

Statistical Tests of Taylor's Hypothesis: An Application to Precipitation Fields

BO LI*

National Center for Atmospheric Research, Boulder, Colorado

ADITYA MURTHI, KENNETH P. BOWMAN, AND GERALD R. NORTH

Department of Atmospheric Sciences, Texas A&M University, College Station, Texas

MARC G. GENTON⁺

Department of Econometrics, University of Geneva, Geneva, Switzerland

MICHAEL SHERMAN

Department of Statistics, Texas A&M University, College Station, Texas

(Manuscript received 2 January 2008, in final form 20 June 2008)

ABSTRACT

The Taylor hypothesis (TH) as applied to rainfall is a proposition about the space–time covariance structure of the rainfall field. Specifically, it supposes that if a spatiotemporal precipitation field with a stationary covariance $\text{Cov}(\mathbf{r}, \tau)$ in both space \mathbf{r} and time τ moves with a constant velocity \mathbf{v} , then the temporal covariance at time lag τ is equal to the spatial covariance at space lag $\mathbf{r} = \mathbf{v}\tau$ that is, $\text{Cov}(\mathbf{0}, \tau) = \text{Cov}(\mathbf{v}\tau, 0)$. Qualitatively this means that the field evolves slowly in time relative to the advective time scale, which is often referred to as the frozen field hypothesis. Of specific interest is whether there is a cutoff or decorrelation time scale for which the TH holds for a given mean flow velocity \mathbf{v} . In this study, the validity of the TH is tested for precipitation fields using high-resolution gridded Next Generation Weather Radar (NEXRAD) reflectivity data produced by the WSI Corporation by employing two different statistical approaches. The first method is based on rigorous hypothesis testing, while the second is based on a simple correlation analysis, which neglects possible dependencies between the correlation estimates. Radar reflectivity values are used from the southeastern United States with an approximate horizontal resolution of 4 km \times 4 km and a temporal resolution of 15 min. During the 4-day period from 2 to 5 May 2002, substantial precipitation occurs in the region of interest, and the motion of the precipitation systems is approximately uniform. The results of both statistical methods suggest that the TH might hold for the shortest space and time scales resolved by the data (4 km and 15 min) but that it does not hold for longer periods or larger spatial scales. Also, the simple correlation analysis tends to overestimate the statistical significance through failing to account for correlations between the covariance estimates.

1. Introduction

Rainfall fields have complex spatial and temporal structures that span a wide range of scales. Many practical problems, such as precipitation estimation and data assimilation, require quantitative knowledge not just of the mean and variance of precipitation but also of its space–time covariance structure. Additionally, theoretical and modeling studies of precipitation can benefit from comparisons with high-quality observational estimates of precipitation covariance fields. A comprehensive understanding of the space–time structure of

* Current affiliation: Department of Statistics, Purdue University, West Lafayette, Indiana.

⁺ Current affiliation: Department of Statistics, Texas A&M University, College Station, Texas.

Corresponding author address: Bo Li, Department of Statistics, Purdue University, 150 North University Street, West Lafayette, IN 47907.

E-mail: boli@purdue.edu

rainfall fields has been a subject of considerable interest over the last decade or so and constitutes one of the most challenging problems in precipitation research.

Based on studies of turbulent flow, (Taylor 1938, p. 478) proposed a simple model of the covariance structure of a turbulent current with a constant mean background velocity. He hypothesized that small-scale turbulence might be carried along by the mean flow in such a way that the temporal covariance of a variable at time lag τ would be equal to the spatial covariance at space lag $\mathbf{r} = \mathbf{v}\tau$, where \mathbf{v} is the mean velocity of the flow. Specifically, consider a stationary spatiotemporal random field $Z(\mathbf{x}, t)$ at spatial location $\mathbf{x} \in \mathbb{R}^d$, $d \geq 1$, and time $t \in \mathbb{R}$. Let $C(\mathbf{r}, \tau) = \text{Cov}[Z(\mathbf{x}, t), Z(\mathbf{x} + \mathbf{r}, t + \tau)]$ denote its stationary covariance function for spatial lag \mathbf{r} and temporal lag τ . In this case the Taylor hypothesis (TH) can be written

$$C(\mathbf{0}, \tau) = C(\mathbf{v}\tau, 0). \quad (1)$$

Note that the mean velocity \mathbf{v} may not be known a priori. Although we consider TH in a stationary random field for simplicity, it is worth mentioning that TH can also be defined for a nonstationary random field where the covariance function and velocity \mathbf{v} have to be location specific (e.g., Burghlea et al. 2005). Under all circumstances, TH assumes the existence of a large mean flow compared to the velocity fluctuation, which allows the latter to be frozen into the former. TH is also of interest for many other fields such as the cloud and radiation test bed measurements (Sun and Thorne 1995) other than rainfall.

Several studies have attempted to test the TH for atmospheric convection by using radar observations of precipitation. Based on empirical correlations, Zawadzki (1973) argued that Taylor's hypothesis is plausible for precipitation data for temporal lags less than 40 min. His results seemed to indicate that spatial correlations exhibited more memory than temporal correlation. That study was based on only 11 radar scans from one radar for a single storm event, so its generality is not known. Gupta and Waymire (1987) and Cox and Isham (1988) studied the validity of the hypothesis for various theoretical space-time covariance models, but they did not test the TH with observations. Crane (1990) observed that the TH held in rainfall fields up to a time scale of around 30 min for spatial scales less than 20 km, after which it broke down.

Poveda and Zuluaga (2005) tested the validity of the TH for a set of 12 storms observed in southwestern Amazonia, Brazil, during the January–February 1999 Wet Season Atmospheric Mesoscale Campaign. They concluded that the TH did not hold in 9 out of the 12 studied storms, but that it did hold for 3 storms up to time scales of around 10–15 min. That time scale is

related to the life cycles of convective cells in the region. The generality of these results, however, is open to interpretation; their conclusions are not based on a statistical comparison of the spatial and temporal covariances. This points to the need for a formal statistical testing procedure to assess Taylor's hypothesis and the application of rigorous statistical methods to sizable datasets. Aside from rainfall, the study method may also be useful in investigating the cloud structure and the thermodynamics of moisture transport related to precipitation processes (Sun and Thorne 1995).

In this study, we use a rigorous statistical approach to test the TH based on the asymptotic joint normality of covariance estimators derived by Li et al. (2008). The method is applied to high-resolution gridded Next Generation Weather Radar (NEXRAD) reflectivity data. These results are compared with a simple estimate of statistical significance based on the assumption of independence of the covariance estimates.

2. Data

As a prototype observational dataset to test Taylor's hypothesis, we use high-resolution gridded NEXRAD reflectivity data produced by the WSI Corporation. Data from all available operating radars are routinely merged onto a longitude–latitude grid for the conterminous United States with an approximate horizontal resolution of $4 \text{ km} \times 4 \text{ km}$. The radar reflectivity data are maps depicting the highest reflectivity measured above each grid box, computed from scans at multiple elevation angles. Reflectivity values Z are transformed logarithmically using $\text{dBZ} = 10 \log_{10}(Z)$ and discretized with 5 dBZ precision from 0 to 75 dBZ. Multiple levels of quality control to remove ground clutter and false echoes, along with multiple volume scans, which nominally require 5 to 6 min, are used to produce gridded analyses with 15-min temporal resolution (96 time steps per day). For this study, we use radar reflectivity in dBZ rather than rain rates. Reflectivity is an observable parameter, so it is physically reasonable to ask whether it obeys Taylor's hypothesis. Calculations using estimated rain rates R yield similar results (further discussion below).

Although atmospheric convection is a three-dimensional phenomenon, we treat the precipitation as two-dimensional by using the radar rainfall composites described in the preceding paragraph. The national radar grids are 3661×1837 grid cells. For our analysis, we select a subarea of the grid and a time period during which there is substantial rainfall. We chose 4 days of gridded radar reflectivities from the southeastern United States ($30^\circ\text{--}40^\circ\text{N}$, $75^\circ\text{--}100^\circ\text{W}$) for the period 2 to 5 May 2002 ($4 \text{ days} \times 96 \text{ time steps per day} = 384 \text{ time steps}$).

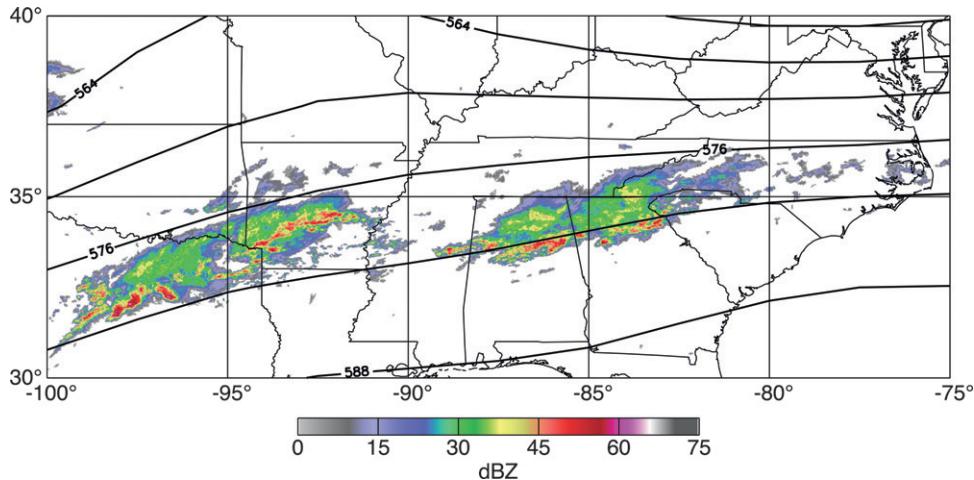


FIG. 1. Composite radar reflectivity map at 1200 UTC 3 May 2002. Contours are 500-hPa geopotential height in dam.

The selected region has 1308×558 grid cells. During this period there are no time steps with complete missing grids, although individual radars may not have been continuously available. A cold front moved into the region from the northwest early in the study period and then became stationary. Multiple mesoscale convective systems propagated eastward along the frontal boundary during the period. A sample radar image is shown in Fig. 1. Superimposed on the radar image are 500-hPa geopotential height contours from the National Centers for Environmental Prediction (NCEP) reanalysis (Kalnay et al. 1996). An animation of the time evolution of the radar reflectivity and geopotential height is available online (<http://csrcp.tamu.edu/hiaper/archive/Taylor/radar.mov>) (54 MB). The geopotential height is linearly interpolated in time to the radar analysis times from the 6-hourly NCEP reanalysis.

The time mean of the radar reflectivity, shown in Fig. 2, illustrates the spatial heterogeneity of the data during the study period. The highest mean reflectivities occur in the center of the region, while the southern part has few or no echoes. Minor artifacts (circular radar footprints) are visible from the procedure that merges the individual radars into the gridded mosaic. During the period of study, the synoptic-scale flow is predominantly westerly, with only weak wave disturbances. The 500-hPa height field in Fig. 1 is typical. The time-mean zonal and meridional wind components for the study period, averaged over the study area, are plotted in Fig. 3 as a function of pressure. Winds are from the NCEP reanalysis, which has $2.5^\circ \times 2.5^\circ$ longitude–latitude resolution. The average directional wind shear is very small during this period, with the flow at levels above the surface generally moving slightly north of east. The

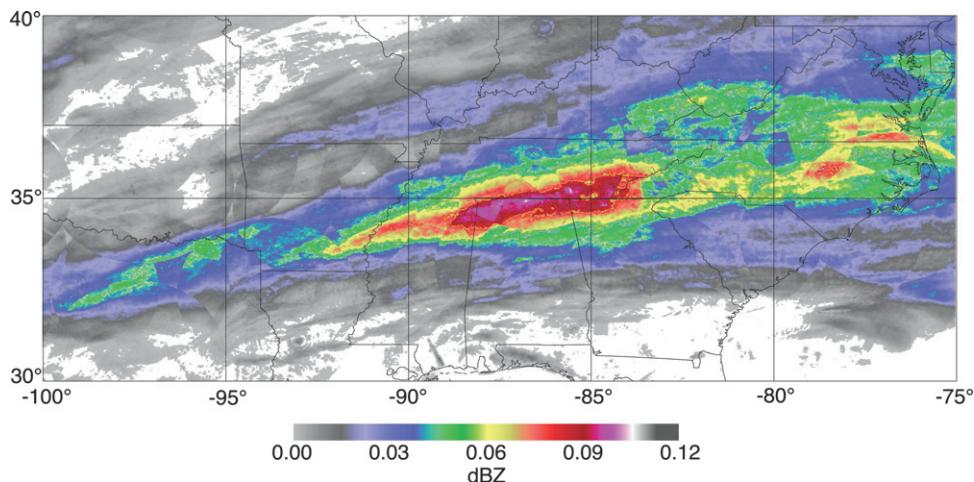


FIG. 2. Time-mean composite radar reflectivity map for 2 May 2002 to 5 May 2005.

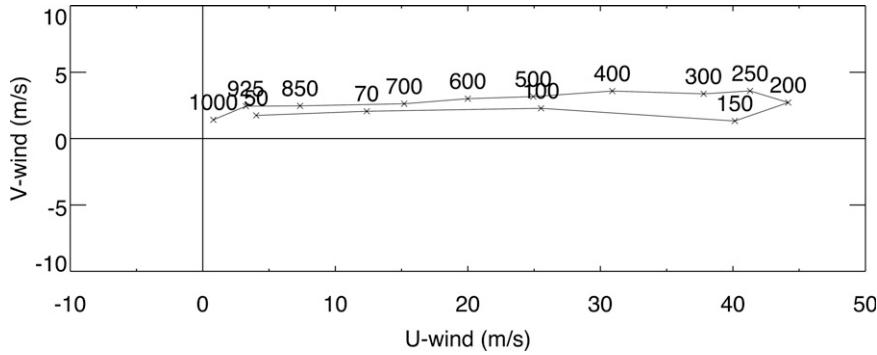


FIG. 3. Time-mean area-mean wind velocity as a function of pressure from the NCEP reanalysis. Numerical labels indicate pressure levels in hPa.

mean wind speed increases from the surface up to a maximum of $\sim 42 \text{ m s}^{-1}$ near the tropopause at 200 hPa and then decreases at higher levels in the stratosphere. From this we can anticipate that convection and the observed radar echoes will propagate approximately eastward with speeds typical of the midtropospheric flow speed of approximately 20 to 30 m s^{-1} .

3. Methods

a. Hypothesis testing

Our method for testing the Taylor hypothesis is developed based on the asymptotic joint normality of sample space–time covariance estimators derived by Li et al. (2008). Assume $Z(\mathbf{x}, t)$ is a strictly stationary space–time random field with covariance function $C(\mathbf{r}, \tau) = \text{Cov}[Z(\mathbf{x}, t), Z(\mathbf{x} + \mathbf{r}, t + \tau)]$, where \mathbf{r} and τ denote an arbitrary spatial lag and time lag, respectively. Let Λ be a set of space–time lags such as $\Lambda = [(\mathbf{r}_1, \tau_1), \dots, (\mathbf{r}_m, \tau_m)]$, where m denotes the number of its elements. Let $\hat{C}(\mathbf{r}, \tau)$ denote an estimator of $C(\mathbf{r}, \tau)$. For simplicity, we choose $\hat{C}(\mathbf{r}, \tau)$ as the moment estimator defined by $\hat{C}(\mathbf{r}, \tau) = (1/N)\sum_{\mathbf{x}}\sum_t [Z(\mathbf{x}, t) - \bar{Z}][Z(\mathbf{x} + \mathbf{r}, t + \tau) - \bar{Z}]$,

where \bar{Z} denotes the mean of $Z(\mathbf{x}, t)$ and N is the total number of summands. This choice of estimator works well in testing properties of the covariance function (see Li et al. 2007). Let $\mathbf{G} = [C(\mathbf{r}, \tau), (\mathbf{r}, \tau) \in \Lambda]$, and let $\hat{\mathbf{G}} = [\hat{C}(\mathbf{r}, \tau), (\mathbf{r}, \tau) \in \Lambda]$ denote the estimator of \mathbf{G} . Li et al. (2008) derived that the appropriately standardized and centered $\hat{\mathbf{G}}$ has an asymptotic multivariate normal distribution in a variety of space–time contexts.

b. Test with a given v

We write Taylor’s hypothesis as

$$H_0: C(\mathbf{0}, \tau) - C(\mathbf{v}\tau, 0) = 0, \quad \text{for any } \tau. \quad (2)$$

Observe that H_0 is a contrast of covariances and thus can be rewritten in the form of $\mathbf{A}\mathbf{G} = \mathbf{0}$, where \mathbf{A} is a contrast matrix of row rank q , say.

For example, if

$$\Lambda = [(\mathbf{0}, \tau_1), (\mathbf{0}, \tau_2), (\mathbf{0}, \tau_3), (\mathbf{v}\tau_1, 0), (\mathbf{v}\tau_2, 0), (\mathbf{v}\tau_3, 0)], \quad (3)$$

then

$$\mathbf{G} = [C(\mathbf{0}, \tau_1), C(\mathbf{0}, \tau_2), C(\mathbf{0}, \tau_3), C(\mathbf{v}\tau_1, 0), C(\mathbf{v}\tau_2, 0), C(\mathbf{v}\tau_3, 0)]^T. \quad (4)$$

Define

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & -1 \end{bmatrix}, \quad (5)$$

then we have $\mathbf{A}\mathbf{G} = \mathbf{0}$ under the null hypothesis. Replacing $C(\cdot)$ with the estimator $\hat{C}(\cdot)$ in (2), we obtain a contrast vector for testing H_0 as the estimated left-hand side of (2), $\mathcal{C} = \hat{C}(\mathbf{0}, \tau) - \hat{C}(\mathbf{v}\tau, 0)$. Apparently, \mathcal{C} can be rewritten into the form of $\mathbf{A}\hat{\mathbf{G}}$. We form the test statistic

(TS) based on the contrasts of $\hat{\mathbf{G}}$ and obtain the distribution of TS under the null hypothesis (Li et al. 2008) as

$$\text{TS} = a_N(\mathbf{A}\hat{\mathbf{G}})^T(\mathbf{A}\hat{\Sigma}\mathbf{A}^T)^{-1}(\mathbf{A}\hat{\mathbf{G}}) \longrightarrow \chi_q^2, \quad (6)$$

in distribution as $N \rightarrow \infty$, for a matrix \mathbf{A} with row rank q , and an appropriate sequence of normalizing constants a_N . We follow Li et al. (2008) and estimate $\hat{\Sigma}$ using subsampling techniques. The choice of subblock size is described in Carlstein (1986). In terms of our

precipitation data, the rich temporal replicates allow us to consider the asymptotics in time dimension and form overlapping subblocks using a moving subblock window along time. Specifically, let N be the total number of time steps, $a_N = \sqrt{N}$ in (6), and the optimal block length for subblocks is $[2\gamma/(1-\gamma^2)]^{2/3}(3N/2)^{1/3}$, where γ can be estimated by $\hat{\gamma} = \hat{C}(\mathbf{0}, 1)/\hat{C}(\mathbf{0}, 0)$. Covariance estimates obtained from each subblock constitute the sample to estimate Σ . The statistical significance of the resulting test statistics can be assessed based on the large sample χ^2 distribution of the test statistic.

4. Models of the autocovariance function for precipitation

a. Isotropic case

To provide some physical insight into the covariance function for precipitation, we develop a simple mathematical model of an evolving precipitation field following Cahalan et al. (1982), North and Nakamoto (1989), Bell et al. (1990), and Bell and Kundu (1996). The model includes propagation (advection), damping, diffusion, and a white-noise stochastic forcing. The evolution of the dependent variable $R(\mathbf{x}, t)$, which represents rain rate, is given by the equation

$$\frac{\partial R}{\partial t} + \mathbf{v} \cdot \nabla R - D\nabla^2 R = -bR + f(\mathbf{x}, t), \quad (7)$$

where \mathbf{v} is a constant advective velocity, D is the diffusion coefficient, b is the damping rate, and f is a stochastic forcing term. The analytical solution to this model can be used directly to compute the covariance function.

The equation is solved in the spectral domain (\mathbf{k}, ω) , where $\mathbf{k} = (k_x, k_y)$ is the spatial wavenumber and ω is the temporal frequency. The random forcing $f(\mathbf{x}, t)$ is a

stationary Gaussian random variable that is white in both space and time; that is,

$$\langle f(\mathbf{x}, t) \rangle = 0 \quad \text{and} \quad \langle f(\mathbf{x}, t)f(\mathbf{x} + \mathbf{r}, t + \tau) \rangle = f_0^2 \delta(\mathbf{r}, \tau), \quad (8)$$

where angle brackets indicate the ensemble mean.

If we define the Fourier transform of $R(\mathbf{x}, t)$ as

$$\tilde{R}(\mathbf{k}, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R(\mathbf{x}, t) e^{-i(\mathbf{k}\cdot\mathbf{x} + \omega t)} d^2\mathbf{x} dt, \quad (9)$$

and take the Fourier transform of (7), we get

$$\frac{1}{2\pi}(i\omega + i\mathbf{k} \cdot \mathbf{v} + D|\mathbf{k}|^2 + b)\tilde{R} = \tilde{f}, \quad (10)$$

where $\tilde{f} = f_0$ is a constant that specifies the magnitude of the white-noise forcing. Solving for \tilde{R} yields

$$\tilde{R}(\mathbf{k}, \omega) = \frac{2\pi\tilde{f}}{(i\omega + i\mathbf{k} \cdot \mathbf{v} + D|\mathbf{k}|^2 + b)}. \quad (11)$$

The variance or power spectrum density S is the magnitude of the complex solution

$$S(\mathbf{k}, \omega) = |\tilde{R} \cdot \tilde{R}^*| = \frac{4\pi^2|\tilde{f}|^2}{(\omega + \mathbf{k} \cdot \mathbf{v})^2 + (D|\mathbf{k}|^2 + b)^2}, \quad (12)$$

where \tilde{R}^* indicates the complex conjugate. Note that the variance in (12) is positive for all (\mathbf{k}, ω) and hence is consistent with the definition of a valid covariance function (Gneiting et al. 2007). Finally, the covariance function in the physical domain is obtained by taking the inverse Fourier transform of S :

$$\begin{aligned} C(\mathbf{r}, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\mathbf{k}, \omega) e^{i(\mathbf{k}\cdot\mathbf{r} + \omega\tau)} d^2\mathbf{k} d\omega \\ &= 2\pi \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{|\tilde{f}|^2}{(\omega + \mathbf{k} \cdot \mathbf{v})^2 + (D|\mathbf{k}|^2 + b)^2} e^{i(\mathbf{k}\cdot\mathbf{r} + \omega\tau)} d^2\mathbf{k} d\omega, \end{aligned} \quad (13)$$

where $\mathbf{r} = (r_x, r_y)$ and τ are appropriate lags in space and time. We solve the ω integral by contour integration by first making a substitution $\omega' = \omega + \mathbf{k} \cdot \mathbf{v}$ and then dropping the prime to simplify the notation. The final integral is obtained as

$$C(r_x, r_y, \tau) = \pi|\tilde{f}|^2 e^{-b\tau} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{e^{i[k_x(r_x - u|\tau|) + k_y(r_y - v|\tau|)]}}{(D|\mathbf{k}|^2 + b)} e^{-D|\mathbf{k}|^2|\tau|} dk_x dk_y. \quad (14)$$

We first explore the above integral analytically by expanding the complex exponential in (14). We then test the validity of TH by examining whether $C(0, 0, \tau) = C(u|\tau|, v|\tau|, 0)$ for a specific $\mathbf{v} = (u, v)$ and τ . Evaluating each term separately yields

$$C(0, 0, \tau) = \pi|\tilde{f}|^2 e^{-b|\tau|} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\cos(k_x u|\tau| + k_y v|\tau|)}{(D|\mathbf{k}|^2 + b)} e^{-D|\mathbf{k}|^2|\tau|} dk_x dk_y \tag{15}$$

and

$$C(r_x = u|\tau|, r_y = v|\tau|, 0) = \pi|\tilde{f}|^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\cos(k_x u|\tau| + k_y v|\tau|)}{(D|\mathbf{k}|^2 + b)} dk_x dk_y. \tag{16}$$

The TH holds if (15) and (16) are equal, for instance,

$$\pi|\tilde{f}|^2 e^{-b|\tau|} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\cos(k_x u|\tau| + k_y v|\tau|)}{(D|\mathbf{k}|^2 + b)} e^{-D|\mathbf{k}|^2|\tau|} dk_x dk_y \stackrel{?}{=} \pi|\tilde{f}|^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\cos(k_x u|\tau| + k_y v|\tau|)}{(D|\mathbf{k}|^2 + b)} dk_x dk_y. \tag{17}$$

The Taylor ‘‘frozen-field’’ hypothesis requires that the turbulent structure of the field being advected (rain in our case) evolves slowly compared to the advective time scale τ . In the model given in (7), the evolution of the rain field is controlled by the diffusion operator and

the damping. We investigate the covariance function in the limit of negligible diffusion ($D \rightarrow 0$). In this limit, the evolution of the rainfall field is controlled entirely by the damping and the stochastic forcing. Taking the limit and canceling like terms, we see that

$$e^{b|\tau|} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cos[k_x u|\tau| + k_y v|\tau|] dk_x dk_y \neq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cos[k_x u|\tau| + k_y v|\tau|] dk_x dk_y. \tag{18}$$

These two expressions are identical, except for the factor of $e^{-b|\tau|}$ on the left-hand side; that is, the TH is not satisfied for this model because the damping changes the *magnitude* of the correlation field, even if the *shape* does not change as the correlation is advected downwind. Thus, for a model like this it is not possible to satisfy TH, and the space and time covariances will differ in proportion to the exponential of the ratio of the temporal lag τ and the damping time scale $1/b$.

b. Extension to anisotropic case

The degree of anisotropy is controlled by the magnitude of the diffusion term, which in the general case has directional dependence and consists of four components:

$$\mathbf{D} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix}. \tag{19}$$

For simplicity, we assume that the diffusion or stretching is only along the two major axes, that is, $D_{xy} = D_{yx} = 0$. Thus in this case (7) can be rewritten as

$$\frac{\partial R}{\partial t} + \mathbf{v} \cdot \nabla R - D_{xx} \frac{\partial^2 R}{\partial x^2} - D_{yy} \frac{\partial^2 R}{\partial y^2} = -bR + f(\mathbf{x}, t). \tag{20}$$

We simplify the above equation by dividing throughout by b and scale the resulting equation with the following nondimensional variables:

$$t' = bt, u' = \left(u\sqrt{\frac{b}{D_{xx}}}\right), v' = \left(v\sqrt{\frac{b}{D_{yy}}}\right), x' = \left(\frac{x}{\sqrt{\frac{D_{xx}}{b}}}\right) \text{ and } y' = \left(\frac{y}{\sqrt{\frac{D_{yy}}{b}}}\right) \text{ to obtain}$$

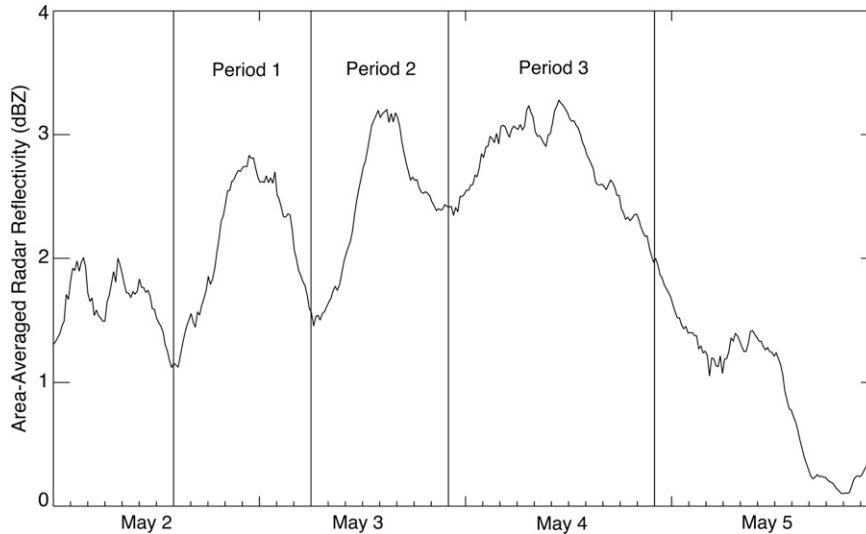


FIG. 4. Time series of area-mean radar reflectivity, 2–5 May 2002. The subperiods selected for the study are (i) 1400 UTC 2 May–0600 UTC 3 May, (ii) 0600–2200 UTC 3 May, and (iii) 2200 UTC 3 May–2200 UTC 4 May. Note here that we did not test TH for 5 May because there was little or no rain throughout the domain.

$$\frac{\partial R}{\partial t'} + \mathbf{v}' \cdot \nabla' R - \frac{\partial^2 R}{\partial x'^2} - \frac{\partial^2 R}{\partial y'^2} = -R + f'(\mathbf{x}, t). \quad (21)$$

Following an analysis similar to the one outlined in the previous section, we obtain the covariance function for the anisotropic case in terms of nondimensional parameters as

$$C(r'_x, r'_y, \tau') = \pi |\tilde{f}'|^2 e^{-\tau'} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{e^{i[k'_x(r'_x - u'|\tau') + k'_y(r'_y - v'|\tau')]} e^{-|\mathbf{k}'|^2 |\tau'|}}{(|\mathbf{k}'|^2 + 1)} dk'_x dk'_y, \quad (22)$$

where $r'_x = r_x / \sqrt{D_{xx}/b}$, $r'_y = r_y / \sqrt{D_{yy}/b}$, and $\tau' = b\tau$. The form of the above covariance function closely resembles that of (14); not surprisingly, the TH does not hold for this case either.

5. Analysis of radar data

a. Covariance calculations

The time series of instantaneous area-averaged reflectivity for the study region is plotted in Fig. 4. Several periods of heavier rain are apparent, and a period of little rain can be seen on 5 May. The analysis methods described above are applied to the entire 4-day period and to the three subintervals of heavier rain indicated by vertical lines and labeled periods 1 through 3 in Fig. 4. For each of the periods of analysis, the time mean for that period at each point

$$\overline{\text{dBZ}}(\mathbf{x}) = \frac{1}{N} \sum_{t=1}^N \text{dBZ}(\mathbf{x}, t) \quad (23)$$

is removed, and all calculations are done with reflectivity anomalies $\text{dBZ}' = \text{dBZ} - \overline{\text{dBZ}}$.

The time-lagged covariance between the reflectivity anomaly $\text{dBZ}'_i(t)$ at a reference point (designated by the subscript i) and the anomaly at a test point (designated j) for a given time lag τ is computed using

$$C(\mathbf{r}_{ij}, \tau) = \begin{cases} \frac{1}{N} \sum_{t=1}^{N-\tau} \text{dBZ}'_i(t) \text{dBZ}'_j(t + \tau), & \tau \geq 0 \\ \frac{1}{N} \sum_{t=1-\tau}^N \text{dBZ}'_i(t) \text{dBZ}'_j(t + \tau), & \tau < 0, \end{cases} \quad (24)$$

where the space lag $\mathbf{r}_{ij} = \mathbf{x}_j - \mathbf{x}_i$ is the vector from point i to point j , and N is the number of observations in the time series. For a given reference point i , $C(\mathbf{r}_{ij}, \tau_n)$ is calculated for all test points j on the data grid within a $2^\circ \times 2^\circ$ rectangular region centered on the reference point (107×113 grid points) and additionally for all

time lags (τ_n) from -3 to $+3$ h. With a time step size of 15 min, this amounts to discrete time lag indices (n) between ± 12 .

To estimate the average covariance function for the entire domain $\hat{C}(\mathbf{r}, \tau)$, we start off by selecting 2000 random but uniformly distributed reference points throughout the entire domain. The number 2000 is somewhat arbitrary, but it accurately represents the dataset spatially. Then, for each reference point, we prescribe the $2^\circ \times 2^\circ$ rectangular region around it and calculate the mean of this region. If this temporally and spatially averaged value is greater than 4 dBZ, then we treat the point as a valid reference location; otherwise it is rejected. This omits from consideration any reference points where little or no rain fell during the period of analysis. By repeating this process for all 2000 reference points, we end up with around 300 *valid* reference locations. Then the calculation in (24) is repeated for these 300 locations and averaged over all locations. Reference points are chosen to be at least 1° away from the boundaries of the domain to avoid edge effects.

Correlations \hat{c} are estimated using

$$\hat{c} = \frac{\hat{C}}{\hat{\sigma}^2}, \quad (25)$$

where $\hat{\sigma}^2$ denotes the estimated variance of the entire precipitation anomaly field. Error bars for \hat{C} and \hat{c} are roughly estimated by computing the standard deviation of the individual estimates of C or c for the 300 reference points. These correlation estimates based on \hat{c} are then tested for the validity of the TH and are subsequently compared with those from the method described in section 3, which takes into account the correlations between the individual covariance estimates.

b. Structure of the covariance field

The mean space–time covariance function $\hat{C}(\mathbf{r}, \tau)$ and correlation $\hat{c}(\mathbf{r}, \tau)$ are estimated for the entire 4-day study as described in section 5a. The spatial structure of the empirical space–time correlation \hat{c} estimated over all pairs that have the same spatial and temporal lag is shown in Figs. 5, 6. Figure 5 shows contours of \hat{c} as a function of spatial lag at $\tau = 0$, while Fig. 6 shows \hat{c} for $\tau = 0, 15, 30,$ and 45 min. Figure 5 reveals that the \hat{c} has an approximately elliptical shape with the major axis oriented somewhat north of east. The elliptical shape (anisotropy) of \hat{c} indicates the mean orientation of the precipitation areas during this period on the scale shown. As in the theoretical model, the peak observed at $\tau = 0$ (Fig. 6a) decays with time as it moves upwind (downwind) with decreasing (increasing) lag (Figs. 6b–d).

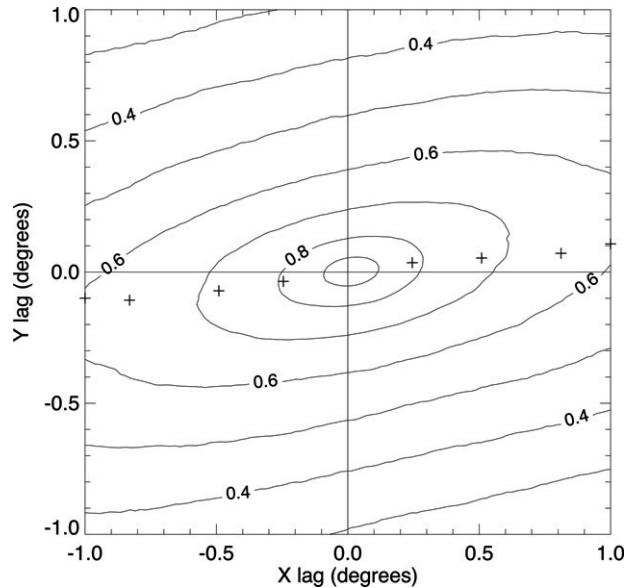


FIG. 5. Averaged space–time correlation (\hat{c}) for $\tau = 0$. The plus signs indicate the positions of $\max[\hat{c}(\mathbf{r}, \tau)]$ at $\tau = \pm 15, 30, 45,$ and 60 min, respectively.

To qualitatively compare the general shape and structure of \hat{c} with that of the anisotropic model (20) described in section 4b, we also evaluate the model correlation function (22) numerically for different values of the parameters D, b, u, v, \mathbf{r} , and τ . The goal is not to exactly reproduce the observed values, but to understand the nature of correlation functions with similar shape. Figure 7 illustrates the shape of (22) for $b = 0.25 \text{ min}^{-1}$, $u = 1.6 \times 10^{-4} \text{ deg min}^{-1}$, $v = 2.5 \times 10^{-5} \text{ deg min}^{-1}$, $D_{xx} = 0.20 \text{ deg}^2 \text{ min}^{-1}$ and $D_{yy} = 0.075 \text{ deg}^2 \text{ min}^{-1}$ for spatial lags (r_x, r_y) that range from -1.0° to $+1.0^\circ$ and temporal lags $\tau = 0, 15, 30, 45$ min, respectively. While solving (14) numerically, the limits of the integration were truncated to get a finite correlation function. The shape of the correlation function is similar to the observations. For fields that have this type of covariance structure, with a localized peak in the covariance that decays as it is advected downstream, we generally would not expect the field to satisfy the Taylor hypothesis.

In Fig. 5 the maximum correlation at $\tau = 0$ occurs at the origin, as expected. As the magnitude of τ increases from zero in the positive or negative direction, the peak of the correlation function shifts upstream or downstream, respectively, depending on the sign of the lag; the maximum correlation values decrease with increasing lag, as shown in Fig. 6. The plus symbols in Fig. 5 represent the locations of the peak correlations at time lags of $\pm 15, 30, 45,$ and 60 min. The plus signs are approximately colinear and equally spaced, indicating that

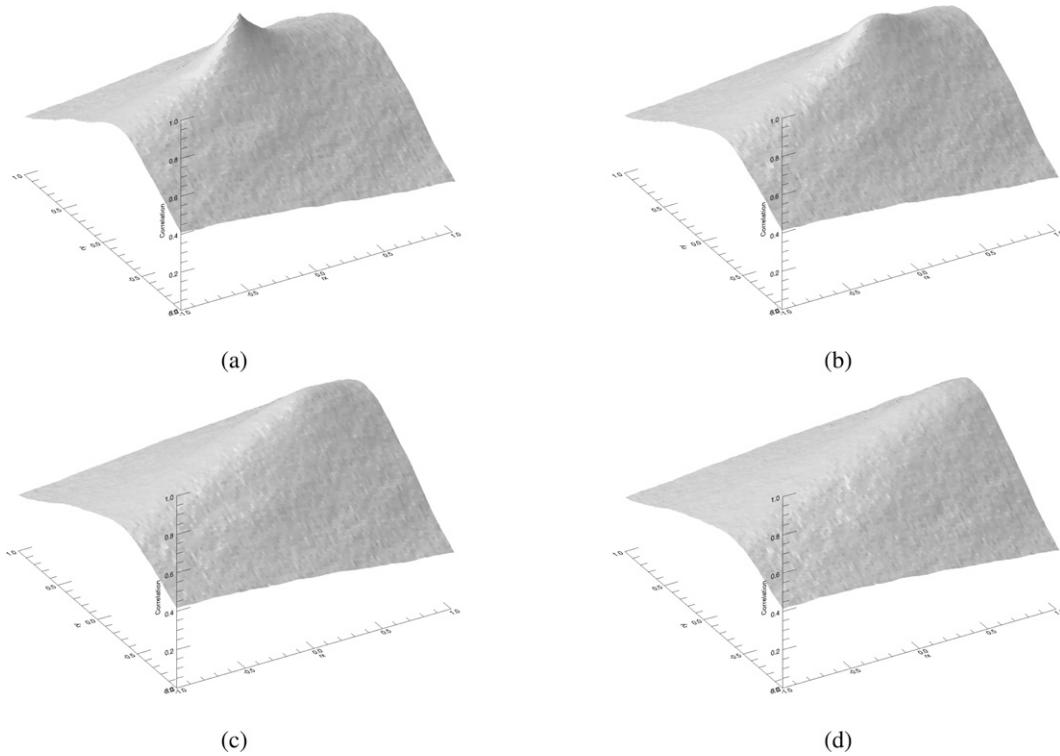


FIG. 6. Averaged space-time correlation fields \hat{c} at (a) $\tau = 0$, (b) $\tau = 15$, (c) $\tau = 30$, and (d) $\tau = 45$ min. Isoleths of \hat{c} are approximately elliptical in shape and oriented somewhat north of east (see Fig. 5). The field decays in both space and time as it translates with the wind velocity.

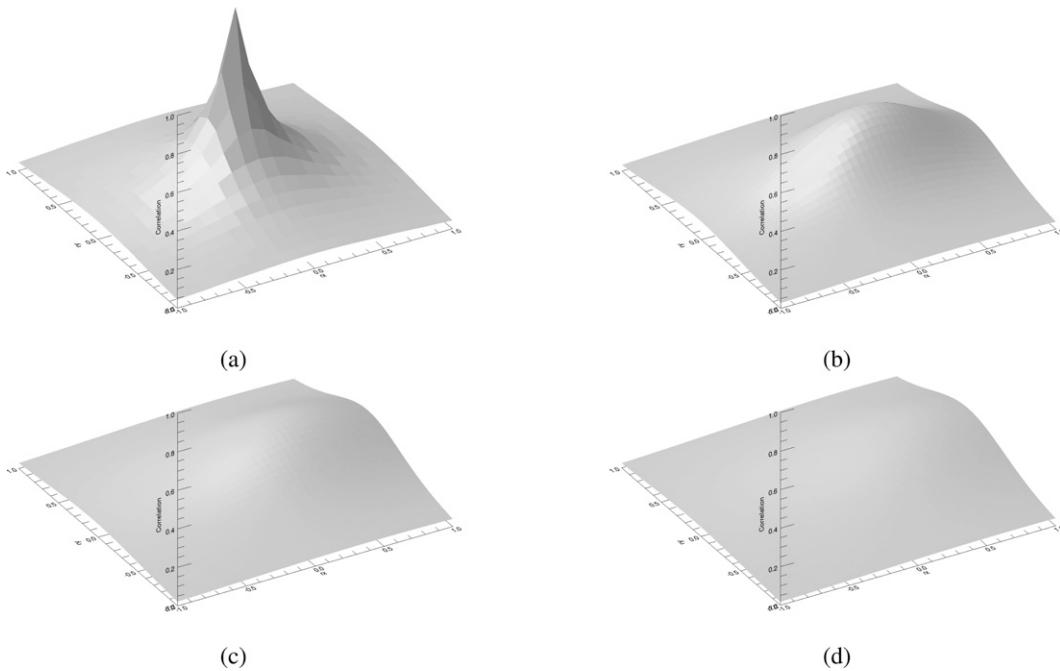


FIG. 7. Correlation fields (cor) for precipitation model (7) at (a) $\tau = 0$, (b) $\tau = 15$, (c) $\tau = 30$, and (d) $\tau = 45$ min, respectively. It is assumed here that $u = 0.22641^\circ \text{ day}^{-1}$, $v = 0.0357^\circ \text{ day}^{-1}$, $b = 0.25 \text{ min}^{-1}$, $D_{xx} = 0.20 \text{ deg}^2 \text{ min}^{-1}$, and $D_{yy} = 0.075 \text{ deg}^2 \text{ min}^{-1}$. The correlation decays as it is advected by the wind in the r_x lag direction.

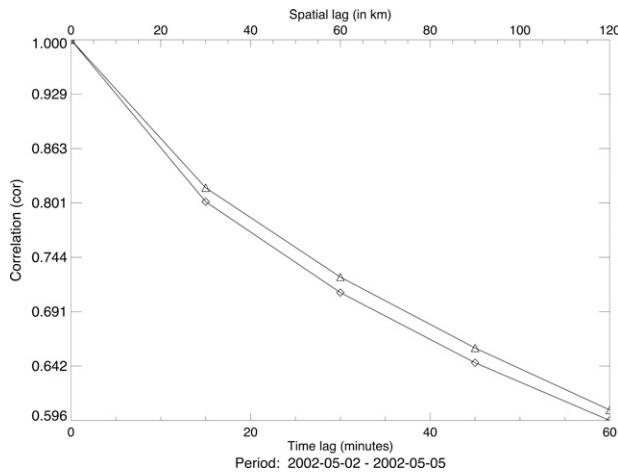


FIG. 8. Averaged space–time correlation (\hat{c}) as a function of space and time lags for the entire 4-day period of study. The triangles correspond to $\hat{c}(\mathbf{v}\tau, 0)$ and the diamonds correspond to $\hat{c}(\mathbf{0}, \tau)$ for $\tau = 0, 15, 30, 45, 60$ min, respectively. The results indicate that the TH does not hold, even for short time intervals.

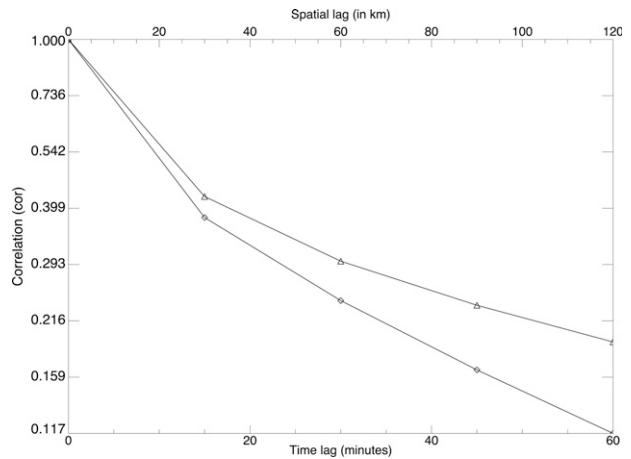


FIG. 9. Space–time correlation (cor) for the precipitation model as a function of time lag. The triangles correspond to $\text{cor}(\mathbf{v}\tau, 0)$ and the diamonds correspond to $\text{cor}(\mathbf{0}, \tau)$ for $\tau = 0, 15, 30, 45, 60$ min, respectively. These curves serve to reinforce the analytical results described in the paper that the TH does not hold for this model.

the advective velocity is approximately independent of lag. The average velocity vector \mathbf{v} can be estimated by using the vectors from the origin to the plus signs, divided by the corresponding time lag τ . The mean velocities are found to have magnitudes between 25 and 30 m s^{-1} ($\sim 25.6^\circ \text{ day}^{-1}$) oriented $\sim 83^\circ$ from north. This is consistent with the mean wind plotted in Fig. 3 and with the motion of precipitation features visible in an animation of the radar reflectivity maps. We use this value of $\mathbf{v} = 25.6^\circ \text{ day}^{-1}$ (83° east of north) to compute $\hat{c}(\mathbf{v}\tau, 0)$ for the entire 4-day period of study. Due to some variability of \mathbf{v} with time, we use individual velocity estimates obtained from the correlation functions for each subperiod when testing the TH for the three subperiods. Note that the propagation velocity \mathbf{v} and the principal axis of the correlation function c are not oriented in the same direction, nor is there any reason to expect them to be.

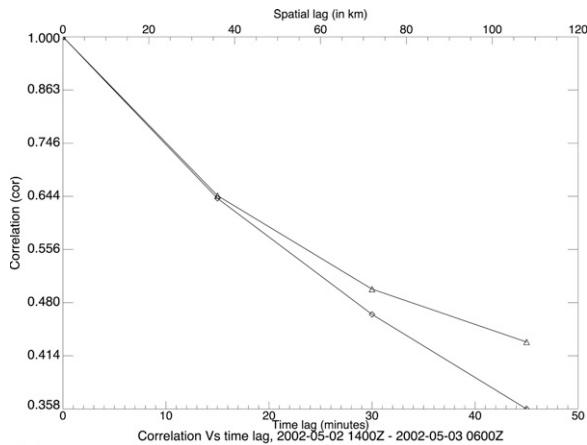
c. Testing the Taylor hypothesis

Figure 8 shows the space-lagged correlations [$\hat{c}(\mathbf{v}\tau, 0)$, triangles] and time-lagged correlations [$\hat{c}(\mathbf{0}, \tau)$, diamonds] as a function of time lag for the entire 4-day period plotted on a logarithmic scale. These curves show that the TH in general does not hold for the 4-day time period and hence for large space and time scales. It is important to mention here that although the curves do appear close (especially at smaller time lags), the error bars (not shown) around the mean correlations corresponding to the shortest time lag of 15 min are small and do not overlap, thereby indicating that the curves are indeed statistically different. Figure 9 shows

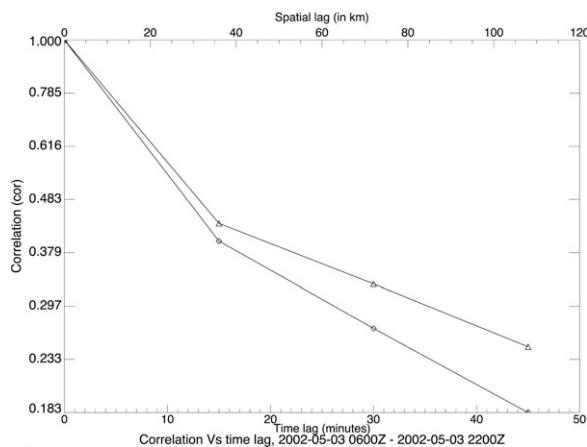
curves of correlation plotted as a function of time lag for the precipitation model (20). The correlation curves in Fig. 9 are similar to those in Fig. 8, with the rate of decay being controlled by the magnitude of the damping term b .

In addition, we test the TH for each of the smaller time periods illustrated in Fig. 4, during which the heaviest rainfall occurs. Figure 10a demonstrates a subperiod for which the TH held up to 15 min (35 km), while Figs. 10b,c show subperiods for which it did not hold for even short space and time scales.

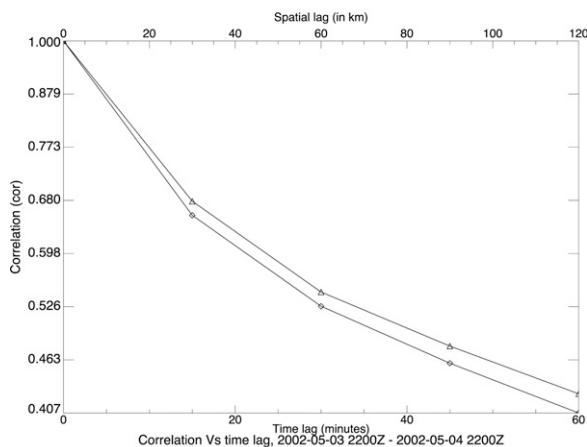
To assess the TH using both methods, a table containing probability values (p values) pertaining to the statistical significance of the difference between the averaged correlations, that is, $\hat{c}(\mathbf{0}, \tau) - \hat{c}(\mathbf{v}\tau, 0)$, is constructed for each testing period at various time lags using both a standard Student’s t test and the method from section 3. These tables serve to reinforce the results illustrated in the Figs. 8, 10a–c in that the TH does not hold for such systems. Table 1 lists p values generated using the Student’s t test, while Table 2 lists those generated using the method of Li et al. (2007). The convention used here is that a p value $p^* \geq 0.05$ implies that the difference is not significant at the 5% level, and the TH cannot be rejected, while $p^* < 0.05$ indicates that the difference is significant at the 5% level and the TH is rejected. The results from both statistical methods show that the TH does not hold for the full 4-day period, but it is admitted for up to 15 min (35 km) for one subperiod (period 1), as indicated by large p^* values (in boldface). The TH is not rejected in only one of the 16 period-and-lag combinations tested. Although the results from the two methods agree in all cases, it is



(a)



(b)



(c)

FIG. 10. Averaged space–time correlation (\hat{c}) as a function of time and space lags for time period (a) 1400 UTC 2 May–0600 UTC 3 May, (b) 0600–2200 UTC 3 May, and (c) 2200 UTC 3 May–2200 UTC 4 May. (a) Shows that the TH holds for at least 15 min (35 km), while (b) and (c) show that the TH does not hold even for time scales shorter than 15 min. The triangles denote $\hat{c}(\mathbf{v}\tau, 0)$, while the diamonds denote $\hat{c}(\mathbf{0}, \tau)$.

TABLE 1. Significant p values at various τ from Student’s t test. Bold indicates values that are not significant at the 5% level.

Time period	τ			
	15 min	30 min	45 min	60 min
4-day period	5.6×10^{-4}	6.4×10^{-5}	5.1×10^{-6}	1.1×10^{-3}
Time period 1	4.2×10^{-1}	7.7×10^{-3}	2.5×10^{-2}	—
Time period 2	7.8×10^{-3}	3.4×10^{-9}	0.000	—
Time period 3	3.2×10^{-9}	6.8×10^{-7}	3.2×10^{-5}	1.0×10^{-2}

important to note that the results from the Student’s t test (Table 1) tend to overestimate the significance at a particular τ compared to the method of Li et al. (2007; Table 2). This can be attributed to the inability of the correlation method to account for the correlation between the correlation estimates or, equivalently, to overestimating the number of degrees of freedom. Using the standard NEXRAD Z – R relationship ($Z = 300R^{1.4}$), space–time correlations were also computed using rain rates to gauge whether the choice of variable affects the conclusions. The statistical significance of the results has the same pattern as in Table 1. In only one case is the TH not rejected (period 1 at 15 min), from which we conclude that the TH does hold for this period up to 15 min (35 km). In Table 1, p^* values for time periods 1 and 2 have not been shown for $\tau = 60$ min because the velocity estimates for either case resulted in computing locations (and corresponding correlation estimates) that were beyond the $2^\circ \times 2^\circ$ moving window in the spatial domain. The hypothesis-testing approach (Table 2), however, used a slightly larger moving window and thus reported correlations (and hence p^* values) at $\tau = 60$ min as well.

6. Summary and discussion

Taylor’s hypothesis provides a simple model of the covariance function for fluid variables in a uniform flow. It also implies a relationship between the time scale of small-scale variations in the fluid compared to the advective time scale of the mean flow. This paper compares two approaches to testing the validity of the TH for a geophysical fluid flow by using radar obser-

TABLE 2. Significant p values at various τ from hypothesis testing in section 3b. Bold indicates values that are not significant at the 5% level.

Time period	τ			
	15 min	30 min	45 min	60 min
4-day period	5.0×10^{-3}	1.9×10^{-2}	4.7×10^{-2}	2.0×10^{-3}
Time period 1	5.8×10^{-1}	1.0×10^{-3}	0.000	0.000
Time period 2	2.3×10^{-2}	0.000	0.000	0.000
Time period 3	1.1×10^{-4}	7.2×10^{-5}	2.0×10^{-4}	3.7×10^{-4}

vations of rainfall. The first is based on a statistically rigorous procedure (Li et al. 2007), while the second is based on the assumption of independence of the covariance estimates, which demonstrably does not hold in this case. The first method does not require any assumptions about the data distribution and tests the null hypothesis, given by $H_0: \hat{C}(\mathbf{0}, \tau) - \hat{C}(\mathbf{v}\tau, 0) = 0$ for the mean advection velocity \mathbf{v} and time lag τ . The results indicate that both methods agree well with the analytical model described in section 4a in that the TH does not hold for fields characterized by advection, diffusion and decay. The TH does appear to hold in one case out of 16 (period 1 for a lag of 15 min), but testing at the 5% level, we would expect this in one case out of 20. Nevertheless, there is the possibility that the TH might hold for shorter spatial and temporal scales than what is resolved by the data (4 km and 15 min). The simple Student's t test tends to overestimate the significance of the difference between correlation estimates by not accounting for the correlation between those estimates. This is reflected by the fact that the p values from the Student's t test are considerably smaller than those obtained by the hypothesis testing procedure in section 3a.

The failure of the Taylor hypothesis for the data analyzed here could be due to several factors. First, the background flow velocity \mathbf{v} may not be constant in space or time (Poveda and Zuluaga 2005), although in this case the flow is relatively steady during the study period. Second, the observed variable may evolve with a time scale shorter than the advection time scale (Waymire et al. 1984). In either case, the results in this paper raise questions about the validity of the TH for radar rainfall data as reported previously by Zawadzki (1973) and Poveda and Zuluaga (2005), at least for the space and time scales resolved by the data used here.

Acknowledgments. The authors wish to thank David Ahijevych for assistance with accessing the gridded NEXRAD data. This research was supported in part by a CMG National Science Foundation Grant ATM-0620624.

REFERENCES

- Bell, T. L., and P. K. Kundu, 1996: A study of the sampling error in satellite rainfall estimates using optimal averaging of data and a stochastic method. *J. Climate*, **9**, 1251–1268.
- , A. Abdullah, R. L. Martin, and G. R. North, 1990: Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space-time stochastic model. *J. Geophys. Res.*, **95D**, 2195–2205.
- Burghlelea, T., E. Serge, and V. Steinberg, 2005: Validity of the Taylor hypothesis in a random spatially smooth flow. *Phys. Fluids*, **17**, 103101, doi:10.1063/1.2077367.
- Cahalan, R. F., D. A. Short, and G. R. North, 1982: Cloud fluctuation statistics. *Mon. Wea. Rev.*, **110**, 26–43.
- Carlstein, E., 1986: The use of subsample values for estimating the variance of a general statistic from a stationary sequence. *Ann. Stat.*, **14**, 1171–1179.
- Cox, D. R., and V. Isham, 1988: A simple spatial-temporal model of rainfall. *Proc. Roy. Soc. London*, **A415**, 317–328.
- Crane, R. K., 1990: Space-time structure of rain rate fields. *J. Geophys. Res.*, **95**, 2011–2020.
- Gneiting, T., M. G. Genton, and P. Guttorp, 2007: Geostatistical space-time models, stationarity, separability and full symmetry. *Statistics of Spatio-Temporal Systems, Monogr. Stat. Appl. Probab.*, No. 107, Chapman & Hall/CRC Press, 151–175.
- Gupta, V. K., and E. Waymire, 1987: On Taylor's hypothesis and dissipation in rainfall. *J. Geophys. Res.*, **92**, 9657–9660.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Li, B., M. G. Genton, and M. Sherman, 2007: A nonparametric assessment of properties of space-time covariance functions. *J. Amer. Stat. Assoc.*, **102**, 736–744.
- , —, and —, 2008: On the asymptotic joint distribution of sample space-time covariance estimators. *Bernoulli*, **14**, 228–248.
- North, G. R., and S. Nakamoto, 1989: Formalism for comparing rain estimation designs. *J. Atmos. Oceanic Technol.*, **6**, 985–992.
- Poveda, G., and M. D. Zuluaga, 2005: Testing Taylor's hypothesis in Amazonian rainfall fields during the WETAMC/LBA experiment. *Adv. Water Resour.*, **28**, 1230–1239.
- Sun, C.-H., and L.R. Thorne, 1995: Inferring spatial cloud statistics from limited field-of-view, zenith observations. *Proc. Fifth Atmospheric Radiation Measure (ARM) Science Team Meeting*, San Diego, CA, ARM, 331–334. [Available online at http://www.arm.gov/publications/proceedings/conf05/extended_abs/sun_ch.pdf.]
- Taylor, G. I., 1938: The spectrum of turbulence. *Proc. Roy. Soc. London*, **A164**, 476–490.
- Waymire, E., V. K. Gupta, and I. Rodriguez-Iturbe, 1984: A spectral theory of rainfall intensity at the meso- β scale. *Water Resour. Res.*, **20**, 1453–1465.
- Zawadzki, I. I., 1973: Statistical properties of precipitation patterns. *J. Appl. Meteor.*, **12**, 459–472.

Copyright of *Journal of Hydrometeorology* is the property of *American Meteorological Society* and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.