# Analysis of time series generated by low-order models of atmospheric dynamics

Alexander Gluhovsky*

**Abstract**
A novel class of time series models for atmospheric data is introduced based on recent results on statistical properties of dynamical systems.

**Key Words:** Time series, dynamical systems, Lorenz model, gyrostatic low-order models

## 1. Introduction

Time series analysis traditionally plays an important role in atmospheric and climate studies. Observed series, however, are often prohibitively short, with only one realization typically available, and standard models prove inadequate since conventional statistical methods involve strong assumptions that are rarely met in real data (e.g., Ghil et al. 2002). These problems are illustrated in Section 2 with a typical example that motivates a new approach to atmospheric time series analysis based on low-order models of atmospheric dynamics.

Fortunately, in addition to the flood of data, we also have the equations of atmospheric dynamics, which are a reliable part of our knowledge. Following the pioneering work by Kolmogorov (see Pasini and Pelino, 2000), Lorenz (1963, 1982), and Obukhov (1969, 1973), a helpful way to deal with formidable difficulties posed by these partial differential equations is to approximate them with finite systems of ordinary differential equations, the so-called low-order models (LOMs). This brings us to dynamical systems, where considerable progress has been achieved in understanding their statistical properties (e.g., Collet and Eckmann 2006). A dynamical system is a set equipped with a time evolution, which may be given, for example, by a LOM. Such dynamical systems may have invariant measures as well as other statistical properties, like ergodicity, mixing, etc. In particular, statistical properties have been found out for the celebrated Lorenz (1963) model,

$$\begin{aligned} \dot{x} &= \sigma(y-x), \\ \dot{y} &= -xz+rx-y, \\ \dot{z} &= xy-bz. \end{aligned} \tag{1}$$

Specifically, it has been proved recently that the Lorenz flow possesses a physical ergodic invariant probability measure (Arajo et al., 2009) and satisfies the central limit theorem (Holland and Melbourne 2007, Arajo and Varandas 2012).

Lorenz found that much of the irregularity in observed time series, traditionally attributed to random forcing by an immense number of variables, could be generated by nonlinear interactions of just three variables. This has radically changed our understanding of atmospheric dynamics and turbulence, but many attempts to extend model (1) to obtain larger, more realistic models often led to LOMs exhibiting unphysical behaviors.

In this talk, we suggest *physically sound* extensions of model (1) as novel time series models for atmospheric studies. Section 3 briefly reviews such extensions developed by the author in the form of coupled systems known in mechanics as Volterra gyrostats

*Department of Earth, Atmospheric, and Planetary Sciences and Department of Statistics, Purdue University, West Lafayette, IN 47907

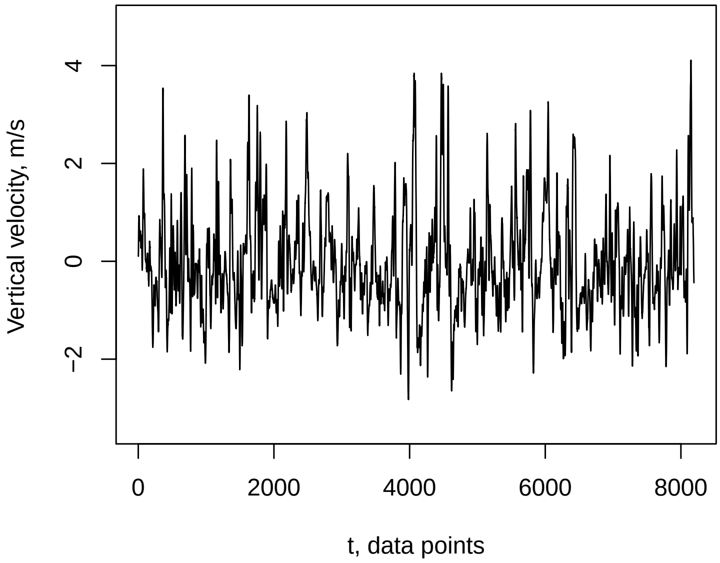


**Figure 1**: Record of 20-Hz aircraft vertical velocity measurements over Lake Michigan. Figure from Gluhovsky (2011).

(Gluhovsky 2006, and references therein), while Section 4 demonstrates in a pilot study how such gyrostatic LOMs may serve as a viable alternative to models commonly used in atmospheric time series analysis.

## 2. Motivating Example

### 2.1 Observational Data

Consider a record in Fig.1 of the vertical velocity of wind in a convective boundary layer during an outbreak of a polar air mass over the Great Lakes region. The record consists of 8192 data points over about 29 km across Lake Michigan, 50 m above the lake, and it has passed a test for stationarity from Gluhovsky and Agee (1994). The sample mean, variance, skewness, and kurtosis of the vertical velocity computed from this record are $-0.04$, $1.06$, $0.83$, and $4.10$, respectively. The elevated skewness and kurtosis may indicate nonlinearities in the underlying data generating mechanism (DGM), but these sample characteristics are just point estimates of the true values of the parameters, and therefore confidence intervals (CIs) are employed to learn how far one can trust such numbers.

Here is the problem with CIs for parameters of atmospheric time series, which are produced by the inherently nonlinear system. The *target* coverage probability is attained only if the assumptions underlying the method for the CI construction are met. Since for atmospheric time series this is rarely the case, the *actual* coverage probability may differ from the target level (sometimes considerably). For example, when the DGM is linear, CIs for the mean or the variance of the time series may be found analytically, but the common practice of computing CIs from fitted linear models may result in erroneous CIs when the real DGM is, in fact, nonlinear (Gluhovsky and Agee 2007). Moreover, CIs for the skewness cannot be based on linear models that imply zero skewness. Thus, there is need for nonlinear models, but selecting an appropriate one is problematic.

**Table 1**: Parameters of the model time series $X_t$ (Eq.(4)) distribution vs sample characteristics of the observed series $W_t$ in Fig. 1.

| | $X_t$ | $X_t$ at $a = 0.145$ | $W_t$ |
|---|---|---|---|
| Mean | $M = 0$ | $0$ | $-0.04$ |
| Variance | $V = 1 + 2a^2$ | $\approx 1.04$ | $1.06$ |
| Skewness | $S = \frac{6a+8a^3}{V^{3/2}}$ | $\approx 0.84$ | $0.83$ |
| Kurtosis | $K = \frac{3+60a^2+60a^4}{V^2}$ | $\approx 3.95$ | $4.10$ |

### 2.2 Subsampling Confidence Intervals

One may, of course, resort to resampling methods; subsampling (Politis et al. 1999), for example, where the available record of $n$ observations is divided into overlapping blocks of the same length, $b$:

$$\underbrace{X_1, ..., X_b}_{b},\ ...,\ \underbrace{X_i, ..., X_{i+b-1}}_{b},\ ...,\ \underbrace{X_{n-b+1}, ..., X_n}_{b}. \quad (2)$$

These blocks, *subsamples*, retain the dependence structure of the time series, and subsampling provides CIs with asymptotically correct coverage from a single record (without having to fit a model, linear or nonlinear, and to make questionable assumptions about the DGM), when

$$b \to \infty \qquad \text{and} \qquad b/n \to 0 \qquad \text{as} \qquad n \to \infty. \quad (3)$$

But real records are typically too short to satisfy conditions (3), so that approximating nonlinear models are needed to assess the *actual* coverage and to fix accordingly the subsampling CIs.

#### *2.2.1 Approximating Model*

For the vertical velocity data in Fig.1, the following model from Lenschow et al. (1994) was employed in Gluhovsky (2011),

$$X_t = Y_t + a(Y_t^2 - 1), \quad (4)$$

where $Y_t$ is a first order autoregressive process (AR(1)),

$$Y_t = \phi Y_{t-1} + \epsilon_t, \quad (5)$$

$0 < \phi < 1$ and $a$ are constants, and $\epsilon_t$ is a white noise process with mean $0$ and variance $\sigma_\epsilon^2$). In simulations, $\sigma_\epsilon^2 = 1 - \phi^2$ (so that $\sigma_Y^2 = 1$), the records contain 2048 data points, and $\phi = 0.83$, the latter to imitate the dependence structure of the vertical velocity time series in Fig.1 as characterized by autocorrelation functions. Since at $a = 0.145$, the first four moments of $X_t$ (in model (4)) are close to the corresponding sample characteristics of the observed series (see Table1), one may presume that the model is adequate for fixing subsampling CIs. Still, there is no guarantee that the simulated and real data are similar in all aspects.

Thus, in any case, models are needed for observed time series, but linear ones are often inadequate, while finding appropriate nonlinear models is challenging since there are a myriad of possibilities to explore. For this reason, LOMs with chaotic behavior are considered below, whose dynamics is inherently related to that of atmospheric systems.

# 3. Gyrostatic LOMs of Atmospheric Dynamics

## 3.1 The Simplest Volterra Gyrostat as a Mechanical Analog of the Lorenz Model

Consider a classical mechanical system known as the Volterra gyrostat (Volterra 1899, Wittenburg 1977), which can be written as (e.g., Gluhovsky 2006)

$$\begin{aligned} \dot{x}_1 &= px_2x_3 + bx_3 - cx_2, \\ \dot{x}_2 &= qx_2x_3 + cx_1 - ax_3, \\ \dot{x}_3 &= rx_2x_3 + ax_2 - bx_1; \\ & p + q + r = 0. \end{aligned} \tag{6}$$

Note that unlike linear friction terms, linear terms in Eqs.(6) (*linear gyrostatic terms*) do not affect the conservation of energy nor the conservation of phase space volume.

The Lorenz (1963) model (Eqs.(1)) was shown to be equivalent to the simplest Volterra gyrostat ($r = b = c = 0$ in Eqs.(6)) with added constant forcing and linear friction (Gluhovsky 1982):

$$\begin{array}{r|r|l} \dot{x}_1 = & -x_2x_3 & -\alpha_1 x_1 + F, \\ \dot{x}_2 = & x_3x_1 - x_3 & -\alpha_2 x_2, \\ \dot{x}_3 = & x_2 & -\alpha_3 x_3. \end{array} \tag{7}$$

It was later found (see Gluhovsky (2006) and references therein) that effective LOMs for atmospheric circulations and turbulence could be developed as systems of coupled gyrostats (6). These always have a quadratic integral of motion (interpreted as some form of energy), which eliminates unphysical behaviors that have often plagued LOMs obtained through *ad hoc* Galerkin truncations, while increasing the order of approximation just adds more gyrostats to the resulting LOM.

## 3.2 Gyrostatic Form of the Charney-DeVore (1979) Model

Another feature of systems of coupled gyrostats, of particular importance for this study, is that mechanisms peculiar to atmospheric dynamics (e.g., stratification, rotation , topography, shear, magnetohydrodynamic effects) bring about linear gyrostatic terms in gyrostats that form the LOM. For example, the popular Charney-DeVore (1979) model for the quasi-geostrophic potential vorticity equation involves two such mechanisms, topography and rotation, and accordingly, the gyrostatic form of the model (Gluhovsky et al., 2002),

$$\begin{array}{l|r|r|l|l|} \dot{x}_1 = f_1 - \alpha x_1 & +b_1x_3 & & & \\ \dot{x}_2 = \quad -\alpha x_2 & +q_1x_3x_1 - a_1x_3 & & +q_3x_4x_6 & \\ \dot{x}_3 = \quad -\alpha x_3 & -q_1x_1x_2 + a_1x_2 - b_1x_1 & & & -q_3x_4x_5 \\ \dot{x}_4 = f_2 - \alpha x_4 & & & +p_3x_6x_2 + b_3x_6 & -p_3x_5x_3 \\ \dot{x}_5 = \quad -\alpha x_5 & & +q_2x_6x_1 - a_2x_6 & & -r_3x_3x_4 \\ \dot{x}_6 = \quad -\alpha x_6 & & -q_2x_1x_5 + a_2x_5 & +r_3x_2x_4 - b_3x_4 & \end{array} \tag{8}$$

consists of gyrostats that have two kinds of linear gyrostatic terms: those with coefficients $a_i$ are due to topography, those with $b_i$ are caused by rotation. Note that all gyrostats in the model are different from the gyrostat in Eqs.(7).

# 4. Gyrostatic LOMs as Time Series Models

Returning now to time series models for the vertical velocity series in Fig.1, let us see how gyrostatic LOMs may enter the picture here.

### 4.1 Time Series Generated by the Lorenz Model

The basic mechanism responsible for producing time series in Fig.1 is the Rayleigh-Bénard convection, so model (1) (or rather its gyrostatic equivalent, LOM (7)) was a natural one to begin with. This did produce non-Gaussian kurtosis of $K = 2.3$, but the skewness remained equal to zero, both far off the sample characteristics of the observed time series $(S = 0.83, K = 4.1)$ with the subsampling CIs $(0.56, 1.1)$ and $(3.7, 4.5)$, respectively (from Gluhovsky (2011) using approximating model described in Section 2.2.1).

### 4.2 Time Series Generated by a Gyrostatic Extension of the Lorenz Model

Whereas the Rayleigh-Bénard convection is its principal mechanism, the dynamics over Lake Michigan involves a hoist of others, such as large-scale vertical motion, cloud top entrainment instability, latent heat release, and gravity waves (Atkinson and Zhang 1996), which would have resulted in new linear gyrostatic terms in the LOM had we attempted to derive it from the governing equations. Putting this off for the future, just one new pair of linear gyrostatic terms was added in model (7) as representing all such mechanisms,

$$\begin{array}{rl|l|l} \dot{x}_1 = & -x_2x_3 & +\,.35x_3 & -\alpha_1 x_1 + F, \\ \dot{x}_2 = & \quad x_3x_1 - x_3 & & -\alpha_2 x_2, \\ \dot{x}_3 = & \qquad\qquad x_2 & -\,.35x_1 & -\alpha_3 x_3, \end{array} \tag{9}$$

which has resulted in $(S = 0.81, K = 4.20)$. As it also shares some fundamental physics with the original system, even this simplest model could be helpful for the developing statistical tools to analyze atmospheric data sets, while larger systems of coupled gyrostats, being much closer to the original system, should be even more useful.

## 5. Concluding Remarks

In summary, I was trying to draw attention to gyrostatic LOMs as a viable alternative to time series models commonly used in atmospheric studies. On one hand, one can easily generate numerous records of any length with such LOMs (unattainable with more complex atmospheric models). And on the other hand, gyrostatic LOMs are derived from the original governing equations, whose fundamental properties they inherit, thereby infusing more physics into atmospheric time series analysis, with obvious benefits for developing statistical tools for trends, long memory and extremes. This is in contrast to models borrowed from traditional time series analysis, which take no advantage of the physics contained in the equations.

*Acknowledgments.* This work was supported by NSF Grant AGS-1050588.

## REFERENCES

Araújo, V., Pacifico, M. J., Pujals, E. R., and Viana, M. (2009), "Singular-hyperbolic attractors are chaotic," *Trans. Amer. Math. Soc.*, 361, 2431–2485.

Araújo, V., and Varandas, P. (2012) "Robust exponential decay of correlations for singular-flows," Preprint arXiv:1101.2915v3.

Atkinson, B. W., and Zhang, J. W. (1996), "Mesoscale shallow convection in the atmosphere," *Rev. Geophys.*, 34, 403–43.

Charney, J. G., and DeVore, J. G. (1979), "Multiple flow equilibria in the atmosphere and blocking," *J. Atmos. Sci.*, 36, 1205-1216.

Collet, P., and Eckmann, J.-P. (2006), *Concepts and Results in Chaotic Dynamics: A Short Course*, Springer.

Ghil, M., Allen, M. R., Dettinger, M. D., Ide, K, Kondrashov, D., Mann, M. E., Robertson, A.W., Saunders, A, Tian, Y., Varadi, F., and Yiou, P. (2002), “Advanced spectral methods for climate time series,” *Rev. Geophys.*, 40, 1-41.

Gluhovsky, A. (1982), “Nonlinear systems that are superpositions of gyrostats,” *Sov. Phys. Doklady*, 27, 823-825.

Gluhovsky, A., and Agee, E. (1994), “A definitive approach to turbulence statistical studies in planetary boundary layers,” *J. Atmos. Sci.*, 51, 1682-1690.

Gluhovsky, A., Tong, C, and Agee, E. M. (2002), “Selection of modes in convective low-order models,” *J. Atmos. Sci.*, 59, 1383-1393.

Gluhovsky, A. (2006), Energy-conserving and Hamiltonian low-order models in geophysical fluid dynamics, *Nonlin. Processes Geophys.*, 13, 125-133.

Gluhovsky, A., and Agee E. M. (2007), “On the analysis of atmospheric and climatic time series, *J. Appl. Meteorol. Climatol.*, 46, 1125-1129.

Gluhovsky, A. (2011), “Statistical inference from atmospheric time series: detecting trends and coherent structures,” *Nonlin. Processes Geophys.*, 18, 537–544.

Holland, M., and Melbourne, I. (2007), “Central limit theorems and invariance principles for Lorenz attractors,” *J. Lond. Math. Soc.*, 76(2), 345-364.

Lenschow, D., Mann, J., and Kristensen, L. (1994), “How long is long enough when measuring fluxes and other turbulence statistics?” *J. Atmos. Oceanic Technol.*, 11, 661–673.

Lorenz, E. N. (1963), “Deterministic nonperiodic flow,” *J. Atmos. Sci.*, 20, 130-141.

Lorenz, E. N. (1982), “Low-order models of atmospheric circulations,” *J. Meteorol. Soc. Jpn.*, 60, 255-267.

Obukhov, A. M. (1969), “On integral characteristics in hydrodynamic type systems,” *Sov. Phys. Dokl.*, 14, 32-35.

Obukhov, A. M. (1973), “On the problem of nonlinear interactions in fluid dynamics,” *Gerlands Beitr. Geophys.*, 82, 282-290.

Pasini, A., and Pelino, V. (2000), “A unified view of Kolmogorov and Lorenz systems,” *Phys. Lett. A*, 275, 435-446.

Politis, D. N., Romano, J. P., and Wolf, M. (1999), *Subsampling*, Springer.

Volterra, V. (1899), “Sur la théorie des variations des latitudes,” *Acta Math.*, 22, 201-358.

Wittenburg, J. (1977), *Dynamics of Systems of Rigid Bodies*, B. G. Teubner.