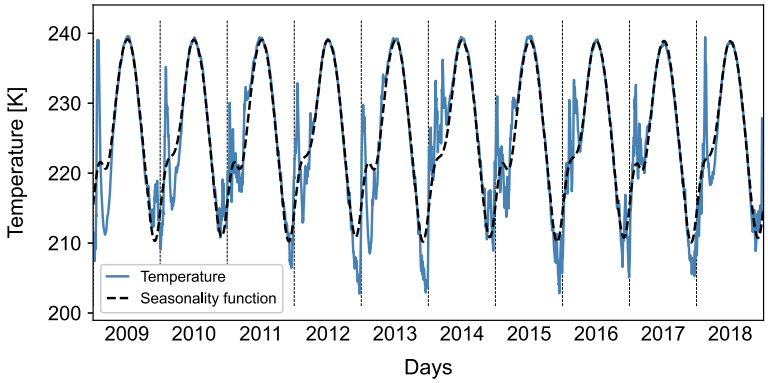


Fig. 1 Daily zonal mean stratospheric temperature, $S(t)$, over a region R (see Sect. 3.2) the last 10 years (1 January 2009 to 31 December 2018) with the fitted seasonality function $\Lambda(t)$. The vertical lines represent each of the 10 years

process should be generalized to include a time-dependent speed of mean reversion. Finally, conclusions and suggestions for future work are provided in Sect. 5.

2 The Structure of a Stratospheric Temperature Model

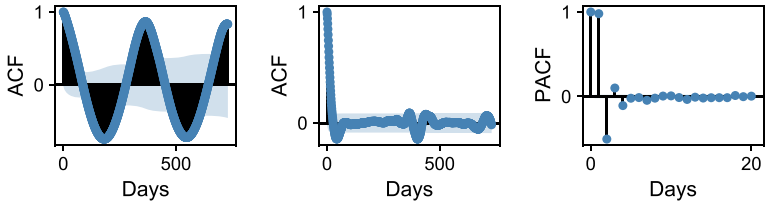
In this section, it is argued that stratospheric temperature exhibits an autoregressive behavior. This motivates the use of autoregressive models. For completeness, the definitions of discrete- and continuous-time autoregressive processes are recalled. It is also explained how these two kinds of processes are connected.

2.1 Autoregressive Behavior of Stratospheric Temperature Dynamics

Inspection of daily zonal mean stratospheric temperature data over 60°N and 10 hPa, from now on referred to as the stratospheric temperature data, $S(t)$, clearly indicates a seasonal pattern. This is illustrated in Fig. 1 which displays 10 years of stratospheric temperature data. The corresponding autocorrelation function (ACF), computed over the period from 1 January 1979 to 31 December 2018, with lags up to 730 days (2 years), is presented in Fig. 2a. The ACF pattern confirms a stratospheric temperature seasonal behavior.

Inspired by prior work in stochastic modeling of surface temperature and wind dynamics, in the context of financial weather contracts (e.g., Benth and Šaltytė Benth 2005, 2009, 2013; Benth et al. 2008; Benth and Taib 2013), a long-term seasonality and trend function is fit to the stratospheric temperature data; see Sect. 3.3. The long-term seasonality and trend function is from now on referred to as the seasonality function. Deseasonalized temperature is obtained by subtracting the fitted seasonality function from the original dataset.

Denote by $Y(t)$ the deseasonalized version of the stratospheric temperature, $S(t)$. Further, define $\Lambda(t) : [0, T] \to \mathbb{R}$ to be a bounded and continuously differentiable



(a) ACF of S(t) **(b)** ACF of Y(t) **(c)** PACF of Y(t)

Fig. 2 ACF of stratospheric temperature, $S(t)$, and ACF and PACF of deseasonalized stratospheric temperature, $Y(t)$ [see Eq. (1)]

(deterministic) seasonality function. Thus, stratospheric temperature is modeled as

$$S(t) = \Lambda(t) + Y(t). \tag{1}$$

Note that although the seasonality function $\Lambda(t)$ is deterministic, the stratospheric temperature, $S(t)$, and the deseasonalized temperature, $Y(t)$, are stochastic. Let Ω be a scenario space. Then, both $S(t)$ and $Y(t)$ depend on some scenario $\omega \in \Omega$, that is, $S(t) \triangleq S(t, \omega)$, $Y(t) \triangleq Y(t, \omega)$. For notational convenience, the scenario ω is suppressed from the notation for the remaining part of the paper. In Sect. 3.3, a review of possible seasonal effects is given prior to the explicit definition of the seasonality function $\Lambda(t)$. There, it will also be shown that a truncated Fourier series with a linear trend is an appropriate choice for the seasonality function [see Eq. (18)].

Studying the ACF and partial autocorrelation function (PACF) of the deseasonalized temperature data, $Y(t)$ (Fig. 2b, c), it is found that the deseasonalized temperature dynamics follows an AR process. For completeness, the definition of AR processes is given in the next section. Further, the PACF of the deseasonalized stratospheric temperature (Fig. 2c) indicates that an AR(4) model should be used to capture significant memory effects (see Levendis (2018) for an introduction to AR modeling and the interpretation of ACF and PACF plots). This means that the direct memory effects in stratospheric temperature last for 4 days in this model. However, due to the recursive properties of AR processes, the total memory effect is actually longer. As explained in Sect. 2.2, there is a transformation relation between (discrete-time) AR(p) and (continuous-time) CAR(p) processes. This means that, by removing the seasonal behavior in stratospheric temperature data, the resulting deseasonalized stratospheric temperature data can be modeled by a CAR process. It will be shown that $Y(t)$ can be approximated by a CAR(4) model.

2.2 Non-Gaussian CAR(p) Processes and Their Connection to AR(p) Processes

In this section, AR and CAR processes are defined. Motivated by the observed ACF and PACF for the deseasonalized stratospheric temperature data (see Sect. 2.1), these processes are natural components in a model for stratospheric temperature. They are also part of surface temperature models applied in energy markets contexts (e.g., Benth and Šaltytė Benth 2005, 2013, Ch. 4; Benth and Taib 2013; Benth et al. 2008, Ch. 10).

Suppose $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, P)$ is a complete filtered probability space. Let $\boldsymbol{X}(t) = \{\boldsymbol{X}(t)\}_{t\in\mathbb{R}^+}$ be a stochastic process in $\mathbb{R}^p$, $p \in \mathbb{N}$, defined by a multidimensional non-Gaussian OU process with time-dependent volatility. That is, $\boldsymbol{X}(t)$ is given by the solution of the stochastic differential equation (SDE)

$$\mathrm{d}\boldsymbol{X}(t) = A\boldsymbol{X}(t)\mathrm{d}t + \boldsymbol{e}_p\sigma(t-)\mathrm{d}L(t), \tag{2}$$

where $\boldsymbol{e}_p$ is the p-th unit vector in $\mathbb{R}^p$, $\sigma(t) : \mathbb{R}^+ \to \mathbb{R}^+$ is a càdlàg, $\mathscr{F}_t$-adapted function, $L(t) = \{L(t)\}_{t\in\mathbb{R}^+}$ is a Lévy process, and A is the $p \times p$ coefficient matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\alpha_p & -\alpha_{p-1} & -\alpha_{p-2} & \cdots & -\alpha_1 \end{bmatrix}, \tag{3}$$

where $\alpha_k, k = 1, \ldots, p$, are positive constants. By the multidimensional Itô formula, the solution of Eq. (2) is given by

$$\boldsymbol{X}(s) = \exp\left(A(s-t)\right)\boldsymbol{x} + \int_t^s \exp(A(s-u))\boldsymbol{e}_p\sigma(u-)\mathrm{d}L(u),$$

where $s \geq t \geq 0$ and $\boldsymbol{X}(t) \triangleq \boldsymbol{x} \in \mathbb{R}^p$. Let $p > j > q \in \mathbb{N}$. Then, the Lévy-driven stochastic process $Y(t) = \{Y(t)\}_{t\in\mathbb{R}^+}$ defined as $Y(t) \triangleq \boldsymbol{b}'\boldsymbol{X}(t)$, for a transposed vector $\boldsymbol{b} \in \mathbb{R}^{p\times 1}$ with elements satisfying $b_q \neq 0$ (sometimes, $b_q = 1$ is assumed) and $b_j = 0$, is called a (Lévy-driven) CARMA(p, q) process (e.g., Brockwell and Lindner 2015; Brockwell 2001, 2004, 2014). The simplified version of a CARMA(p, q) process $Y(t)$ where $\boldsymbol{b} = \boldsymbol{e}_1$, meaning $q = 1$ and

$$Y(t) = X_1(t), \tag{4}$$

is called a CAR(p) process. Note that p is the direct time lag dependence in $Y(t)$. As seen in Benth and Šaltytė Benth (2005) and Benth and Šaltytė Benth (2009), for example, the CAR(p) model framework is suitable for capturing surface temperature and wind evolution. Therefore, it is used in modeling of weather dynamics, for example in relation to financial weather contracts. In the current paper, it will be proved that the deseasonalized stratospheric temperature, $Y(t)$, can be modeled by a CAR(p) process as in Eq. (4).

Now, for a discrete-time framework version, consider an AR(p) process given by

$$X(t) = \beta_1 X(t-1) + \beta_2 X(t-2) + \cdots + \beta_p X(t-p) + e(t), \tag{5}$$

where $X(t) \in \mathbb{R}$ is the value of the AR process at times $t = 0, 1, \ldots$, and β_k, $k = 1, \ldots, p$, are constant coefficients and $e(t)$ are i.i.d. random error terms. The

dynamics of a Lévy-driven CARMA(p, q) process, see Eq. (2), can be expressed as

$$\mathrm{d}X_q(t) = \begin{cases} X_{q+1}(t)\mathrm{d}t & \text{if } q = 1, \ldots, p-1 \\ -\sum_{q=1}^{p} \alpha_{p-q+1} X_q(t)\mathrm{d}t + \sigma(t-)\mathrm{d}L(t) & \text{if } q = p. \end{cases} \tag{6}$$

By discretization of the expression in Eq. (6) (Benth et al. 2008; Benth and Taib 2013), a transformation relation between (discrete-time) AR(p) processes and the corresponding (continuous-time) CAR(p) processes is obtained, that is, a transformation relation between $X(t)$ in Eq. (5) and $Y(t)$ in Eq. (4). See Benth et al. (2008), Ch. 10, for a detailed derivation. Note that the connection between the AR and CAR processes is primarily useful because the continuous version, CAR, allows for deriving analytical results more easily, via stochastic analysis. For instance, Benth et al. (2008) uses CAR processes to model surface temperature, which is later used to price options. To the best of our knowledge, financial products based directly on stratospheric data are not commonly available. However, as the state of the stratosphere is connected to long-term surface weather forecasting, the CAR model may be of interest for pricing financial weather contracts with long-term maturity. Further developments may also aim for a stratospheric temperature model where a control is involved. This means a situation where one may affect the stratospheric temperature directly or indirectly, for example by carbon emissions.

Now consider the special case when $p = 4$, which will be proven to be well suited for modeling of stratospheric temperature, as assumed from observations in Sect. 2.1. The dynamics of the CAR(4) process, see Eq. (2), can be written as

$$\begin{aligned} \begin{bmatrix} \mathrm{d}X_1(t) \\ \mathrm{d}X_2(t) \\ \mathrm{d}X_3(t) \\ \mathrm{d}X_4(t) \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\alpha_4 & -\alpha_3 & -\alpha_2 & -\alpha_1 \end{bmatrix} \cdot \begin{bmatrix} X_1(t)\mathrm{d}t \\ X_2(t)\mathrm{d}t \\ X_3(t)\mathrm{d}t \\ X_4(t)\mathrm{d}t \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \sigma(t-)\mathrm{d}L(t) \end{bmatrix} \\ &= \begin{bmatrix} X_2(t)\mathrm{d}t \\ X_3(t)\mathrm{d}t \\ X_4(t)\mathrm{d}t \\ -\alpha_4 X_1(t)\mathrm{d}t - \alpha_3 X_2(t)\mathrm{d}t - \alpha_2 X_3(t)\mathrm{d}t - \alpha_1 X_4(t)\mathrm{d}t + \sigma(t-)\mathrm{d}L(t) \end{bmatrix}. \end{aligned}$$

Note that the dynamics have the form as described in Eq. (6). By the transformation relation between AR(p) processes and CAR(p) processes (Benth et al. 2008), it is found that the model coefficients of the CAR(4) process are given by

$$\begin{aligned} &\alpha_1 = 4 - \beta_1, \alpha_2 = -3\beta_1 - \beta_2 + 6, \\ &\alpha_3 = -3\beta_1 - 2\beta_2 - \beta_3 + 4, \alpha_4 = -\beta_1 - \beta_2 - \beta_3 - \beta_4 + 1. \end{aligned} \tag{7}$$

The matrix A, see Eq. (3), is referred to as the speed of mean reversion throughout the paper. This concept was introduced through half-life computations for Brownian motion-driven (one-dimensional) OU processes in Clewlow and Strickland (2000),

Sect. 2.4. That is, for some $s > t$ and a drift coefficient α, the formula

$$s - t = \frac{\ln(2)}{\alpha}, \tag{8}$$

gives the time until a shock $X_{(1)}(t)$ away from the process' long-term mean returns half-way back to this long-term mean (Benth and Taib 2013). For an OU process, the drift coefficient α is the only variable affecting the half-life. As large α gives shorter half-life, and smaller α gives longer half-life, α is referred to as speed of mean reversion. In the current paper, non-Gaussian CAR (CARMA) processes are considered rather than standard OU processes. The half-life formula for non-Gaussian CARMA processes is state-dependent (Benth and Taib 2013), meaning that the time s in Eq. (8) is a stopping time. Denote this stopping time by τ. The special case when the non-Gaussian CARMA process is a CAR process (the process considered in the remaining parts of this paper) gives a half-life formula of the form

$$\boldsymbol{e}_1' \left(\exp(A(\tau - t)) - \frac{1}{2} I \right) \boldsymbol{X}(t) = 0, \tag{9}$$

(Benth and Taib 2012). Solving this equation for τ analytically is difficult, and hence it is not clear how the coefficient matrix A affects the process' half-life. The coefficient matrix A will still be referred to as the speed of mean reversion, where each matrix element α_i is assumed to be a contribution to the speed of mean reversion.

A CAR(4) model driven by the multidimensional OU process in Eq. (2) assumes constant speed of mean reversion. In Sect. 4, it will be shown that this assumption is not valid for our dataset. The stratospheric temperature data indicate a seasonal varying pattern in speed of mean reversion from month to month. Based on this observation, an extended model framework is proposed; that is, a CAR(4) model driven by a multidimensional OU process with time-varying speed of mean reversion. The theorem below gives an explicit formula for the (unique) solution of the multidimensional OU SDE driven by a Lévy process with time-dependent speed of mean reversion.

Theorem 1 Let $\boldsymbol{X}(t)$ be given by the multidimensional OU process

$$\mathrm{d}\boldsymbol{X}(t) = A(t)\boldsymbol{X}(t)\mathrm{d}t + \boldsymbol{e}_4 \sigma(t-)\mathrm{d}L(t), \tag{10}$$

where $A(t)$ is the 4×4-matrix

$$A(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\alpha_4(t) & -\alpha_3(t) & -\alpha_2(t) & -\alpha_1(t) \end{bmatrix}. \tag{11}$$

Then,

$$\boldsymbol{X}(t) = \exp\left(\int_0^t A(s)\mathrm{d}s\right)\boldsymbol{x} + \exp\left(\int_0^t A(s)\mathrm{d}s\right)\int_0^t \exp\left(-\int_0^s A(u)\mathrm{d}u\right)\boldsymbol{e}_4\sigma(s-)\mathrm{d}L(s), \tag{12}$$

where $\boldsymbol{x} \triangleq \boldsymbol{X}(0)$.

Proof Rewrite the SDE in Eq. (10) by use of the Itô-Lévy decomposition, to find that

$$\begin{aligned}\mathrm{d}\boldsymbol{X}(t) =& A(t)\boldsymbol{X}(t)\mathrm{d}t + \boldsymbol{e}_4\sigma(t-)\left(a\mathrm{d}t + bdB(t) + \int_{\mathbb{R}} z\bar{\boldsymbol{N}}(\mathrm{d}t,\mathrm{d}z)\right)\\ =& \left(A(t)\boldsymbol{X}(t) + \boldsymbol{e}_4 a\sigma(t-)\right)\mathrm{d}t + \boldsymbol{e}_4 b\sigma(t-)dB(t) + \boldsymbol{e}_4\sigma(t-)\int_{\mathbb{R}} z\bar{\boldsymbol{N}}(\mathrm{d}t,\mathrm{d}z).\end{aligned}$$

By definition, $\boldsymbol{X}(t)$ is a multidimensional Itô-Lévy process. Apply the multidimensional Itô formula on $\mathrm{d}\boldsymbol{Y}(t) \triangleq d\left(\exp\left(-\int_0^t A(s)\mathrm{d}s\right)\boldsymbol{X}(t)\right)$. By defining $Y(t) \triangleq f(t, \boldsymbol{X}(t))$, it is found by the dominated convergence theorem and the fundamental theorem of calculus that

$$\begin{aligned}\frac{\partial f}{\partial t}(t,\boldsymbol{X}(t)) =& \frac{\partial}{\partial t}\sum_{k=1}^{\infty}\frac{1}{k!}\left(-\int_0^t A(s)\mathrm{d}s\right)^k \boldsymbol{X}(t)\\ =& -\sum_{k=1}^{\infty}\frac{1}{(k-1)!}\left(-\int_0^t A(s)\mathrm{d}s\right)^{k-1} A(t)\boldsymbol{X}(t)\\ =& -\exp\left(-\int_0^t A(s)\mathrm{d}s\right)A(t)\boldsymbol{X}(t).\end{aligned}$$

Furthermore, note that

$$\frac{\partial f}{\partial x_i}(t,\boldsymbol{X}(t)) = \exp\left(-\int_0^t A(s)\mathrm{d}s\right)\boldsymbol{e}_i \quad \text{and} \quad \frac{\partial^2 f}{\partial x_i \partial x_j}(t,\boldsymbol{X}(t)) = 0,$$

for all i and i, j, respectively. The remaining terms coming from the Itô formula are trivial. Thus, one finds that

$$\begin{aligned}\mathrm{d}\boldsymbol{Y}(t) =& -\exp\left(-\int_0^t A(s)\mathrm{d}s\right)A(t)\boldsymbol{X}(t)\mathrm{d}t\\ &+\exp\left(-\int_0^t A(s)\mathrm{d}s\right)(X_2(t)\boldsymbol{e}_1 + X_3(t)\boldsymbol{e}_2 + X_4(t)\boldsymbol{e}_3\\ &+(-\alpha_4(t)X_1(t) - \alpha_3(t)X_2(t) - \alpha_2(t)X_3(t) - \alpha_1(t)X_4(t) + a\sigma(t-))\boldsymbol{e}_4)\,\mathrm{d}t\\ &+\exp\left(-\int_0^t A(s)\mathrm{d}s\right)\boldsymbol{e}_4 b\sigma(t-)dB(t)\end{aligned}$$

$$
\begin{aligned}
&+\exp\left(-\int_0^t A(s)\mathrm{d}s\right)\int_{|z|<R}\{\boldsymbol{X}(t-)+\boldsymbol{e}_4\sigma(t-)z-\boldsymbol{X}(t-)-\boldsymbol{e}_4\sigma(t-)z\}\,\nu(\mathrm{d}z)\mathrm{d}t\\
&+\exp\left(-\int_0^t A(s)\mathrm{d}s\right)\int_{\mathbb{R}}\{\boldsymbol{X}(t-)+\boldsymbol{e}_4\sigma(t-)z-\boldsymbol{X}(t-)\}\,\bar{\boldsymbol{N}}(\mathrm{d}t,\mathrm{d}z)\\
&=\exp\left(-\int_0^t A(s)\mathrm{d}s\right)(A(t)\boldsymbol{X}(t)-A(t)\boldsymbol{X}(t)\\
&\quad+\boldsymbol{e}_4\sigma(t-)\left(a\mathrm{d}t+b\mathrm{d}B(t)+\int_{\mathbb{R}}z\bar{\boldsymbol{N}}(\mathrm{d}t,\mathrm{d}z)\right)\Bigg)\\
&=\exp\left(-\int_0^t A(s)\mathrm{d}s\right)\boldsymbol{e}_4\sigma(t-)\mathrm{d}L(t).
\end{aligned}
$$

Hence, from the definition of $Y(t)$, when $\boldsymbol{x}\triangleq\boldsymbol{X}(0)$,

$$
\begin{aligned}
\boldsymbol{X}(t)&=\exp\left(\int_0^t A(s)\mathrm{d}s\right)\boldsymbol{x}\\
&\quad+\exp\left(\int_0^t A(s)\mathrm{d}s\right)\int_0^t\exp\left(-\int_0^s A(u)\mathrm{d}u\right)\boldsymbol{e}_4\sigma(s-)\mathrm{d}L(s).
\end{aligned}
$$

□

3 Stochastic Modeling of Daily Zonal Mean Stratospheric Temperature

The aim of the following sections is to fit a CAR model to the daily zonal mean stratospheric temperature data obtained from the ECMWF ERA-Interim reanalysis product.

3.1 Methodology for Deriving and Fitting a Stochastic Model to Stratospheric Temperature Data

This section describes the data analysis applied in Sects. 3.3–3.5 to fit the model in Eq. (1) to ERA-Interim stratospheric temperature reanalysis data (described in Berrisford et al. (2011) and specified in Sect. 3.2). Applying this methodology shows that the model in Eq. (1) is suitable to model stratospheric temperature when $Y(t)$ is a non-Gaussian CAR(4) process.

Assume that a dataset of stratospheric temperatures indexed by time is given, and denote this by $\mathscr{S}$. A detailed description of the stratospheric temperature dataset used in this paper will be given in Sect. 3.2. The main steps of the data analysis of $\mathscr{S}$ are as follows:

1. Fit a deterministic continuous seasonality function $\Lambda(t)$ to $\mathscr{S}$. Subtract $\Lambda(t)$ from $\mathscr{S}$ to obtain a dataset of deseasonalized stratospheric temperatures, denoted $\mathscr{S}_d$.
2. Fit an AR(p) model to $\mathscr{S}_d$ with the choice of p based on the PACF of the dataset. Subtract the fitted AR(p) model from $\mathscr{S}_d$ to obtain a dataset of residuals, $\mathscr{E}$.

3. Compute the empirical expected values of the squared residuals each day over the year (assumed to be 365 days) to construct an approximation of the time-varying volatility function, $\sigma(t)$.
4. Divide $\mathscr{E}$ by $\sigma(t)$ to obtain a dataset of $\sigma(t)$-scaled residuals, denoted as $\hat{\mathscr{E}}$. Find the probability distribution of the elements in $\hat{\mathscr{E}}$ (by statistical analysis).

As the goal of this work is to obtain a continuous-time stochastic model for stratospheric temperature, notation corresponding to continuous functions will be used in the more detailed explanation that follows.

Assume that the stratospheric temperature, $S(t)$, is given by Eq. (1), $Y(t)$ being a CAR(p) process as in Eq. (4). The lag $p = 4$ is chosen based on observations in Sect. 2.1, meaning that the direct memory effect reaches over 4 days. Then, by the transformation relation between CAR(4) and AR(4) processes in Eq. (7), the deseasonalized temperature is given by

$$X(t) = S(t) - \Lambda(t) = \sum_{k=1}^{4} \beta_k X(t-k) + e(t). \tag{13}$$

The seasonality function $\Lambda(t)$ is fit by least squares to simulate the seasonal behavior of the stratospheric temperature data in $\mathscr{S}$, and then subtracted from the stratospheric temperature data to find the discrete version of deseasonalized stratospheric temperature. The deseasonalized temperature is given by the AR(4) process $X(t)$ in Eq. (13), with random error terms (residuals) $e(t)$. Therefore, by use of least squares, an AR(4) model is fit to the deseasonalized stratospheric temperature data in $\mathscr{S}_d$, and then subtracted to find the residuals dataset $\mathscr{E}$. Mathematically, the residuals are given by

$$e(t) = X(t) - \sum_{k=1}^{4} \beta_k X(t-k). \tag{14}$$

In Sect. 3.5, yearly heteroscedasticity is observed in the squared residuals. This means that the daily variance values (over the year) of the dataset $\mathscr{E}$ are time-dependent. Therefore, a time-varying volatility function $\sigma(t)$ is approximated and divided on the residuals in $\mathscr{E}$ to obtain the $\sigma(t)$-scaled residuals

$$\epsilon(t) = \frac{e(t)}{\sigma(t)}. \tag{15}$$

That is, $\epsilon(t)$ represents the data in $\hat{\mathscr{E}}$. Recall from Sect. 2.2 [Eq. (5)], that $\epsilon(t)$ are i.i.d. random variables. The mean value of residuals each day d during the year, $d \in [1, 365]$ (see Sect. 3.2), is assumed to be constant. Therefore, the variance each day during the year is given by

$$\mathrm{Var}(e_d(t)) = \left(E[e_d^2(t)] - E[e_d(t)]^2 \right), \tag{16}$$

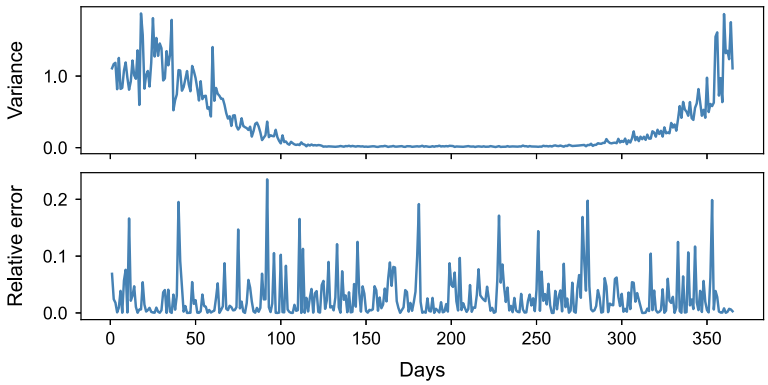


Fig. 3 Variance each day of the year computed with the approximation in Eq. (17), and the corresponding relative error when compared with the definition of variance [Eq. (16)]. The mean absolute percentage error is 2.8%

where $E[e_d(t)]$ represents the empirical mean value of residuals at day d. The magnitude of $E[e_d(t)]^2$ is insignificant compared to $E[e_d^2(t)]$; see Fig. 3 for an illustration of this. Hence, the approximation

$$\mathrm{Var}(e_d(t)) \simeq E[e_d^2(t)] \tag{17}$$

is used to fit an appropriate time-varying variance function, $V(t)$, for $t \in [1, 365]$. See Sect. 3.5 for a thorough explanation of how to compute $E[e_d^2(t)]$ empirically. When $E[e_d^2(t)]$ is computed for each d, the function $V(t)$ is fit to the values by use of three heavily truncated Fourier series (by least squares), which are connected by two sigmoid functions. The time-varying volatility function is finally computed as $\sigma(t) = \sqrt{V(t)}$. When $\hat{\mathcal{E}}$ is obtained, an appropriate probability density function (pdf) describing the distribution of the $\sigma(t)$-scaled residuals has to be found. Finally, the last step is to introduce a stochastic process which is able to replicate the behavior of the particular pdf. The $\sigma(t)$-scaled residuals function $\epsilon(t)$ is represented by this stochastic process in the CAR model.

From this analysis, an appropriate driving stochastic process for the CAR(4) model is obtained. However, at this point (due to doing time series analysis) the model is given by Eq. (13), and is thus a discrete model. The CAR(4) model, which is given in Eq. (4), is found by applying the transformation relation in Eq. (7). A detailed description of this data analysis methodology applied to the ERA-Interim stratospheric temperature reanalysis data (see Sect. 3.2) is given in Sects. 3.3–3.5.

3.2 The Zonal Mean Stratospheric Temperature Dataset

The aim of this section is to describe the stratospheric temperature dataset analyzed in the remainder of this paper.

Define a spherical coordinate system such that $(r_{\mathrm{tot}}, \theta, \varphi)$ represents a point in the atmosphere. Let $r_{\mathrm{tot}} = r_E + r$ represent the altitude from the center of the Earth,

where r_E is the radius of the Earth, and r is the distance from Earth's surface to the atmospheric point of interest. Further, θ represents the longitude and φ the latitude. With the presented notation, the region of interest in this paper can be defined as

$$R \triangleq \left\{(r_{\text{tot}}, \theta, \varphi) | r = 10\,\text{hPa} \simeq 30\,\text{km}, \theta \in [-180^\circ\text{E}, 180^\circ\text{E}), \varphi = 60^\circ\text{N}\right\}.$$

The region R is an area bounded by a circumpolar line in the extratropical stratosphere. The pressure level 10 hPa corresponds roughly to 30-km altitude. Stratospheric dynamics in this region are highly variable, and they depend on the state of the stratospheric polar vortex.

Enhanced probing and representation of the stratosphere in atmospheric models and numerical weather prediction systems has potential to enhance surface weather predictions on weekly to monthly timescales (e.g., Pedatella et al. 2018; Karpechko et al. 2016; Butler et al. 2019). Maybe the most striking example of stratospheric influence on the surface is the extreme event of sudden stratospheric warming (SSW), where an abrupt disruption in the stratospheric winter circulation occurs, accompanied by a stratospheric temperature increase of several tens of degrees. SSWs are detectable in the region R. Through stratosphere–troposphere coupling, the effects of SSWs can extend to the troposphere, with increased probability of shifts in the jet stream and storm tracks, further affecting the expected precipitation and surface temperatures. This phenomenon can, for example, be manifested as harsher winter weather regimes on continental North America and Eurasia (Baldwin et al. 2021). This may have impact on several sectors in society and industry.

With the purpose of deriving a stochastic stratospheric temperature model, a dataset $\mathscr{D}$ from the ERA-Interim reanalysis model product is retrieved from ECMWF (Berrisford et al. 2011), such that each temporal data point $d_i \in \mathscr{D}$ represents the spatial mean stratospheric temperature over R. As the spatial mean is taken over the full circumpolar interval, this is denoted as the zonal mean. The zonal mean properties of the stratosphere at the 10-hPa pressure level are commonly considered in stratospheric diagnostics and when studying stratospheric events such as SSWs and beyond (Butler et al. 2015). The subscript i represents a measurement every 6 h from midnight, and the zonal mean is taken over θ at a 0.5° spacing. That is, $\mathscr{D}$ contains four zonal mean temperature measurements every day within the interval $T \in$ [1 January 1979, 31 December 2018]. For computational convenience, all data from 29 February each leap year are excluded from $\mathscr{D}$, such that the length of each year is constant. All stated specifications of $\mathscr{D}$ are collected in Table 1. Further, define the dataset

$$\mathscr{S} \triangleq \left\{S_k : S_k = E[d|D_k]\right\},$$

where D_k are subsets of $\mathscr{D}$ containing four data points every given day k in the time interval T (except days 29 February), and $E[d|D_k]$ is the empirical mean. That is, $\mathscr{S}$ contains daily zonal mean stratospheric temperatures over the region R for days in the time interval T. This is the time series analyzed in the remainder of this paper. As 29 February is excluded each leap year, the total number of data points in $\mathscr{S}$ is 14, 600.

Table 1 Specifications of the stratospheric temperature dataset $\mathscr{D}$

Date	Grid	Pressure level	Time	Area	Unit
1 January 1979 to 31 December 2018	0.5°	10 hPa	00:00, 06:00, 12:00, 18:00	60°N and [− 180°E, 180°E)	Kelvin

A plot of $\mathscr{S}$ for the last 10 years (from 1 January 2009 to 31 December 2018) with a fitted seasonality function was shown in Fig. 1.

3.3 Fitting a Seasonality Function to Stratospheric Temperature Data

In this section, seasonality in stratospheric temperature data is analyzed. Seasonality in this setting means (deterministic) periodic repetitive patterns of temperature dynamics over time. A deterministic seasonality function will be fit to the dataset $\mathscr{S}$, with the aim of further analyzing deseasonalized temperature data where these periodic repetitive patterns are removed. Although the stratosphere is typically characterized by variations on longer timescales than the troposphere, there are oscillation, or atmospheric tide, patterns present at these altitudes as well. In addition to the directly forced cycles causing seasonal effects, the phenomenon of quasi-biennial oscillations (QBO), for example, is a nearly periodic phenomenon in the stratosphere. The QBO period is variable, but averages to about 28 months. This is a phenomenon occurring in the equatorial stratosphere. Still, the QBO can affect stratospheric conditions from pole to pole, and even has effects on the breaking of wintertime polar vortices, leading to SSWs (Vallis 2017; Baldwin et al. 2001).

Seasonal effects complicate stochastic modeling because they cause non-stationarity. In the current study, daily zonal mean stratospheric temperatures over 40 years are considered, meaning that the periodic phenomena of interest are the yearly cycle and the QBO. Non-stationarity can also result from long-term effects of greenhouse gases and ozone, anthropogenic forcings that cause a stratospheric cooling trend (Cnossen et al. 2015; Danilov and Konstantinova 2020; Steiner et al. 2020). The first step in deriving a stochastic stratospheric temperature model is to fit a seasonality function to the data, and then to subtract this to remove the non-stationary effects.

As the yearly cycle is the most pronounced phenomenon (Figs. 1 and 2a), a Fourier series with a period of 365 days is chosen as seasonality function. Further, the long-term decreasing trend in stratospheric temperature is approximately linear, meaning that a linear function should be present in the seasonality function as well. Based on these considerations, a seasonality function $\Lambda(t)$ is defined as in Eq. (18).

In the following, the continuous version of daily zonal mean stratospheric temperature data $S_i \in \mathscr{S}$ (as described in Sect. 3.2) is denoted by $S(t)$, where $t \in \mathbb{R}^+$. Let $S(t)$ be given by the stochastic model in Eq. (1) with seasonality function

$$\Lambda(t) = c_0 + c_1 t + \sum_{k=1}^{n} \left(c_{2k} \cos(k\pi t/365) + c_{2k+1} \sin(k\pi t/365)\right), \qquad (18)$$

Table 2 Seasonality function, $\Lambda(t)$, parameters [see Eq. (18)] for daily zonal mean stratospheric temperature at 60°N and 10 hPa between 1 January 1979 and 31 December 2018

c_0	c_2	c_4	c_6	c_8	c_{10}	c_{12}	c_{14}	c_{16}	c_{18}	c_{20}
226.15	− 0.05	− 12.09	0.23	1.88	0.33	0.16	0.13	− 0.09	− 0.15	− 0.01
c_1	c_3	c_5	c_7	c_9	c_{11}	c_{13}	c_{15}	c_{17}	c_{19}	c_{21}
− 0.000072	− 0.11	1.63	− 0.23	2.81	− 0.04	1.54	0.14	0.45	0.05	0.11

where $c_0, c_1, c_2, \ldots, c_{2n+1}$ are constants. The choice of $\Lambda(t)$ is made based on the discussion above, where c_1 captures the slope of the long-term cooling of the stratosphere (corresponding to global warming of the troposphere), and where the constant term c_0 represents the average level at the beginning of the time series $\mathscr{S}$. The constants $c_2, \ldots, c_{2k+1}$ describe the yearly cycle as weights in the truncated Fourier series.

The seasonality function $\Lambda(t)$ is fit to the time series in $\mathscr{S}$ with $n = 10$ (using least squares), with resulting parameters given in Table 2. Figure 1 displays the fitted seasonality function $\Lambda(t)$ together with the last 10 years of the times series in $\mathscr{S}$. The value of c_0 in Table 2 indicates that the daily zonal mean stratospheric temperature over the region R (see Sect. 3.2) was approximately 226.15 K (− 47.00 °C) in 1979. The negative value of c_1 confirms the long-term cooling effect of the stratosphere. The c_1 value found corresponds to a daily zonal mean stratospheric temperature decrease of approximately 1.05 K (equivalent to a change of 1.05°C) over the last 40 years at 60°N and 10 hPa. This is consistent with Steiner et al. (2020), estimating an overall cooling of the stratosphere of about 1–3 K over the same time span.

3.4 Fitting an AR Model to Deseasonalized Stratospheric Temperature Data

Having deseasonalized the stratospheric temperature dataset $\mathscr{S}$, the next step is to fit an AR model to the deseasonalized dataset $\mathscr{S}_d$. Temperature tends to have a mean-reverting property over time, a property that can be modeled by an AR(p) process (Benth et al. 2008). Based on the discussion in Sect. 2.1, suppose that the deseasonalized stratospheric temperature $Y(t) = S(t) - \Lambda(t)$ can be modeled by an AR(p) process as represented in Eq. (5), where the random error terms $e(t)$ represent the model residuals. The empirical ACF and PACF of the deseasonalized stratospheric temperature data illustrated in Fig. 2b, c confirm that it is appropriate to model $Y(t)$ by an AR(p) process, and indicate that $p = 4$ is needed to explain the time series evolution (see Levendis 2018). An AR(4) model is fit to the deseasonalized stratospheric temperature data in $\mathscr{S}_d$ by use of least squares. The resulting AR(4) model parameters are presented in Table 3. By use of Eq. (7), the corresponding CAR(4) process is calculated, and the resulting model parameters for this continuous model are presented in Table 3 as well. Based on the reasoning in Benth et al. (2008), preservation of stationarity of the CAR(4) model depends on the properties of the time-dependent volatility function $\sigma(t)$. However, as long as all eigenvalues of the matrix A have a negative real part, it is ensured that the modeled temperature on average will coincide

Table 3 AR(4) model parameters, and parameters of its continuous counterpart, when fitted to daily zonal mean stratospheric temperature over 60°N and 10 hPa in the period 1 January 1979 to 31 December 2018

AR(4) parameters			
β_1	β_2	β_3	β_4
1.55	− 0.75	0.28	− 0.11
CAR(4) parameters			
α_1	α_2	α_3	α_4
2.45	2.10	0.56	0.03

with the seasonality function $\Lambda(t)$ when time approaches infinity. This is because, as will be shown in Sect. 3.5, the Lévy-driven CAR(4) model generates NIG-distributed random variables with mean zero, a property which is preserved for the model in the long run when the eigenvalues have a negative real part. The eigenvalue equation

$$\lambda^4 + \alpha_1\lambda^3 + \alpha_2\lambda^2 + \alpha_3\lambda + \alpha_4, \tag{19}$$

has roots $\lambda_{1,2} = -1.01 \pm 0.43i$, $\lambda_3 = -0.36$ and $\lambda_4 = -0.07$, and the stationarity condition is therefore satisfied.

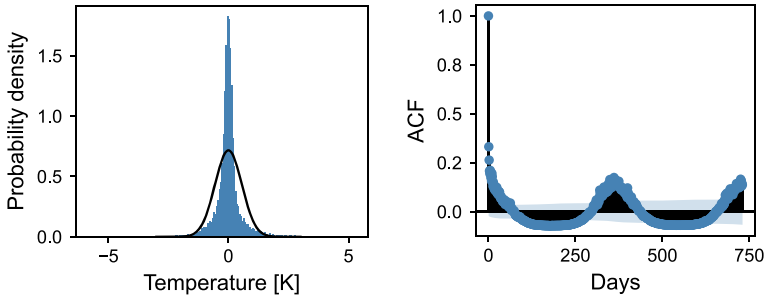
3.5 Analyzing the Residuals

In this section, the residuals (random error terms) in the dataset $\mathscr{E}$ are analyzed to determine the appropriate stochastic driving process of the CAR(4) model for stratospheric temperature.

In computing the parameters of the CAR(4) model in Sect. 3.4, the deterministic mean-reverting property of the stratospheric temperature is found. A suitable stochastic driving process for the model residuals, corresponding to the random error terms in Eq. (5), still remains to be found. As derived in Sect. 3.1, the model residuals are given by

$$e(t) = X(t) - \sum_{k=1}^{4} \beta_k X(t-k). \tag{20}$$

The approach to find a suitable stochastic driving process is therefore to empirically determine the stratospheric temperature model residual distribution in $\mathscr{E}$. In Levendis (2018), there is a statement that residuals of AR(p) models approach white noise for larger p. Therefore, it is reasonable as a first guess to assume that $e(t)$ is distributed as i.i.d. $N(0, 1)$. A normal fit is performed on $\mathscr{E}$; however, by Fig. 4a, it is clear that the data are not normally distributed. Further, Fig. 4b indicates a seasonally varying empirical ACF of squared residuals. Since the distributional mean value is close to zero, this is a sign of seasonal heteroscedasticity in the distributional variance (see Sect. 3.1). A similar seasonal pattern in the empirical ACF of squared residuals was observed by Benth et al. (2008) in daily average surface temperature data in Sweden.



(a) Distribution of e(t) **(b)** ACF of e^2(t)

Fig. 4 The distribution of stratospheric temperature model residuals $e(t)$ [Eq. (20)] with a fitted normal distribution, and ACF of the squared model residuals $e^2(t)$

Still assuming the stratospheric temperature model residuals to be normally distributed random variables, but with a time-varying variance rather than the constant 1, $e(t)$ is rewritten as

$$e(t) = \sigma(t)\epsilon(t). \tag{21}$$

Here, $\epsilon(t)$ are distributed as i.i.d. $N(0, 1)$, and $\sigma(t)$ is a yearly (see Fig. 4b) time-varying deterministic function. To adjust for the heteroscedasticity, the volatility function $\sigma(t)$ must be defined explicitly. To do so, the same approach as in Benth et al. (2008) is used: Daily residuals over 40 years are organized into 365 groups, one group for each day of the year. This means that all observations on 1 January are collected into group 1, all observations on 2 January into group 2, and so on until all days of all years are grouped together. Recall that observations on 29 February were removed each leap year, such that each year contains 365 data points. By computing the empirical mean of the squared residuals in each group, an estimate of the expected squared residual each day of the year is found, corresponding to an estimate of the daily variance, as explained in Sect. 3.1 [Eq. (17)]. The resulting 365 estimates of daily variance yields an estimate of the time-varying variance function, $V(t)$, over the year. This is illustrated in Fig. 5. The yearly heteroscedasticity is clearly visible.

Recall that, by definition, the volatility function $\sigma(t)$ is the square root of the time-varying variance function. With the aim of obtaining $\sigma(t)$, an analytic function is fit to the empirically computed expected value of squared residuals to find a proper function $V(t)$. Figure 5 illustrates that the volatility in the stratospheric temperature variance is much higher in winter than in summer, as seen in Haynes (2005). This, as well as the shape of the estimated daily variance (expected squared residuals) over the year, makes function estimation with Fourier series more challenging than simply fitting a single Fourier series. To properly fit a function $V(t)$, the year is split into three parts. Each part represents the winter/spring season, summer season, and autumn/winter season, respectively. A local variance test is performed to find appropriate seasonal endpoints. That is, the summer season variance is low and stable compared to the two other seasons, and so the summer season endpoints are set where the local variance hits a given limit, δ. Define the local variance as

$$v(n) \triangleq \frac{1}{2n-1} \sum_{i=0}^{2n-1} (x_i - \mu)^2,$$

where n represents the degree of locality. The number of elements included in the sum is odd, $2n-1$, such that the local variance for each element is based on a symmetric number of neighbors on each side. The test is performed as follows: First, the limit δ is defined such that mid-summer local variances do not exceed δ. Second, with estimated expected squared residuals given as $[V_1, V_2, \ldots, V_{364}, V_{365}]$, the local variance $v(n)_k$ is computed for each point V_k in $[V_n, V_{n+1}, \ldots, V_{364-n}, V_{365-n}]$ (each endpoint is cut with $n-1$ elements for computability). Third, an array $K = [k_i \in \{n, n+1, \ldots, 365-n\} : v(n)_{k_i} < \delta]$ is constructed (sequentially in time), such that the index (that is, day) of elements with satisfactory small local variance is known. Fourth, based on the array K, a collection $\mathscr{K} = \{(k_i - k_{i-1}, k_i) \in \mathbb{N} \times \{n, n+1, \ldots, 365-n\} : \forall\, k_i \in K\}$ is constructed for stability purposes. Finally, all pairs in $\mathscr{K}$ where $k_i - k_{i-1} > 1$ (day) are printed such that stability of the condition $v(n)_k < \delta$ can be evaluated manually.

The analysis is performed with $n = 5$ and $\delta = 0.0002$, and gives cutoffs at days 115 and 288, corresponding to 25 April and 15 October, respectively. For simplicity, the cutoffs are set at the following whole month. That is, the three seasons winter/spring, summer, and autumn/winter are defined to be in the intervals [1 January, 30 April], [1 May, 31 October], and [1 November, 31 December], respectively. A function is fit to the estimate of $V(t)$ for each of the three seasons by use of the truncated Fourier series

$$w_f(t) = d_0 + \sum_{k=1}^{2} (d_{2k-1} \cos(fk\pi t/365) + d_{2k} \sin(fk\pi t/365)), \qquad (22)$$

where $d_0, \ldots, d_4$ are constants and f is a given parameter adjusting the series frequency. By manual inspection, the function $w_f(t)$ for each of the three seasons is chosen as $w_{0.44(1)}(t)$, $w_{2.0}(t)$, and $w_{0.44(2)}(t)$, respectively, and Table 4 displays the fitted parameters.

The transitions from winter/spring to summer and from summer to autumn/winter should be smooth in order to obtain a smooth yearly time-varying volatility function $\sigma(t) : [1, 365] \to \mathbb{R}$. This is achieved by connecting the three functions $w_{0.44(1)}(t)$, $w_{2.0}(t)$, and $w_{0.44(2)}(t)$ with two sigmoid functions as in Kitchin (2013). The sigmoid function $\omega(x)$ and the connective function $\xi(x)$ are given by

$$\omega(x) = \frac{1}{1 + \exp\left(-(\frac{x-a}{b})\right)} \quad \text{and} \quad \xi(x) = \big(1 - \omega(x)\big) f_1(x) + \omega(x) f_2(x),$$

where a and b are shift and scaling constants, respectively, and $f_1(x)$ and $f_2(x)$ are two functions that are to be connected. By connecting the functions $w_{0.44(1)}(t)$ and $w_{2.0}(t)$ with $a = 120$ and $b = 2$, and connecting the functions $w_{2.0}(t)$ and $w_{0.44(2)}(t)$ with $a = 304$ and $b = 5$, a smooth function $\sigma(t)$ is found. The resulting volatility function is illustrated as $\sigma^2(t)$ (or $V(t)$) in Fig. 5, together with the estimated daily variances during the year.

Table 4 Parameters of Fourier series, $w_f(t)$ [see Eq. (22)], fitted to each of the three seasons winter/spring, summer, and autumn/winter

Winter/spring ($f = 0.44$)				
d_0	d_1	d_2	d_3	d_4
− 507.58	633.29	269.90	− 124.82	− 130.75
Summer ($f = 2.0$)				
d_0	d_1	d_2	d_3	d_4
0.092	0.107	0.023	0.034	0.015
Autumn/winter ($f = 0.44$)				
d_0	d_1	d_2	d_3	d_4
13.36	− 263.03	86.00	91.21	102.50

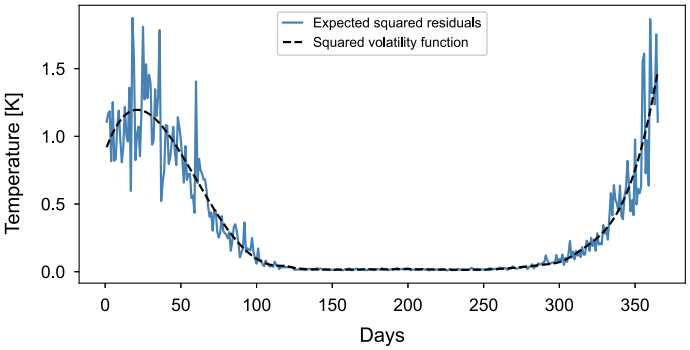
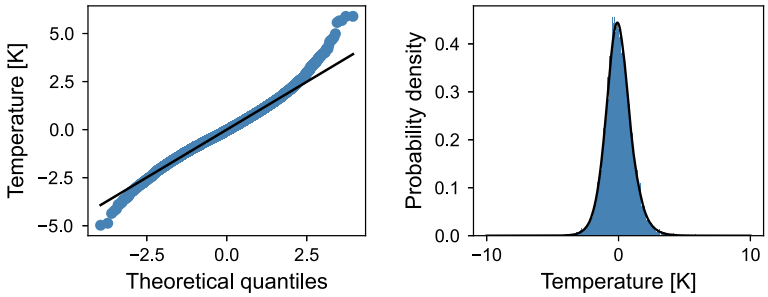


Fig. 5 Estimation of expected squared residuals (estimated variance) each day of the year illustrated together with a fitted function

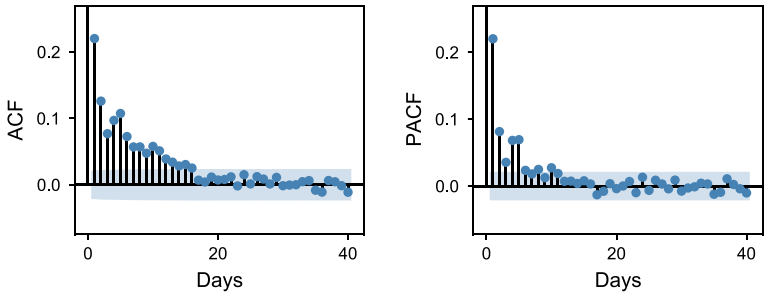
With an explicit expression for the volatility function $\sigma(t)$, the $\sigma(t)$-scaled residuals $\epsilon(t)$ [as described in Eq. (21)] can be studied. The distribution of the $\sigma(t)$-scaled stratospheric temperature model residuals in $\hat{\mathcal{E}}$ is compared to the normal distribution with a Q–Q plot in Fig. 6a. The above hypothesis about $\epsilon(t)$ being i.i.d. $N(0, 1)$ random variables does not hold, as the Q–Q plot illustrates heavy tails and a slightly skewed distribution. Also, the Kolmogorov–Smirnov (KS) test with statistic 0.027 and p-value $1.57 \cdot 10^{-9}$ gives significance for rejecting the hypothesis about $\epsilon(t)$ being standard normal. A fit with the NIG distribution is further performed, and the resulting pdf is illustrated in Fig. 6b. The KS test with statistic 0.0064 and p-value 0.57 does not reject the null hypothesis that $\epsilon(t)$ represents NIG-distributed random variables. Note that the KS test is meant to provide indicative results, rather than conclusive results from a carefully planned statistical experiment.

The result of $\epsilon(t)$ representing NIG-distributed random variables supports the hypothesis of using a Lévy process as the driving process for the stratospheric temperature model, as proposed in Sect. 2.2. However, as shown in Fig. 7a, b, the squared $\sigma(t)$-scaled residuals are partially autocorrelated in approximately five lags, meaning increments of $\epsilon(t)$ fail to be independently distributed. These memory effects indicate the use of a stochastic volatility function as described in Benth and Šaltytė Benth



(a) Normal QQ-plot of $\varepsilon(t)$ **(b)** NIG fit of $\varepsilon(t)$

Fig. 6 Fitted distributions to the $\sigma(t)$-scaled stratospheric temperature model residuals $\epsilon(t)$; see Eq. (21)



(a) ACF of $\varepsilon^2(t)$ **(b)** PACF of $\varepsilon^2(t)$

Fig. 7 ACF and PACF of the squared $\sigma(t)$-scaled stratospheric temperature model residuals $\epsilon^2(t)$; see Eq. (21)

(2009), rather than a deterministic yearly time-varying volatility function. To generalize the proposed model in Sect. 2.2, a possibility would be to model $e(t)$ as a normal variance-mean mixture with an inverse Gaussian stochastic volatility, as this process is approximately NIG-distributed (Barndorff-Nielsen and Shephard 2001; Benth and Taib 2013, Sect. 4). However, as the memory effects are rather small (except in the first lag), it could be appropriate to assume that there are no significant memory effects in the variance. A possibility is therefore to assume a deterministic volatility function, where the driving process for the stratospheric temperature model is a NIG Lévy process (e.g., Barndorff-Nielsen 1997a, b; Barndorff-Nielsen and Shephard 2001). Further study of this aspect is beyond the scope of the current paper and is left as a topic for further research.

4 Analyzing the Speed of Mean Reversion

In this section, it is shown that the assumption of constant speed of mean reversion for stratospheric temperature is erroneous. A generalization of the proposed stratospheric temperature model dynamics in Eq. (2), correcting this erroneous assumption, is pre-

sented. More specifically, a special case of the dynamics in Eq. (10) is proposed as a replacement to drive the stratospheric temperature model in Eq. (4).

4.1 Methodology for Analyzing Speed of Mean Reversion

In the previous sections, it was shown that the deseasonalized stratospheric temperature, $Y(t)$, follows a mean-reverting stochastic process. That is, deseasonalized stratospheric temperature dynamics is given by the OU process in Eq. (2). The matrix A holds parameters of speed of mean reversion, meaning that the rate at which the stratospheric temperature reverts back to its long-term mean is given by the elements of A; see Sect. 2.1 [Eqs. (8) and (9)]. In the previous sections the speed of mean reversion was assumed to be constant, and thus independent of time. To check the validity of this assumption, a similar stability analysis of speed of mean reversion as in Benth and Šaltytė Benth (2005) will be performed in the following. The stability analysis exploits the transformation relation between CAR and AR models (see Sect. 2.2), meaning that it is applied on computed AR parameters $\beta(t)$. That is, the mean value, $E[\beta(t)]$, and standard deviation, $\sqrt{\mathrm{Var}(\beta(t))}$, of fitted AR($p$) parameters are computed empirically over each available year and month. Based on this, the yearly and monthly variation coefficients are found as

$$\Delta = \frac{\sqrt{\mathrm{Var}(\beta(t))}}{E[\beta(t)]}. \tag{23}$$

The yearly and monthly variation coefficients reflect the stability of the speed of mean reversion over years and months, respectively. As the stratospheric temperature model is derived with lags in four days, this analysis will be performed with $p = 4$ for all computed AR parameters. The methodology for analyzing the yearly stability of the speed of mean reversion is as follows: (1) Loop through the 40 years in $\mathscr{S}_d$ and collect deseasonalized daily zonal mean stratospheric temperatures 365 days at a time, such that data for 1 January 1979 to 31 December 1979 are collected in one array, data for 1 January 1980 to 31 December 1980 in one array, and so on until the last array containing data for 1 January 2018 to 31 December 2018. Then (2) collect the 40 arrays containing all data each year in a single array to form the nested array $\boldsymbol{y}$, so $\dim(\boldsymbol{y}) = 1 \times 40 \times 365$. (3) Loop through the 40 arrays in $\boldsymbol{y}$, where an AR(4) model is fit to each of the years 1979 to 2018. The result is 40 arrays of AR parameters $(\beta_1^y, \beta_2^y, \beta_3^y, \beta_4^y)$, $y \in \{1, 2, \ldots, 40\}$. (4) Collect all 40 AR parameters corresponding to the same lag in one array and compute the statistics. That is, make the arrays $\boldsymbol{\beta}_1^y = (\beta_1^1, \beta_1^2, \ldots, \beta_1^{40}), \ldots, \boldsymbol{\beta}_4^y = (\beta_4^1, \beta_4^2, \ldots, \beta_4^{40})$, and compute the empirical mean, standard deviation, and finally the variation coefficients defined in Eq. (23), for each array $\boldsymbol{\beta}_1^y$, $\boldsymbol{\beta}_2^y$, $\boldsymbol{\beta}_3^y$ and $\boldsymbol{\beta}_4^y$. The results from this yearly stability analysis are presented in Table 5. Further, the methodology for the monthly stability analysis of speed of mean reversion is as follows: (1) Define one array for each month: $\boldsymbol{Jan}$, $\boldsymbol{Feb}$, $\boldsymbol{Mar}$, …, $\boldsymbol{Dec}$. (2) Loop through the 40 arrays in $\boldsymbol{y}$ (which is constructed in point 2 for the yearly stability analysis). For all 40 arrays, collect elements 0 to 30 in $\boldsymbol{Jan}$, elements 31 to 58 in $\boldsymbol{Feb}$, elements 59 to 89 in $\boldsymbol{Mar}$, and so on until you

reach elements 334 to 364 which are collected in $\boldsymbol{Dec}$. The resulting arrays are nested arrays of the form

$$\boldsymbol{Jan} = \left[\boldsymbol{Jan}^1, \ldots, \boldsymbol{Jan}^{40}\right], \boldsymbol{Feb} = \left[\boldsymbol{Feb}^1, \ldots, \boldsymbol{Feb}^{40}\right], \ldots,$$
$$\boldsymbol{Dec} = \left[\boldsymbol{Dec}^1, \ldots, \boldsymbol{Dec}^{40}\right].$$

Each array has dimension $1 \times 40 \times n$, where n corresponds to the number of days in that particular month. That is, $n = 31$ for $\boldsymbol{Jan}$, $n = 28$ for $\boldsymbol{Feb}$, and so on. (3) Collect the arrays $\boldsymbol{Jan}, \ldots, \boldsymbol{Dec}$ in a nested array $\boldsymbol{m}$: $\dim(\boldsymbol{m}) = 1 \times 12 \times 40 \times n$. Loop through each of the 480 months in $\boldsymbol{m}$, and fit an AR(4) model to each of the Januaries of the years 1979 to 2018, to each of the Februaries of the years 1979 to 2018, and so on until the last fit is performed on the data of December of 2018. The result is 480 arrays of AR parameters $(\beta_1^m, \beta_2^m, \beta_3^m, \beta_4^m)$, $m \in \{1, 2, \ldots, 480\}$. (4) Make the arrays $\boldsymbol{\beta}_1^m = (\beta_1^1, \beta_1^2, \ldots, \beta_1^{480}), \ldots, \boldsymbol{\beta}_4^m = (\beta_4^1, \beta_4^2, \ldots, \beta_4^{480})$, and for each of them, compute the variation coefficient as defined in Eq. (23). Further, because of the constructed order of the parameters, the arrays $\boldsymbol{\beta}_1^m$, $\boldsymbol{\beta}_2^m$, $\boldsymbol{\beta}_3^m$, and $\boldsymbol{\beta}_4^m$ can be used to analyze the seasonal behavior of the speed of mean reversion. The results computed from this monthly stability analysis are presented in Table 5, where the variability coefficients will reveal any monthly instability of speed of mean reversion. With the intention of detecting any monthly seasonal behavior in the four AR parameters, $\boldsymbol{\beta}_1^m$, $\boldsymbol{\beta}_2^m$, $\boldsymbol{\beta}_3^m$, and $\boldsymbol{\beta}_4^m$ are plotted in Fig. 8a–d.

4.2 Interpretations of the Monthly Stability Analysis

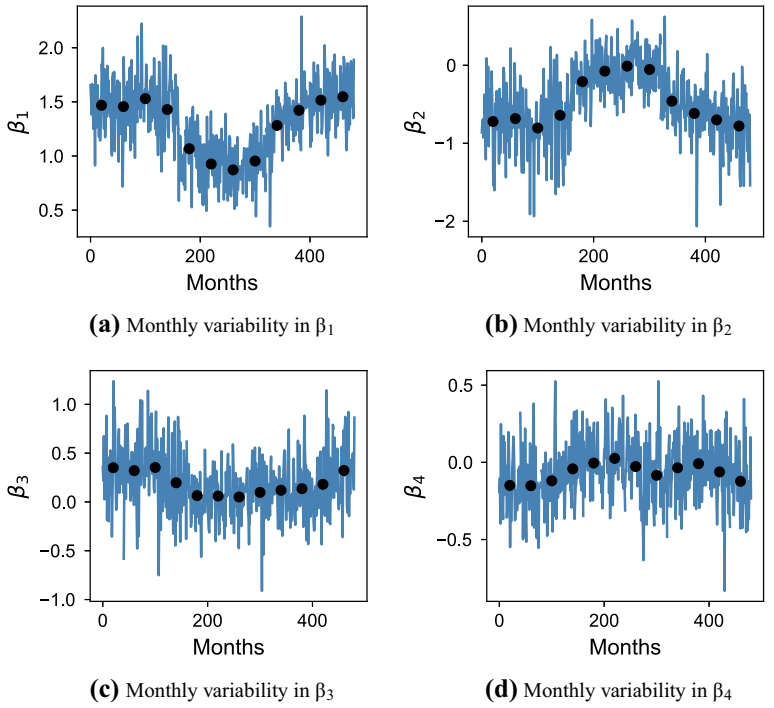
The monthly variation coefficients are more extreme than the yearly ones. Therefore, the remainder of this section will focus on interpreting results from the monthly stability analysis and will incorporate the observed time-varying behavior of the speed of mean reversion into the stratospheric temperature model dynamics in Eq. (10).

The magnitudes of the monthly variation coefficients (see Table 5) for all four AR parameters indicate that the assumption of constant speed of mean reversion in the stratospheric temperature model dynamics [Eq. (2)] is insufficient. The monthly variation coefficient, Δ, of the first lag parameter, β_1, is small compared to Δ of the other three lag parameters. However, as seen in Fig. 8a–d, the magnitude of β_1 is up to many times larger than the magnitudes of β_2, β_3, and β_4 (where the magnitudes of β_1, β_2, β_3, and β_4 are assessed over the contents of $\boldsymbol{\beta}_1^m$, $\boldsymbol{\beta}_2^m$, $\boldsymbol{\beta}_3^m$, and $\boldsymbol{\beta}_4^m$, respectively). This means that larger variability in the latter lag parameters affects the estimated stratospheric temperature less. Despite this observation, all variability coefficients are too large to be ignored, meaning that time-dependent AR parameters should be used rather than constant ones. By the transformation relation between CAR(4) and AR(4) models in Eq. (7), it is clear that the stratospheric temperature model dynamics should be given by the OU process in Eq. (10), rather than the one in Eq. (2).

It is not only the monthly variability coefficients of the AR parameters that suggest time-varying speed of mean reversion. The sequential patterns of $\boldsymbol{\beta}_1^m$ and $\boldsymbol{\beta}_2^m$ in Fig. 8a, b clearly show that the AR parameters β_1 and β_2 are seasonally varying. Both the first,

Table 5 Mean value, standard deviation, and absolute value of variability coefficient for the four parameters of an AR(4) process with yearly and monthly varying parameters, respectively

Parameter	β_1	β_2	β_3	β_4
Yearly				
Mean value	1.54	−0.74	0.28	−0.12
Standard deviation	0.12	0.22	0.18	0.09
abs(Δ)	7.8%	29.9%	64.3%	72.6%
Monthly				
Mean value	1.29	−0.48	0.19	−0.06
Standard deviation	0.34	0.45	0.32	0.20
abs(Δ)	26.6%	94.8%	170.7%	312.9%



(a) Monthly variability in β_1 **(b)** Monthly variability in β_2

(c) Monthly variability in β_3 **(d)** Monthly variability in β_4

Fig. 8 Monthly variability in the four AR parameters of the stratospheric temperature model

β_1, and the second, β_2, AR parameters are smaller in magnitude in summer than in winter. This means that summertime stratospheric temperature is less dependent on the stratospheric temperature the last 2 days than wintertime stratospheric temperature. This tendency is also (albeit less) evident for the third, β_3, and fourth, β_4, AR parameters in Fig. 8c, d.

To the best of our knowledge, there is no previous research specifically on the mean reverting property of stratospheric temperature. In Zapranis and Alexandridis (2008), daily values of speed of mean reversion for surface temperature are estimated by use of a neural network, revealing strong time dependence. Even though daily variation

in the speed of mean reversion is not studied in the current paper, a similar conclusion is reached: the speed of mean reversion of stratospheric temperature is dependent on time. However, Zapranis and Alexandridis (2008) found no signs of seasonal patterns, unlike the current study for stratospheric temperature, where a clear seasonal pattern is observed in the monthly estimated AR parameters. In response to the observation of time dependence in speed of mean reversion of surface temperature, Benth and Khedher (2015) presented a generalized version of the state-of-the-art stochastic models for surface temperatures applied in mathematical finance (see Sect. 2.2), where the speed of mean reversion of the driving (standard) OU process is a stochastic process. A simplified version of this generalized process (however multidimensional, and Lévy-driven rather than Brownian motion-driven) is presented in the current paper to incorporate time variability in speed of mean reversion. That is, the matrix $A(t)$ holding parameters of speed of mean reversion is assumed to be time-dependent and deterministic as presented in Theorem 1.

Before presenting an explicit time-dependent and deterministic matrix $A(t)$ representing the speed of mean reversion of stratospheric temperature, its time-varying behavior will be discussed. As already mentioned, the monthly AR parameters (which the speed of mean reversion depends directly upon) have large variability coefficients, Δ, as well as a seasonal behavior. From the definition of Δ in Eq. (23), one can see that the seasonal behavior increases the computed variability coefficients considerably. By removing the seasonal behavior in the AR parameters (see below), noise is still present, indicating that the speed of mean reversion could be modeled by a stochastic process. However, this is beyond the scope of this paper, and the noise is assumed to be negligible. For this reason, time dependence in speed of mean reversion is assumed to come solely from the seasonal variations.

As discussed in Sect. 3.3, the only long-term (perfectly) periodic phenomenon in the stratosphere is the yearly cycle. This is clearly seen in the stratospheric temperature data presented in Fig. 1. Further, as seen in Fig. 5, this phenomenon affects the variability, as well as volatility in variability, of stratospheric temperature. Physical explanations of this behavior are discussed in Haynes (2005), where the wintertime stratosphere is said to be more disturbed than the summertime stratosphere. Based on this, together with the above discussion concluding stronger speed of mean reversion in winter than in summer, it might be reasonable to assume that the yearly cycle affects the speed of mean reversion of stratospheric temperature as well. Stated in another way, large values of stratospheric temperature variance seem to generate larger dependence on stratospheric temperature the last couple of days. This is a topic for further research.

An explicit deterministic matrix $A(t)$ representing speed of mean reversion of stratospheric temperature is proposed in the following. By introducing time-varying AR(4) parameters such that each month of the year holds fixed parameters, seasonal variability will be adjusted for on a monthly basis. The result of the monthly variation analysis is exploited to define such time-varying AR(4) parameters. That is, define monthly parameter values as (remember that the initial dataset $\mathscr{S}$ contains stratospheric temperature values over 40 years)

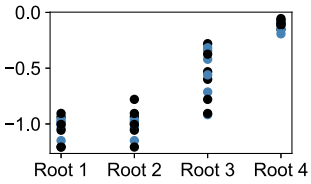


Fig. 9 Real part of the roots of the eigenvalue equations [Eq. (19)] of the monthly-varying CAR(4) parameters

$$\beta_k^M \triangleq \begin{cases} \beta_k^1 \\ \beta_k^2 \\ \vdots \\ \beta_k^{12} \end{cases} = \begin{cases} E[\boldsymbol{\beta}_k^m], & \text{for } m \in [1, 40] \\ E[\boldsymbol{\beta}_k^m], & \text{for } m \in [41, 80] \\ \quad \vdots \\ E[\boldsymbol{\beta}_k^m], & \text{for } m \in [441, 480], \end{cases}$$

where $M \in \{1, 2, \ldots, 12\}$ represents January to December, respectively, $k \in \{1, 2, 3, 4\}$ and $E[\cdot]$ represents the empirical mean. The computed values of β_1^M, β_2^M, β_3^M, and β_4^M are marked in Fig. 8a–d, respectively. Building on the theory in Sect. 2.2, the corresponding CAR(4) model is given by the multidimensional OU process in Eq. (10) (Theorem 1). Time dependence in the functions $\alpha_1(t), \ldots \alpha_4(t)$ does not matter for the transformation relation between CAR and AR models as long as the discretization scheme is chosen properly (Iacus 2008; Benth et al. 2008). That is, in this specific case, the discretization scheme has to be constructed such that the two time points which define the current scheme time step never belong to two different months. Hence, the continuous parameter counterparts, $\alpha_k(t)$, of the β_k^Ms can be computed by the transformation relation in Eq. (7). The $\alpha_k(t)$s can be considered as 12-level step functions, where each step represents a month of the year. The roots of the eigenvalue equation [Eq. (19)] of each of the 12 steps in each of the functions $\alpha_k(t)$ are computed, and their real parts are shown in Fig. 9. As all the roots have a negative real part, the stationarity condition of CAR processes is secured.

Repeating the analysis in Sect. 3.5 with model residuals

$$e(t) = X(t) - \sum_{k=1}^{4} \beta_k^M X(t-k)$$

also gives residuals on the form $e(t) = \sigma(t)\epsilon(t)$, with $\epsilon(t)$ being NIG-distributed random variables with memory effects. This confirms that the proposed CAR(4) model from Sect. 2.2 is suitable to model stratospheric temperature dynamics when driven by the multidimensional OU dynamics in Eq. (10); see Theorem 1.

5 Conclusions and Further Work

In this paper, a novel stochastic model for stratospheric temperature is proposed. By time series analysis, it was shown that stratospheric temperature can be approximated by an AR(4) process added to a deterministic seasonality function. The seasonality function captures periodic (yearly seasonal) effects as well as the long-term

trend to model stratospheric cooling. The scaled model residuals were shown to be NIG-distributed. By exploiting the connection between AR and CAR processes, a continuous-time model for stratospheric temperature was developed. It was shown that a Lévy-driven CAR(4) process with time-dependent volatility is well suited as a continuous-time, stochastic model for deseasonalized stratospheric temperature.

Some ideas for future work include incorporating a volatility which is stochastic, not just time-dependent, into the model. Furthermore, it may be of interest to further investigate why large values of variance in the stratospheric temperature seem to generate a larger dependency on stratospheric temperature the previous days. Developing a continuous-time, stochastic model for stratospheric wind, potentially as a joint model with stratospheric temperature, is also relevant. In addition, based on the model presented in the current paper, one may exploit stratosphere–troposphere coupling in order to develop improved methods for pricing of weather derivatives (on a surface level). A current work in progress is to develop a dual model for stratospheric temperature, where winter and summer temperatures are studied separately. Such a model is particularly useful when analyzing, for example, pure winter phenomena, such as sudden stratospheric warming.

Author contributions MDE is the principal author of this paper. Conceptualization: MDE, SM, SPN, KRD. Methodology: MDE, KRD. Data collection/preparation, formal analysis and investigation: MDE. Writing—original draft preparation: MDE. Writing—review and editing: MDE, KRD, SM, SPN. Supervision: SM, SPN, KRD.

Data availability Data retrieved from https://apps.ecmwf.int/datasets/data/interim-full-daily/levtype%3Dpl/levelist-param%3D10%3B130.128/, 12 October 2020. See data specifications in Table 1.
Code availability The author plans to make the code publicly available when the article is published.

Declarations

Funding The PhD grant of M. D. Eggen is funded by NORSAR. This work was supported by the Research Council of Norway, FRIPRO Young Research Talent SCROLLER project, Grant number 299897, as well as the Research Council of Norway FRIPRO/FRINATEK project MADEIRA, Grant number 274377.

Conflict of interest The authors have no conflicts of interest to declare.

References

Baldwin MP, Ayarzagüena B, Birner T, Butchart N, Butler AH, Charlton-Perez AJ, Domeisen DIV, Garfinkel CI, Garny H, Gerber EP, Hegglin MI, Langematz U, Pedatella NM (2021) Sudden stratospheric warmings. Rev Geophys 59(1):e2020RG000708

Baldwin MP, Birner T, Brasseur G, Burrows J, Butchart N, Garcia R, Geller M, Gray L, Hamilton K, Harnik N, Hegglin MI, Langematz U, Robock A, Sato K, Scaife AA (2019) 100 years of progress in understanding the stratosphere and mesosphere. Meteorol Monogr 59(1):27.1-27.62

Baldwin MP, Dunkerton TJ (2001) Stratospheric harbingers of anomalous weather regimes. Science 294(5542):581–584

Baldwin MP, Gray LJ, Dunkerton TJ, Hamilton K, Haynes PH, Randel WJ, Holton JR, Alexander MJ, Hirota I, Horinouchi T, Jones DBA, Kinnersley JS, Marquardt C, Sato K, Takahashi M (2001) The quasi-biennial oscillation. Rev Geophys 39(2):179–229

Barndorff-Nielsen OE (1997) Normal inverse Gaussian distributions and stochastic volatility modelling. Scand J Stat 24(1):1–13

Barndorff-Nielsen OE (1997) Processes of normal inverse Gaussian type. Finance Stoch 2(1):41–68

Barndorff-Nielsen OE, Shephard N (2001) Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. J R Stat Soc Ser B Stat Methodol 63(2):167–241

Benth FE, Khedher A (2015) Weak Stationarity of Ornstein–Uhlenbeck processes with stochastic speed of mean reversion. In: Podolskij M, Stelzer R, Thorbjørnsen S, Veraart A (eds) The fascination of probability, statistics and their applications. Springer, Cham, pp 153–189

Benth FE, Klüppelberg C, Müller G, Vos L (2014) Futures pricing in electricity markets based on stable CARMA spot models. Energy Econ 44:392–406

Benth FE, Taib CMIC (2012) On the speed towards the mean for carma processes with applications to energy markets

Benth FE, Taib CMIC (2013) On the speed towards the mean for continuous time autoregressive moving average processes with applications to energy markets. Energy Econ 40:259–268

Benth FE, Šaltytė Benth J (2005) Stochastic modelling of temperature variations with a view towards weather derivatives. Appl Math Finance 12(1):53–85

Benth FE, Šaltytė Benth J (2009) Dynamic pricing of wind futures. Energy Econ 31(1):16–24

Benth FE, Šaltytė Benth J (2013) Modeling and pricing in financial markets for weather derivatives. World Scientific, Singapore

Benth FE, Šaltytė Benth J, Koekebakker S (2008) Stochastic modelling of electricity and related markets, vol 11. Advanced series on statistical science & applied probability. World Scientific Publishing Pte, Singapore

Berrisford P, Dee DP, Poli P, Brugge R, Fielding M, Fuentes M, Kållberg PW, Kobayashi S, Uppala S, Simmons A (2011) The ERA-Interim archive Version 2.0 (1):1–23

Brockwell PJ (2001) Lévy-driven CARMA processes. Ann Inst Stat Math 53(1):113–124

Brockwell PJ (2004) Representations of continuous-time ARMA processes. J Appl Probab 41(A):375–382

Brockwell PJ (2014) Recent results in the theory and applications of CARMA processes. Ann Inst Stat Math 66(4):647–685

Brockwell PJ, Lindner A (2015) Prediction of Lévy-driven CARMA processes. J Econom 189(2):263–271

Butler A, Charlton-Perez A, Domeisen DI, Garfinkel C, Gerber EP, Hitchcock P, Karpechko AY, Maycock AC, Sigmond M, Simpson I, Son SW (2019) Sub-seasonal predictability and the stratosphere. In: Robertson AW, Vitart F (eds) Sub-seasonal to seasonal prediction. Elsevier, Amsterdam, pp 223–241

Butler AH, Seidel DJ, Hardiman SC, Butchart N, Birner T, Match A (2015) Defining sudden stratospheric warmings. Bull Am Meteorol Soc 96(11):1913–1928

Clewlow L, Strickland C (2000) Energy derivatives-pricing and risk management. Lacima Publishers, Pretoria

Cnossen I, Laštovička J, Emmert JT (2015) Introduction to special issue on "long-term changes and trends in the stratosphere, mesosphere, thermosphere and ionosphere." J Geophys Res Atmos 120(22):11401–11403

Danilov AD, Konstantinova AV (2020) Long-term variations in the parameters of the middle and upper atmosphere and ionosphere (review). Geomagn Aeron 60(4):397–420

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars ACM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kållberg P, Köhler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette JJ, Park BK, Peubey C, de Rosnay P, Tavolato C, Thépaut JN, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656):553–597

Fu Q, Solomon S, Lin P (2010) On the seasonal dependence of tropical lower-stratospheric temperature trends. Atmos Chem Phys 10(6):2643–2653

Haynes P (2005) Stratospheric dynamics. Annu Rev Fluid Mech 37(1):263–293
Iacus SM (2008) Simulation and inference for stochastic differential equations: with R examples. Springer, New York
Intergovernmental Panel on Climate Change (2014) Climate change 2013—the physical science basis: Working Group I contribution to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press
Jones PD, Wigley TML (1990) Global warming trends. Sci Am 263(2):84–91
Karpechko A, Tummon F, WMO Secretariat (2016) Climate predictability in the stratosphere. WMO Bull 65(1)
Kitchin J (2013) Smooth transitions between discontinuous functions
Levendis JD (2018) Time series econometrics: learning through replication. Springer, Berlin
McCormack JP, Hood LL (1996) Apparent solar cycle variations of upper stratospheric ozone and temperature: latitude and seasonal dependences. J Geophys Res Atmos 101(D15):20933–20944
Pedatella N, Chau J, Schmidt H, Goncharenko L, Stolle C, Hocke K, Harvey V, Funke B, Siddiqui T (2018) How sudden stratospheric warming affects the whole atmosphere. Eos 99
Steiner AK, Ladstädter F, Randel WJ, Maycock AC, Fu Q, Claud C, Gleisner H, Haimberger L, Ho SP, Keckhut P, Leblanc T, Mears C, Polvani LM, Santer BD, Schmidt T, Sofieva V, Wing R, Zou CZ, Cardon C (2020) Observed temperature changes in the troposphere and stratosphere from 1979 to 2018. J Clim 33(19):8165–8194
Sévellec F, Drijfhout SS (2018) A novel probabilistic forecast system predicting anomalously warm 2018–2022 reinforcing the long-term global warming trend. Nat Commun 9(1):3024
Vallis GK (2017) The Stratosphere. In: Atmospheric and oceanic fluid dynamics. Cambridge University Press, pp 627–671
Zapranis A, Alexandridis A (2008) Modelling the temperature time-dependent speed of mean reversion in the context of weather derivatives pricing. Appl Math Finance 15(4):355–386