A dynamical–system description of precipitation over the tropics and the midlatitudes

By JUN-ICHI YANO[1]*, MAARTEN H. P. AMBAUM[2], HELEN F. DACRE[2], and AGOSTINO MANZATO[3], [1]*CNRM UMR3589 (CNRS), Météo-France, Toulouse Cedex, France;* [2]*Department of Meteorology, University of Reading, Reading, UK;* [3]*OSMER — Osservatorio Meteorologico Regionale dell'ARPA Friuli Venezia Giulia, Palmanova, Italy*

ABSTRACT

A dynamical–system approach is proposed to describe the relationship between precipitation and a chosen predictor. This is done by constructing a two dimensional phase space spanned by predictor and predictant. This study uses two sounding data sets from the Tropical Western Pacific and Friuli Venezia Giulia (FVG) over North–East Italy as representatives of the tropics and midlatitudes, in addition to a basin–scale average over the winter–period North Atlantic from global re–analysis data. In contrast to conventional correlation-based approaches, the proposed approach depicts periodic cycles, as well as discharge–recharge cycles as its nonlinear extension. Discharge–recharge cycles for tropical convection are identified by using both the convective available potential energy (CAPE) and the column–integrated water vapor (CIW) as predictors, as well as the baroclinicity for the winter–period North–Atlantic rain. On the other hand, the midlatitude rain, as seen over FVG as well as the winter–period North Atlantic, does not constitute a well–defined periodic cycle either with CAPE or CIW as a predictor. The inferred phase–space trajectories are more deterministic at peripheries of dense data areas rather than at a middle in the phase space. Data–dense areas in phase space, where traditional approaches primarily focus, are associated with more prediction uncertainties in our analysis due to more phase-velocity fluctuations.

Keywords: dynamical system, precipitation, column-integrated water vapor, CAPE, empirical forecast, discharge–recharge

1. Introduction

The atmospheric system is complex, and it is often hard to derive a simple system equation theoretically. Thus, it is a common custom to seek a relationship of a predicted variable with another variable based on observational data. Such an approach is often misleadingly called 'statistical'. It would be better called empirical, because there is nothing particularly statistical in these approaches except for a need for a certain statistical test to evaluate a reliability of such an empirically–derived model (see more in: Ambaum 2010).

The most popular approach would be to seek a linear relationship between two variables for inferring a cause and effect between them, for example, by correlation analysis.[1] A slightly more advanced approach along the same line is to take scatter plots between these two variables to identify a nonlinear relationship between the two

variables. Even better still would be to evaluate the density distribution of these two variables over a phase space spanning them. However, those approaches have their limitations. An example is when two variables constitute a variable pair of a harmonic oscillator. The two variables have no correlation but are clearly related.[2]

The present paper proposes an approach, which is designed to work in such a case. A more general, basic idea of the proposed alternative approach is to infer the cause and effect between two variables, rather than directly seeking a functional relationship (linear or nonlinear) between the variables, based on a dynamical–system description: see Eq. (2.2) below and associated discussions.

Precipitation is one of the most important variables to be predicted in atmospheric science and hydrology, and as a result various empirical approaches have been adopted. The atmospheric precipitation system is so complex that even today, numerical weather predictions

1

regularly fail. A simple empirical prediction scheme is also of theoretical interest, because it may also help us to understand the atmospheric precipitation processes better.

It is found that correlations between the precipitation and the convectively available potential energy (CAPE) are rather weak in observational analyses (Sherwood 1999; Manzato 2003; Zhang and Klein 2010; Barkiđija and Fuchs 2013). This is in spite of the fact that CAPE is expected to be a key variable in describing the convective energy cycle (*cf.*, Emanuel et al. 1994; Rennó and Ingersoll, 1996; Yano and Plant 2012a, 2012b), and substantial fractions of precipitation are expected from convection (see also Smith and Montgomery 2012). On the other hand, quantities less directly linked to the convection dynamics, such as the column-integrated water vapor (CIW)[3] present higher correlations with the precipitation (Raymond 1995; Brown and Zhang 1997; Sherwood 1999; Sherwood and Wahrlich 1999; Manzato 2003; Sobel et al. 2004; Mapes et al. 2006, 2009).

The present paper addresses questions concerning empirical prediction methods by focusing on precipitation. Our more direct motivation is a suggestion by Yano and Plant (2012a) that the convective energy cycle may be described by a process of recharge and discharge in the phase space consisting of the cloud work function and the convective kinetic energy. The usefulness of such a phase space was originally suggested by Arakawa and Schubert (1974) in their development of convection parameterization. Although the theory is sound, it is highly idealized. For this reason, we seek observational support so that a realism of this theory can be assessed.

Here, the cloud work function can practically be considered a potential energy of the system. For the purpose of our analysis, it is replaced by CAPE, which is commonly diagnosed from observational analyses. Another dependent variable of the system, the convective kinetic energy, is replaced by the precipitation rate (rain in short hereafter), because both measure the convective intensity, and are expected to be closely related each other. See Yano et al. (2005) for the physical interpretation of the cloud work function as well as its link to CAPE.

We will, more specifically, ask whether a recharge-discharge cycle can be identified in the phase space of these two variables, even though the correlation between them is rather weak. An important point to be emphasized here is that these two facts do not contradict each other: in fact, under an expected recharge-discharge cycle, no strong correlation between the two variables would be found. As an alternative prediction variable for rain, CIW is also examined.

The present paper examines both tropical and midlatitude data. We expect that a contrast in the dynamics between the tropics and the midlatitudes may lead to different results. In general, tropical rain is more dominated by convective processes than those of midlatitudes. Here, as the introduction so far suggests, a motivation of the present analysis stems from the understanding of convective rain, but the analysis is performed on precipitation data without discriminating between convective and non-convective processes.

The focus of the present study is the mesoscale and the longer time scales. This constraint arises from the fact that CAPE and CIW are evaluated from data based on soundings, directly or indirectly, which are typically performed in an interval of 6–12 hours. As a result, the data do not resolve the convective scale. Nevertheless, we emphasize that data are consistent with typical interests of regional forecasts over time scales of 24–48 hours (*cf.*, Sec. 5.4).

The outline of this paper is as follows. The methodology of the present study, based on a dynamical-system description introduced by Novak et al. (2017), is presented in the next section. Data for the analysis is introduced in Sec. 3, and the analysis results are presented in Sec. 4. The results are further discussed in the last section. Some technical details of analysis methods are given separately in the Appendix.

2. Methodology: Dynamical–system based description

2.1. Principle

A limitation of predicting a variable, y, based on its functional dependence on a predictor, x, linear or nonlinear, is probably best seen when two variables are described by a pure harmonic oscillator:

$$\dot{x} = -y, \tag{2.1a}$$
$$\dot{y} = x, \tag{2.1b}$$

where the dot designates the time derivative. In this case, a circular distribution of the two variables around the equilibrium point, $(x, y) = (0, 0)$, is found in the phase space, with a precise distribution depending on an initial distribution of variables. Thus, no linear correlation is seen between the two variables. This example provides a simple demonstration that no correlation does not necessarily mean no relation between two variables.

The discharge–recharge model proposed by Yano and Plant (2012a) and Ambaum and Novak (2014) can be considered a nonlinear generalization of this linear harmonic oscillator. The system (2.1) may, furthermore, be considered a special case of a general dynamical system of a form:

$$\dot{x} = f(x, y), \tag{2.2a}$$
$$\dot{y} = g(x, y). \tag{2.2a}$$

Here, $f(x, y)$ and $g(x, y)$ are (unspecified) tendencies for the variables, x and y, respectively, that are to be

estimated from observations. We refer to this tendency pair, $(\dot{x}, \dot{y})$, as the *phase velocity* in the following. Note that Eq. (2.2) casts a system into a framework of 'cause and effect' with those given by the right–hand and the left–hand sides, respectively.

Practically it is not possible to derive a closed deterministic dynamical system of the form (2.2) for any two variables in the atmospheric system because of its high-dimensionality, in which these two variables are also inevitably coupled with many other variables (*cf.*, Yano and Mukougawa 1992). The formulation given by Eq. (2.2) would best be considered an averaged description of a system, and other variables may project onto this phase space, either as a stochastic contribution, or as a time mean (*cf.*, Pavliotis and Stuart 2007). From a point of view of the data analysis, in other words, an estimate of the tendency functions, $f(x, y)$ and $g(x, y)$, is associated with the fluctuations, *or* noises, which may be designated by ϵ_x and ϵ_y for these two variables. Thus,

$$\dot{x} = f(x, y) + \epsilon_x, \qquad (2.3a)$$

$$\dot{y} = g(x, y) + \epsilon_y. \qquad (2.3b)$$

We estimate $f(x, y)$, $g(x, y)$, ϵ_x, and ϵ_y by the procedure described in next subsection. Here, $f(x, y)$ and $g(x, y)$ are the signals (i.e. mean tendencies) that we wish to determine; ϵ_x and ϵ_y are the noises, or fluctuations around these deterministic signals, whose magnitudes are to be inferred.

2.2. Analysis Procedure

In general, regardless of a specific data set choice, we consider $N+1$ observations consisting of (x_i, y_i) $(i = 1, \ldots, N + 1)$. Here, the index, i, is arranged following a sequence of the observations, which are performed with an interval of Δt. Among the three data sets adopted in the present study (*cf.*, Sec. 3), $\Delta t = 12$ hours for the Italian data set, and the remaining two are with $\Delta t = 6$ hours.

We evaluate the phase velocities by

$$\dot{x}_{i+1/2} = \frac{x_{i+1} - x_i}{\Delta t} \qquad (2.4a)$$

$$\dot{y}_{i+1/2} = \frac{y_{i+1} - y_i}{\Delta t} \qquad (2.4b)$$

as values at the phase-space points

$$x_{i+1/2} = \frac{x_i + x_{i+1}}{2}$$
$$y_{i+1/2} = \frac{y_i + y_{i+1}}{2}$$

with $i = 1, \ldots, N$.

We measure the spread of observation points by

$$s_x = \left[\frac{1}{N}\sum_{i=1}^{N}(x_{i+1/2} - \bar{x})^2\right]^{1/2},$$

$$s_y = \left[\frac{1}{N}\sum_{i=1}^{N}(y_{i+1/2} - \bar{y})^2\right]^{1/2},$$

where

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N}x_{i+1/2},$$

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N}y_{i+1/2}$$

define the average point of data in the phase space.

With the help of these statistics, we estimate a phase velocity (i.e. evolution tendency) at every point, (x, y), in the phase space by a weighted average of the all available phase-velocity estimates from the observations given by Eq. (2.4) as

$$\dot{x}(x, y) = \frac{\sum_{i=1}^{N}\omega_{i+1/2}\dot{x}_{i+1/2}}{\sum_{i=1}^{N}\omega_{i+1/2}} \qquad (2.5a)$$

$$\dot{y}(x, y) = \frac{\sum_{i=1}^{N}\omega_{i+1/2}\dot{y}_{i+1/2}}{\sum_{i=1}^{N}\omega_{i+1/2}} \qquad (2.5b)$$

with a weight defined by

$$\omega_{i+1/2} = \exp\left[-\frac{1}{2}\left\{\left(\frac{x - x_{i+1/2}}{f_x s_x}\right)^2 + \left(\frac{y - y_{i+1/2}}{f_y s_y}\right)^2\right\}\right].$$

Here we adopt weighting-spread factors $f_x = f_y = 0.35$; this definition means that the data are averaged with a Gaussian weight with an influence radius defined by 35% of the data spread. These values of f_x and f_y ensure that sufficient smoothing occurs to produce well-defined phase space trajectories, without removing too much detail (see Novak et al. 2017 for further details).

This procedure has the advantage of being able to estimate the phase velocities over the phase space continuously and everywhere in the phase space regardless of data availability at any specific point. As a result, the phase velocity is estimated at every point, (x, y), in the phase space with an effective number, $a(x, y)$, of data points defined by

$$a(x, y) = \sum_{i=1}^{N}\omega_{i+1/2}. \qquad (2.6)$$

Effectively, $a(x, y)$ measures a data number 'density' used in the present analysis. Here and hereafter, the arguments (x, y) are omitted from most of the expressions for simplicity except for $\dot{x}$ and $\dot{y}$ in the left hand side, for example, in Eqs. (2.5a) and (2.5b).

Note that the phase velocity estimated by Eqs. (2.5a) and (2.5b) are mean values. The phase velocity associated with an actual evolution of the system fluctuates around these mean values, depending on the instant that the system visits a particular phase–space point. We measure these fluctuations of the above phase-velocity estimates by standard deviations:

$$s_u = \left[\frac{\sum_{i=1}^{N} \omega_{i+1/2}(\dot{x}_{i+1/2} - \dot{x})^2}{a} \right]^{1/2}, \qquad (2.7a)$$

$$s_v = \left[\frac{\sum_{i=1}^{N} \omega_{i+1/2}(\dot{y}_{i+1/2} - \dot{y})^2}{a} \right]^{1/2}. \qquad (2.7b)$$

These fluctuations can be treated as noises for quantifying statistical significance (*cf.*, Eq. 2.10 below) as well as for modelling purposes (*cf.*, Sec. 5.4 below). As a result, a nondeterministic part of the dynamical system (2.3) may be represented by $\epsilon_x = \hat{\epsilon}_x s_u$ and $\epsilon_y = \hat{\epsilon}_y s_v$, where $\hat{\epsilon}_x$ and $\hat{\epsilon}_y$ are random noises with standard deviations of unity.

Considering the fact that the scales (as well as units) for x and y are different, for graphical purposes, we normalize the phase-space velocity by

$$(\hat{\dot{x}}, \hat{\dot{y}}) = (\dot{x}/s_x, \dot{y}/s_y), \qquad (2.8)$$

thus on the plotted phase space, the spread of the velocity is also normalized into

$$(\hat{s}_u, \hat{s}_v) = (s_u/s_x, s_v/s_y) \qquad (2.9)$$

(i.e. normalized spread vector).

The statistical significance of the phase–velocity estimate is influenced by two major factors: effective number, $a(x, y)$, of available points defined by (2.6) and the spread (i.e. variance), $\hat{s}_u^2 + \hat{s}_v^2$, of phase–velocity, measuring fluctuations at each point in the phase space. The estimate is more reliable with increasing $a(x, y)$ and less reliable with increasing $\hat{s}_u^2 + \hat{s}_v^2$. Thus, a statistical significance may be measured by:

$$\beta(x, y) = \left[\frac{a}{\hat{s}_u^2 + \hat{s}_v^2} \right]^{1/2}. \qquad (2.10)$$

3. Data

3.1. Data sources

As data sources, two types of sounding data are considered, from the tropics and the midlatitudes, respectively. As a representative over the tropics, we take soundings over the Intensive Flux Array (IFA) during the TOGA-COARE field campaign over the Western Pacific. The observational period was for four months from 1 November 1992 to 28 February 1993. The IFA is encompassed by four sounding stations: Kapingamarangi (1.0° N, 155° E), Kavieng (2.6° S, 151° E), R/V Kexue 1 (4.0° S, 156° E), and R/V Shiyan 3 (2.3° S, 158° E). For the present study, an integration of soundings into a 6–hourly single-column data set, as arranged at the State University of Colorado (available from the web http://tornado.atmos.colostate.edu/togadata/ifa_data.html: Ciesielski et al. 2003), is used.

In this data set, the rain measurement, P, is not directly available, but it is indirectly estimated from a moisture budget, using the local temporal tendency, $\partial I/\partial t$, and convergence rate, $-\text{Div}(I)$, of CIW (I) as well as the evaporation rate, E, as observationally estimated:

$$P = -\frac{\partial I}{\partial t} - \text{Div}(I) + E.$$

Due to data uncertainty, as a result, the rain occasionally becomes negative in this data set. No attempt is made to correct this unphysical behavior.

As a representative over the midlatitudes, we take soundings from Udine-Campoformido (WMO RDS code 16044, 46.0° N, 13.2° E) over Friuli Venezia Giulia (FVG), North East (N. E.) of Italy for a 12–year period, from 1 January 2006 to 1 January 2018. Soundings are available every 12 hours at 00UTC and 12UTC. As for rain, accumulated precipitation over 6 hours up to the moment of the sounding measured by rain gauges at four near-by stations [Codroripo (46.0° N, 13.0° E), Fagagna (46.1° N, 13.1° E), Talmassons (45.9° N, 13.1° E), and Udine (46.0° N, 13.2° E)] are used. Note that FVG is one of the areas with heaviest rain over central Europe (Feudale and Manzato 2014; Isotta et al. 2014; Manzato et al. 2016; Poelman et al. 2016; Pavan et al. 2019).

In addition to these two types of sounding data, following earlier studies by Ambaum and Novak (2014), and Novak et al. (2017), an average over the North Atlantic during five winter periods (December–February) over 1994–1999 is considered. Data is 6-hourly ERA-Interim data, which is averaged over the North–West Atlantic (30–50° N, 30–80° W). Rain is a 6-hour accumulated value. Their goal has been to examine the storm-track dynamics on a basin–scale average. A question in this part of the study is, in turn, how the precipitation is determined under the midlatitude storm-track dynamics under the same settings.

Table 1 lists the mean and the spread (standard deviation) of these three data sets. These spreads approximately correspond to a quarter of the analysis and plotting ranges adopted in subsequent figures.

3.2. Prediction variables

As prediction variables for rain, we primarily consider CAPE and CIW. For logistic reasons, CAPE is calculated slightly differently for the three data sets considered. The

Table 1. Mean and spread (standard deviation) of data.

Area	Varaible	Mean	Spread
			(Standard Deviation)
		$\bar{x}, \bar{y}$	s_x, s_y
Trop.West.Pacific	rain	8.2 (mm/day)	10.2 (mm/day)
	CAPE	414 (J/kg)	229 (J/kg)
	CIW	51.2 (mm)	4.4 (mm)
FVG	rain	0.97 (mm/6h)	2.9 (mm/6h)
	CAPE	180 (J/kg)	382 (J/kg)
	CIW	19.9 (mm)	9.2
North Atlantic	rain	1.09 (mm/6h)	0.41 (mm/6h)
	CAPE	73.5 (J/kg)	29.0 (J/kg)
	CIW	15.3 (mm)	2.03 (mm)
	baroclinicity	4.6 (m/sec)	1.4 (m/sec)

computation procedures of CAPE for the Tropical Western Pacific and Friuli Venezia Giulia (FVG), respectively, follow the procedures described in Yano and Ambaum (2017) and in Manzato and Morgan (2003). ERA data provides CAPE based on forecasts of 6 and 12 hours. CIW is obtained by a direct vertical integral of the water-vapor mixing ratio multiplied by the air density. In addition to these two primary variables, the baroclinicity is also considered over the winter–period North Atlantic, because that is demonstrated to be a main driving force of the storm tracks by Ambaum and Novak (2014), and Novak et al. (2017). Here, the baroclinicity is simplified and described as a difference of the zonal winds between 700 and 850 hPa.

3.3. Data quality

The TOGA–COARE IFA data set is originally delivered in 1997 (Ciesielski et al. 1997), and we use the updated version delivered in 2003 with further calibrations (Ciesielski et al. 2003). The general validity of this data set is systematically verified by Mapes et al. (2003). Furthermore, precipitation data are extensively verified against independent estimates by Johnson and Ciesielski (2000: see especially their Fig. 7) to establish its reliability. For the purpose of the present study, the IFA precipitation time series is re-verified against the MIT–radar precipitation (Rickenbach and Rutledge 1998) for consistency.

A basic error analysis of re–analysis (ERA–Interim) is presented in Dee et al. (2011), who introduce this data set: see especially their Fig. 23(a), which shows that an expected error over the North Atlantic remains at a level of 2 mm/day, although the tropical errors are substantially higher. Reliability of precipitation from re–analysis

(ERA–Interim) is extensively verified further by various studies (e.g. Simmons et al. 2010; Kållberg 2011; Hawcroft et al. 2012; de Leeuw et al. 2015; Hawcroft et al. 2016). See especially Figs. 1–3 of Hawcroft et al. (2012), which compare the re–analysis rain with GPCP.

The whole analysis is also repeated with the ERA5 data set (Hersbach et al. 2020) to further ensure consistency of these data over three regions. Overall agreements between the adopted data and ERA5 are satisfactory, except for over the Tropical Western Pacific, where substantial errors in re–analysis data are expected.

4. Analysis

4.1. Data distribution

Distributions of data for the present study are quantified in Fig. 1 by the effective number, $a(x, y)$, of available points defined by (2.6). In turn, these distributions can be used for inferring relationships between the two variables as a basis for developing an empirical prediction model for rain, from a point of view of correlations, as discussed in introduction.

Fig. 1 overall confirms what is already known in the literature: CAPE is not well correlated with rain, and a much clearer, positive correlation is found with CIW. Over the Tropical Western Pacific, CAPE appears to be only weakly negatively correlated with rain (-0.25). Over N.E. Italy (FVG), a correlation between CAPE and rain is virtually nonexistant (0.06): the bulk majority of the observed points are aligned along an axis either of CAPE or rain. The analysis during the convective season (May-September) is repeated over N.E. Italy (FVG), and a similar distribution is obtained over a phase space of CAPE and rain, but with a noticeable difference that intensive rain is less frequent.
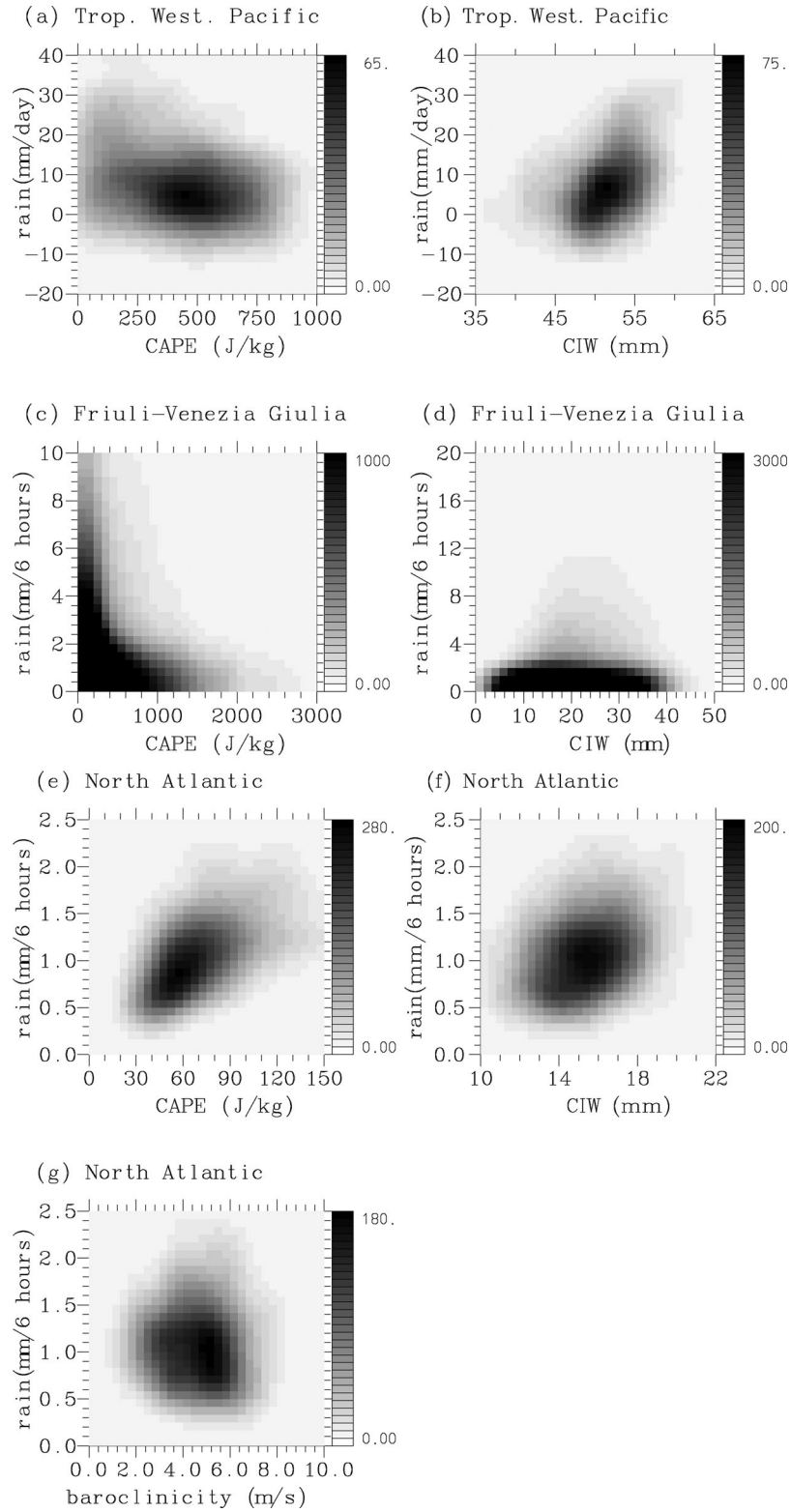
(a) Trop. West. Pacific

(b) Trop. West. Pacific

(c) Friuli–Venezia Giulia

(d) Friuli–Venezia Giulia

(e) North Atlantic

(f) North Atlantic

(g) North Atlantic

Fig. 1.   Effective number of points for analysis defined by Eq. (2.6).
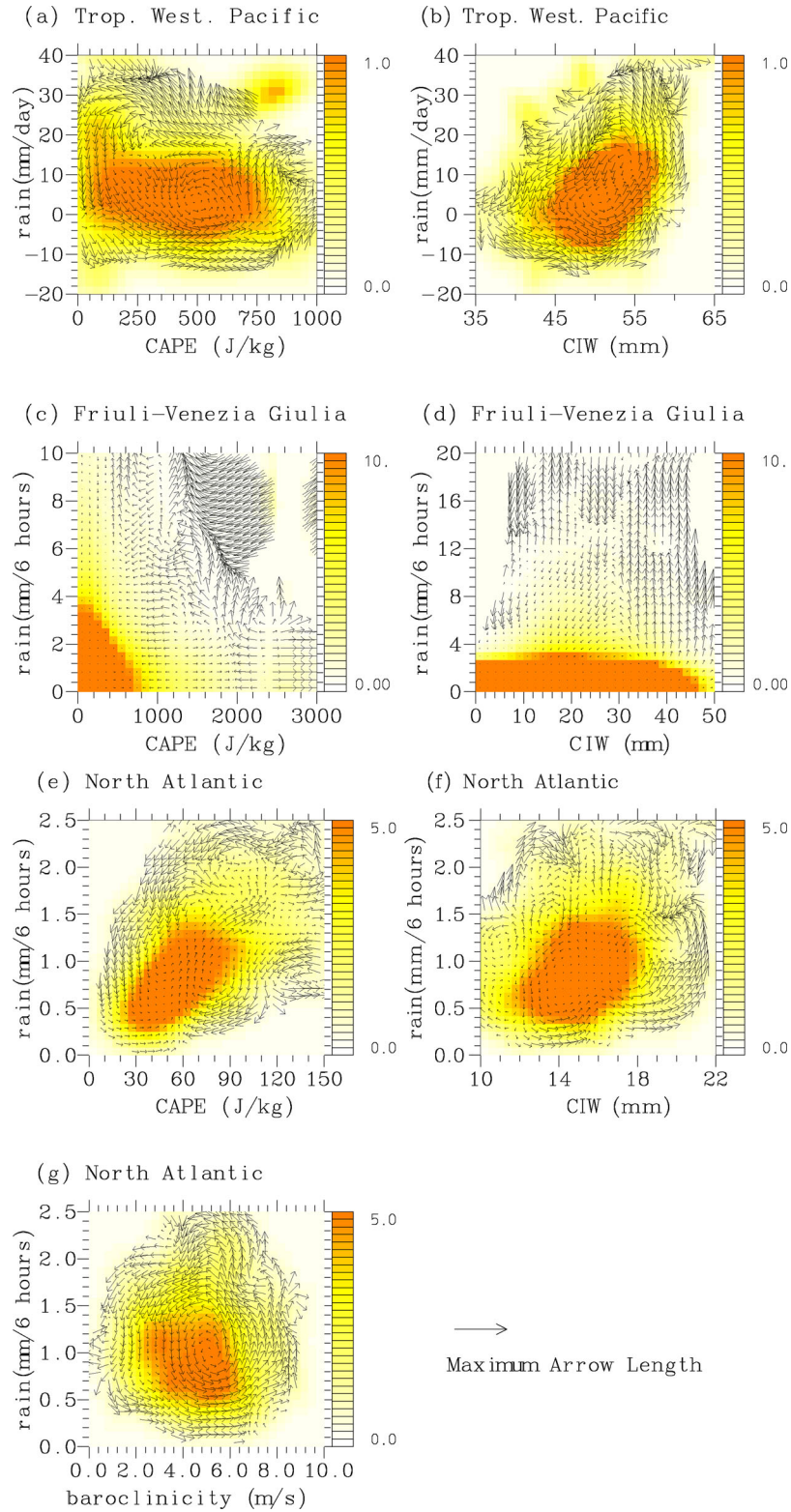
Fig. 2. Phase-velocity vectors: the maximum length indicated at the lower right of the figure is $1\,\mathrm{day}^{-1}$ for (a), (b), $2\,\mathrm{day}^{-1}$ for (c), (d), and $0.5\,\mathrm{day}^{-1}$ for (e), (f), (g). The vectors longer than the maximum are indicated by increasing the arrow head size proportionally. Vectors are marked only over the areas with data density of $a > 1$. Shades show the statistical significance, $\beta(x, y)$, measured by Eq. (2.10).
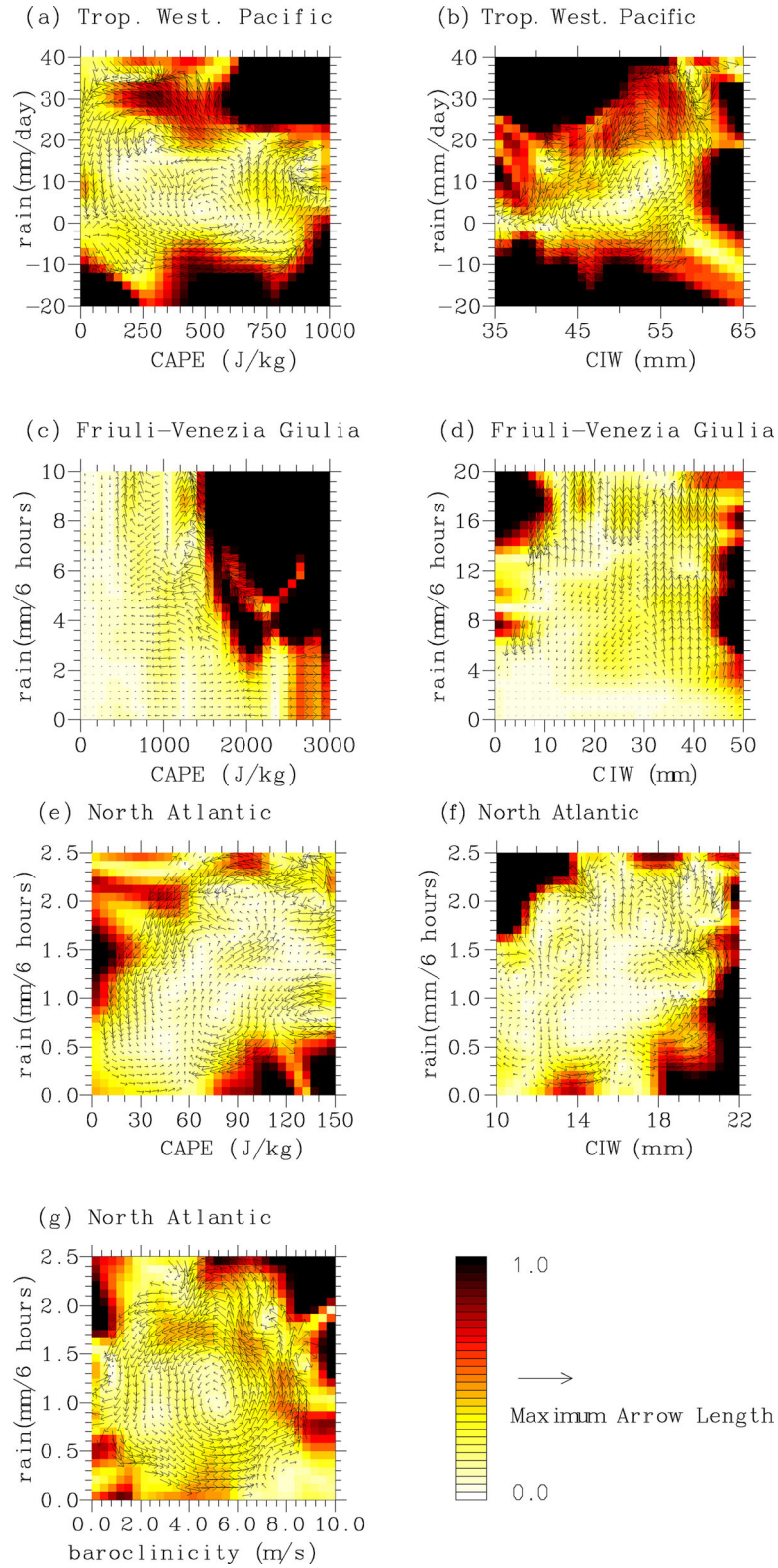
Fig. 3.   Signal-to-noise ratio defined by (4.1) shown by shades. The phase-velocity vectors are also overlain as in Fig. 2.

The rain clearly increases at a faster rate with the increasing CIW over the tropics (with a linear correlation of 0.38): an onset point of rain may be identified with CIW around 50 mm, and increasingly more rain is found beyond this point with the increasing CIW. Such a functional relation is harder to identify over N.E. Italy (FVG), and the linear correlation is weaker (0.17). The most-likely state remains no rain even with increasing CIW. We simply see a greater chance for heavier rains with increasing CIW, associated with an increasing spread of the distribution from a zero-precipitation peak.

The analysis over the winter–period North Atlantic domain leads to qualitatively different results: CAPE is well correlated with rain (0.52) as identified by a linear positive slope in distribution. On the other hand, CIW presents a less clear relation with rain (even though the linear correlation is 0.31). Instead, we find a rather homogeneous, circular-shaped distribution around the modal point. Although the winter–period storm-track system is strongly driven by baroclinicity, it does not represent a clear correlation with rain (-0.03) with a distribution similar to that with CIW. However, as suggested in Sec. 2.1, examining data distributions in this manner is not necessarily a good approach for identifying a controlling variable for a process.

The phase velocities obtained by the procedure of Sec. 2 are plotted in Fig. 2, which is overlain by a measure of statistical significance, $\beta(x, y)$, defined by Eq. (2.10). Vectors are marked only over areas with data density of $a > 1$. Over the tropics (Western Pacific), we clearly identify evidence of a recharge-discharge cycle both with CAPE (a) and CIW (b): both predictors increase without much rain, constituting a recharge phase, until they reach threshold values. These thresholds values are approximately 700 J/kg and 55 mm, respectively, for CAPE and CIW. Above these thresholds, the rain suddenly begins to increase associated only with a weak increase of CAPE and CIW. The tendency continues until the rain rate reaches thresholds, approximately 15 mm/day and 30 mm/day, respectively, with CAPE and CIW. Above this threshold, CAPE and CIW begin to decrease. Over the discharge cycle, rain still tends to increase with CAPE until it almost depletes. With CIW, on the other hand, rain also begins to decrease beyond this point. This discharge phase is completed by returning close to the origin of the phase space. From that point, a recharge cycle begins again.

With the midlatitude soundings (N. E. Italy FVG), CAPE (c) does not form any clear cycle with rain: identified phase velocities are overall horizontal in the phase space, i.e. *only* CAPE changes with little change of the rain rate. A clearly identified pattern is a decrease of CAPE when CAPE is already less than 1200 J/kg,

whereas CAPE is predominantly increasing over 2500 J/kg. A discharge tendency is identified only over a CAPE range of 1200–1800 J/kg with the rain rate above 3 mm/6h.

With CIW obtained by midlatitude soundings (d: N. E. Italy FVG), we can identify an initiation of rain above 30 mm, and a subsequent rapid increase of rain associated with a weak increase of CIW. Above 2 mm/6h, the phase vectors become almost vertical. With the rain above 2 mm/6h and CIW below 30 mm, a discharge cycle is identified. Over a range of 30–40 mm for CIW, the rain increases associated with a decrease of CIW. This tendency suggests that CIW is consumed as rain faster than being supplied, mostly, by low-level convergence. Once CIW reaches below 30 mm, the rain ceases to increase, and gradually turns into a decreasing phase. A recharge branch required for closing the cycle is, however, only identified with weak vectors.

A discharge-recharge cycle of the rain associated with the winter–period North Atlantic storm–track cycle is most clearly identified with baroclinicity (g), and to a lesser extent with CIW (f). With CAPE (g), two branches are identified for discharge for a high and low rain regimes for a given CAPE: both CAPE and rain decrease along these branches. To close the cycles, a recharge phase consisting of an increase of both CAPE and rain must exist, but this cycle cannot be identified clearly in our data, except for a trace identified around 100 J/kg and 1.5 mm/6h, respectively, with CAPE and rain. Furthermore, a reconstructed lower branch of recharge-discharge cycles is physically unintuitive, because a completion of recharge appears to trigger a decrease of rain rather than an increase as expected from this notion.

We should keep in mind that the trajectories shown in Fig. 2 are mean values, as defined by Eqs. (2.5a) and (2.5b): the actual individual trajectory, as traced as a system evolves, substantially fluctuates around this mean trajectory. In other words, the actual evolution of the system is hardly deterministic. The degree of fluctuations of the phase-velocity has been defined by Eqs. (2.7a) and (2.7b). Thus, a degree of such nondeterminism of the system may be measured by a signal-to-noise ratio, $r_{\mathbf{v}}$, defined as

$$r_{\mathbf{v}} \equiv \left[\left(\frac{\dot{x}}{s_u} \right)^2 + \left(\frac{\dot{y}}{s_v} \right)^2 \right]^{1/2}, \qquad (4.1)$$

considering the mean and the degree of fluctuations as signal and noise, respectively. The signal-to-noise ratio (4.1) shown in Fig. 3 is very low, much less than unity in data dense areas (*cf.*, Fig. 1): fluctuations of individual trajectories around the the estimated mean phase-velocities are, most of the time, quite high.

Table 2. Dynamical–system characterisitcs.

Area Predict. Variable	Trop.West.Pacific		FVG		North Atlantic		
	CAPE	CIW	CAPE	CIW	CAPE	CIW	barocl.
λ_1	0.66	0.74	0.10	9.2×10^{-2}	1.5×10^{-2}	3.8×10^{-2}	0.24
λ_2	0.53	0.56	0.30×10^{-3}	1.6×10^{-2}	5.6×10^{-2}	3.1×10^{-2}	0.23
μ_1	-0.17	0.22	-3.2×10^{-2}	9.8×10^{-3}	-4.8×10^{-3}	1.6×10^{-3}	3.4×10^{-3}
μ_2	-0.21	0.14	4.3×10^{-2}	-6.7×10^{-3}	3.1×10^{-2}	1.3×10^{-2}	-6.3×10^{-3}
x_0	397	51	366	19	69	16	4.6
y_0	8.4	7.9	0.36	1.1	1.1	1.2	1.1
ω^2	0.32	0.38	2.8×10^{-4}	1.4×10^{-3}	6.7×10^{-4}	1.0×10^{-3}	5.4×10^{-2}
σ_r	2.0×10^{-2}	3.8×10^{-2}	-3.8×10^{-2}	8.2×10^{-3}	-1.8×10^{-2}	1.4×10^{-3}	4.8×10^{-3}

It also follows that various, 'well–defined' trajectories in Fig. 2 are associated with substantial fluctuations, and hardly deterministic, especially the recharge–discharge cycles identified with various variables. Notably, a discharge branch of the CIW–rain cycle over the FVG (Fig. 2d) is associated with strong fluctuations (i.e. a high noise level). The same is, to some extent, true with the recharge branch of the CAPE–rain cycle over the winter–period North Atlantic (Fig. 2e). In this case, the flow structure itself is not well defined in addition to a high noise level. On the other hand, rather unintuitively, high signal-to-noise ratios tend to be found over data poor areas, partially due to larger mean phase velocities.

Here, it is important to clearly distinguish between *physical* fluctuations of the system, as measured by the signal-to-noise ratio defined by Eq. (4.1), here, and the *statistical* significance of the phase–velocity estimate measured by Eq. (2.10). The former measures an accuracy of predicting the evolution of the system assuming a mean phase velocity, whereas the latter measures an accuracy of statistically estimating a mean value for the phase velocity. It is inaccurate to try to predict an evolution solely based on a mean phase velocity, when the signal-to-noise ratio is very low. However, even in that case, it is still possible to estimate a mean phase velocity accurately, with a high statistical significance, when the data density is high enough.

In general, a high statistical significance of a mean trajectory estimate does not necessarily mean a high degree of determinism of the system. Conversely, a high noise level does not necessarily mean a low statical significance of an estimate. Comparison between Figs. 2 and 3 suggests that these remarks apply rather as the norm rather than the exception in the present analysis.

4.2. Linear dynamics analysis

For understanding the dynamical systems represented by Fig. 2 better, we take the following linear description of the system as an *Ansatz*:

$$\dot{x}' = \mu_1 x' - \lambda_1 y', \qquad (4.2a)$$

$$\dot{y}' = \lambda_2 x' - \mu_2 y'. \qquad (4.2b)$$

Here, the phase-space coordinates (x', y') are defined relative to the fixed point, (x_0, y_0), of a system, indicated by a prime, i.e. $x' = x - x_0$ and $y' = y - y_0$. Signs of the coefficients, λ_j and μ_j ($j = 1, 2$), are defined in such manner that, as shown explicitly in the Appendix A.1, an elliptical oscillator with an axis elongated by a factor, α, tilted from the x-axis by an angle, θ, and with a frequency, ω, is described by

$$\lambda_1 = \omega\left(\alpha \cos^2\theta + \frac{\sin^2\theta}{\alpha}\right), \qquad (4.3a)$$

$$\lambda_2 = \omega\left(\alpha \sin^2\theta + \frac{\cos^2\theta}{\alpha}\right), \qquad (4.3b)$$

$$\mu_1 = \mu_2 = (\alpha - 1/\alpha)\sin\theta\cos\theta. \qquad (4.3c)$$

The coefficients, λ_j and μ_j ($j = 1, 2$), are expected to depend on the phase–space coordinates, (x, y), in general, due to nonlinearity of the system. However, as a first attempt, we estimate these coefficients assuming them to be globally constant. In this case, the phase–velocity data shown in Fig. 2 can be fitted to the linear dynamical system (4.2) by a least–square method. There are six parameters to be estimated with the fixed point coordinates, x_0 and y_0, in addition to the four coefficients, λ_j and μ_j ($j = 1, 2$). The least–square fit is performed by using 30×30 points in phase space, and by using the statistical significance defined by Eq. (2.10) as a weight at each point. The obtained estimates are provided in Table 2. Distributions of local fit errors relative to the magnitude of the local phase–velocity are shown in Fig. 4, along with the estimated fixed points, (x_0, y_0), marked by black circles. Note that the values in Table 2 are after the normalizations defined by Eqs. (2.8) and (2.9).
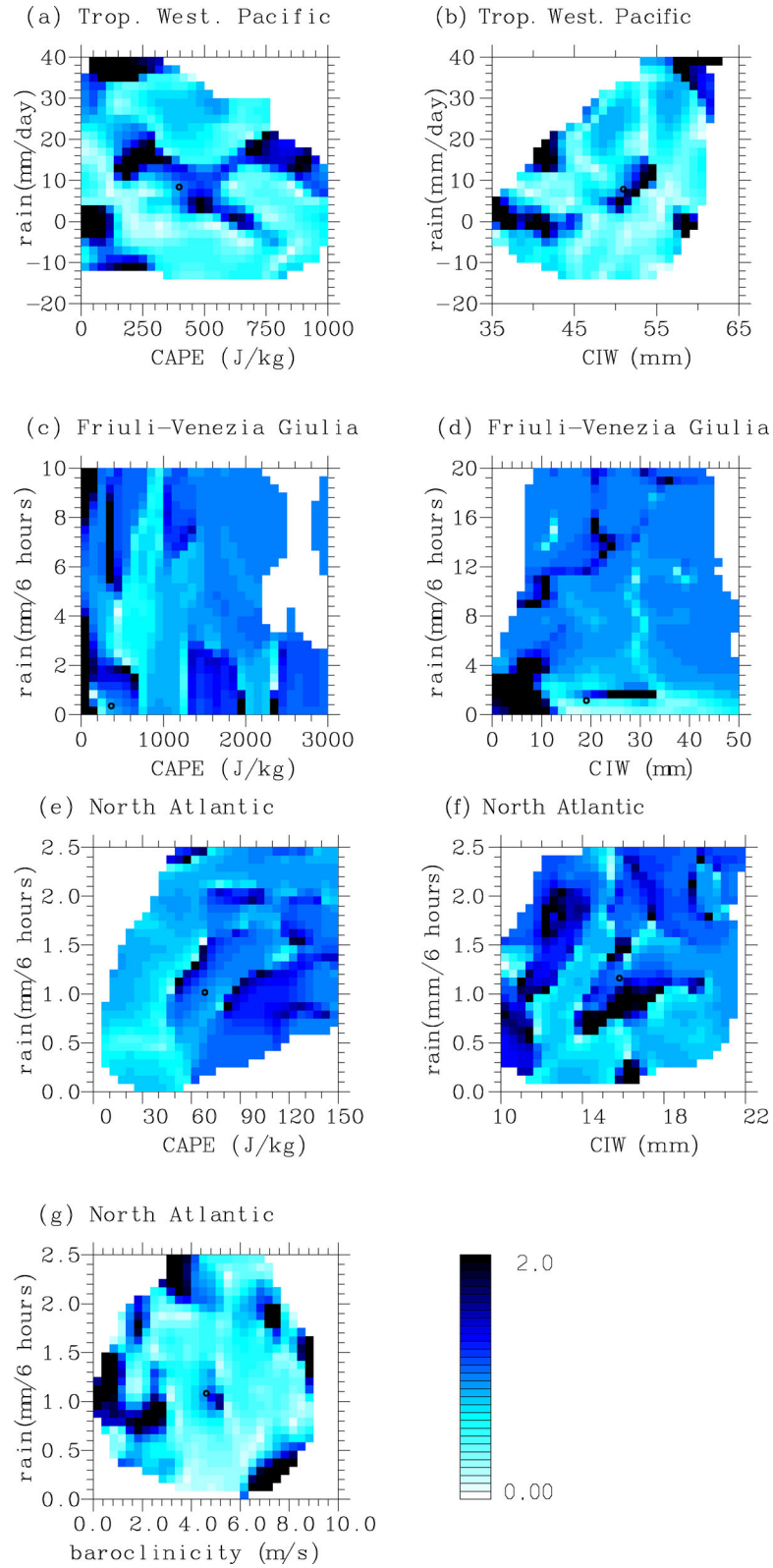
Fig. 4. Error distribution with global fit to the system (4.2). Only areas with $a > 1$ are shaded.

For a system to constitute a coherent periodic cycle, the signs of λ_1 and λ_2 must agree. Furthermore, to form a well-defined elliptical orbit, it requires $\mu_1 \simeq \mu_2$. Table 2 shows that no system satisfies this condition in any clear manner. The system (4.2) can more generally be diagnosed by an eigenvalue analysis, as presented in the Appendix A.2. It shows that the general system (4.2) is described by a frequency, ω, and an exponential tendency, σ_r, defined by

$$\omega^2 = \lambda_1\lambda_2 - \mu_1\mu_2 - \sigma_r^2, \tag{4.4a}$$

$$\sigma_r = \frac{\mu_1 - \mu_2}{2}. \tag{4.4b}$$

These values are also listed in Table 2.

We find that only the systems of the Tropical Western Pacific represent clear periodic cycles with negligible contributions of growth rate. The baroclinicity–rain cycle over the winter–period North Atlantic may also be considered close to a periodic state with the growth rate smaller by an order of magnitude than the frequency. Good periodic cycle descriptions with these three systems are also confirmed by small errors seen in Fig. 4, with the relative errors less than unity over substantial areas in phase spaces. These quantifications are consistent with the periodic cycles that can be identified by eye in Fig. 2 for these three systems.

In the other systems, the contribution of the growth rate is at least comparable to that of the periodicity. The global linear dynamical system is not a good approximation for these systems, either, as also suggested by Fig. 4, in which the relative errors are larger than unity over bulk parts of phase spaces. Furthermore, the growth rates clearly dominate over the periodicity with the CAPE–rain systems both over FVG and the winter–period North Atlantic, and in both cases the systems are damping. These quantifications are consistent with the fact that no clear periodic systems can be identified in the phase–velocity distribution in Fig. 2 for these systems.

To infer the nonlinearity of the system, the same analysis is repeated to define the coefficients of the dynamical system locally, as functions of the phase–space coordinates, (x, y). However, obtained distributions of the coefficients in the phase space present no clear structure to be analysed meaningfully.

5. Discussions

5.1. Limits of the scatter plot approach

To synthesize the present analysis, it may be most helpful to compare between the plots of the data distribution (Fig. 1) and the results from the dynamical–system analysis (Fig. 2). When the data distribution produces a well-defined shape in the phase space, it is tempting to interpret the outline of the shape as the most preferred trajectory of a system. Plots of rain against CAPE and CIW over the Tropical Western Pacific are two good examples. In the former case, rain forms a gentle negative slope against CAPE, suggesting that the rain rate weakly decreases with increasing CAPE (and *vice versa*), although an actual correlation is not that high. In the latter case, it could be interpreted that the rain suddenly begins to increase over a threshold CIW of around 50 mm. However, an inspection of the corresponding phase-velocity fields obtained from the dynamical–system analysis reveals that these identified trajectories from scatter plots constitute only a recharge-phase of a discharge-recharge cycle of the system. The dynamical-system analysis reveals that a discharge phase follows a different preferred trajectory. Less dense data points along the preferred discharge trajectory suggest that, nevertheless, this route is less frequently taken than that for the recharge.

A positive slope identified with CAPE over the winter–period North Atlantic has a different explanation: in this case, the dynamical-system reveals no distinctively preferred evolution tendency along this virtual trajectory. Clear deterministic tendencies for discharge are identified only above and below this virtual trajectory. An opposite example is when a distribution of data is diffused into a circle in the case with baroclinicity over the winter–period North Atlantic. A naive interpretation would suggest no causality between baroclinicity and rain. In contrast, the dynamical–system analysis reveals a well-defined discharge-recharge cycle.

However, the dynamical–system analysis does not always identify a well-defined cycle of discharge and recharge. In the cases both with CAPE and CIW over the N. E. Italy FVG, only the discharge phases are identified. The discharge phase with CAPE is not well-defined with a noticeable divergent tendency of the phase space flows. A recharge phase that must exist to close the cycle is only suggested by less well-defined weak vectors under weak rain regimes. In the case with CAPE over the winter–period North Atlantic, two branches for discharge are identified. An expected recharge phase at the middle of these two branches is only suggested by weak vectors. In the case with CIW over the winter–period North Atlantic, in turn, two branches for recharge are identified, without a well-defined discharge phase.

As a whole, the phase–space trajectories are better defined in extreme, rare situations rather than frequently occurring states. The dynamical-system approach tends to define trajectories better along the periphery of a high

data-density area: phase velocities are larger with larger signal-to-noise ratios. In contrast, high data-density areas tend to be characterized by less well-defined trajectories with smaller signal–to–noise ratios. Thus, data–dense areas are less deterministic than data–scarce areas. Considering the importance of predictions for extreme states, this is another advantage of the dynamical-system approach. An inherent limitation of conventional approaches is that, by focusing on the shape of the data dense structures, less information is gained about the dynamics in the higher amplitude situations.

5.2. Use of CAPE as a predictor of precipitation

It is often argued from theoretical bases (e.g. Emanuel et al. 1994; Rennó and Ingersoll, 1996, see also Yano et al. 2013) that CAPE is a crucial variable for describing the evolution of a convective system associated with its precipitation. However, observational data analyses on midlatitude (e.g. N. E. Italy FVG) and tropics suggests that CAPE has only limited usefulness in general prediction of precipitation. Even from a theoretical point of view, the usefulness of CAPE is limited to the convective precipitation alone.

The present study has shown that CAPE constitutes a well-defined discharge–recharge cycle with rain, as CIW does, over the tropics (Western Pacific). However, in the midlatitudes, no clear periodic cycle with rain is identified either with CAPE or CIW over the N. E. Italy (FVG) as well as over the winter–period North Atlantic. The obtained results are in contrast to earlier studies, which have suggested that CAPE is generally not a good predictor of rain (Sherwood 1999; Manzato 2003; Zhang and Klein 2010; Barkiđija and Fuchs 2013). The present study, in contrast, points to distinctively different roles of CAPE in rain formation between the tropics and the midlatitudes.

A particular problem of using CAPE as a predictor of the convective system may be a fact that it is merely a crude approximation for the cloud work function, which can be shown to be a controlling variable for convection evolution in the context of the mass-flux representation (Arakawa and Schubert 1974; Yano and Plant 2012a, 2012b). Yano et al. (2005) show that the potential energy convertibility (PEC), defined as a diagnosis of the cloud work function from cloud-resolving simulations, presents a much better correlation with convective rain than CAPE. Unfortunately, it is difficult to diagnose PEC from conventional soundings. An alternative approach could be to more directly estimate the cloud work function. However, in this case, a key unknown parameter, entrainment rate must somehow be optimized (cf., Molinari et al. 2012; Sueki and Niino 2016).

5.3. Contrast between the tropics and the midlatitudes

In the present study, as representations of the tropics and the midlatitudes, sounding data over the Western Pacific and those over North-Eastern Italy, Friuli Venezia Giulia (FVG) are considered. The contrast between the results over those two regions would be primarily considered a consequence of different dominant dynamics controlling the precipitation processes over the tropics and the midlatitudes. Over the tropics, rain is expected to primarily be a consequence of convective processes. As a result, the precipitation evolution is expected to follow a recharge––discharge cycle as discussed in Yano and Plant (2012a) for the deep convective system. The result in the CAPE-precipitation phase space supports this expectation. CIW also behaves in a similar manner.

On the other hand, the rain over midlatitudes (N. E. Italy FVG) is primarily controlled by synoptic-scale weather systems. In this case, there is no strong reason to expect a discharge–recharge cycle to be identified, although the majority of rain is still expected to be driven by convection during summer even over this region.[4] The present analysis supports this expectation: no such cycle is identified in CAPE–rain phase space. For CIW, only a discharge phase is identified.

The analysis of the variables averaged over the winter–period North Atlantic, where the storm-track dynamics are most vigorous over the globe, identifies baroclinicity as the most robust variable forming a recharge-discharge cycle with rain. CAPE and CIW present two branches, respectively, of only discharge and recharge phases but without a clear mark of an opposite phase. Although spatial distributions of CAPE and precipitation associated with a synoptic midlatitude storms have been studied (Gray et al. 2011; Glinton et al. 2017), their time evolutions by following the storm cycles are still to be investigated.

The representativeness of the two areas considered with our sounding data needs addressing. The Tropical Western Pacific is an area on the globe where the sea surface temperature is the highest. In this respect, behavior of convection over this area can be considered a stereotype of those over the tropics albeit not necessarily typical. A similar stereotype for midlatitude convection may be considered those over the US Great Plain (cf., Zhang and Klein 2010), where orographical effects can be small. In this respect, geographical characteristics of Friuli Venezia Giulia are rather complex, being flanked by the Alps from north, and facing the Mediterranean in south. For example, a lack of a well-defined recharge branch with CIW may be due to the fact that moist marine air is only available from the Adriatic Sea from the south,

hence only under southerly winds. This interpretation further leads to a speculation that over more continental areas (e.g. Switzerland) the recharge cycle of CIW should be even less deterministic than over FVG.

Clearly more systematic applications of the present methodology to a wider range of geographical locations are required for drawing more general conclusions. The most important contribution of the present study is in suggesting the usefulness of such systematic investigations.

5.4. Towards a dynamical-system prediction model

The present study has been solely focused on data diagnoses. However, the phase-space trajectories identified under the present dynamical–system analysis can equally be applied for prediction purposes. Any operational robustness of this approach is still to be verified by further studies. An important aspect to keep in mind is that the system is fundamentally nondeterministic due to the drastic truncation of the full system to two variables. Perhaps counterintuitively, the nondeterministic part of the phase space tendency tends to be greatest for data dense areas in phase space.

The present dynamical–system formulation suggests a basic prediction strategy different from more traditional approaches. In traditional approaches, rain is *diagnosed* from other predicted variables, such as CAPE and CIW (e.g. Manzato 2005, 2007, 2013). On the other hand, the dynamical system formulation developed here attempts to predict rain along with an adopted prediction variable (predictor). For this purpose, mean trajectories of a system consisting of two variables, for prediction and to be predicted, are constructed in a phase space, along with the associated fluctuations in terms of standard deviations, s_u and s_v (Eqs. 2.7a and 2.7b).

Note that our approach is not equivalent to lagged covariance or lagged correlation approaches: here the statistical information is obtained by averaging events over time, while in our case we average over events in similar locations of the phase space. This allows us to pull together information between situations that are dynamically most similar, even though they may temporally be separated. In this sense, our approach is more aligned with prediction systems based on analogs.

A constructed set of trajectories enables us to perform a prediction based on a dynamical system defined by Eq. (2.3). Here, fluctuations from a deterministic evolution are represented by terms, ϵ_x and ϵ_y. The analysis also provides a measure of fluctuations by Eqs. (2.7a) and (2.7b). These fluctuations can be added as stochastic tendencies in a forecast model so that uncertainties of a forecast can also be directly be quantified. A direct availability of forecast uncertainties also provides a more

objective mean of choosing the best prediction variable for a purpose. The simplest implementation of this prediction model would be to use Eq. (2.3), but neglecting a stochastic term. This provides a straightforward deterministic forecast. However, such a forecast would not be very reliable due to the fact that contributions of stochastic terms (fluctuations) are substantial compared to mean tendencies as shown by direct plots of signal-to-noise ratios (Fig. 3).

Thus, inclusion of stochasticity for representing the uncertainties (standard deviations), s_u and s_v, becomes crucial. Here, stochasticity appears in the system solely as a consequence of a drastic truncation of a system into two variables, rather than from any intrinsic nature of the system. The uncertainty term would be most conveniently treated as Gaussian white noise, although it would be difficult to establish this hypothesis from observations.

The next phase of research along this line would be to evaluate the operational feasibility of a dynamical-system based forecasts system. In spite of wide availability of full numerical weather models, it is still useful and insightful to develop stand-alone simple forecast models that can be applied *in–house* without relying on information from major operational forecast models (*cf.*, Manzato 2005, 2007, 2013).

NOTES

1. A standard simultaneous correlation is, in principle, easily be generalized to include a time lag. However, in the literature, the lag–correlation analysis is not performed as often, presumaly because the lag correlation adds one extra parameter to adjust, and it is often not very obvious how to optimize a lag to obtain the best result.

2. Of course, by introducing an appropriate lag between two variables, a perfect correlation will be obtained. However, in practice, the actual data is always contaminated, and also any given system is never a *perfect* harmonic oscillator. Thus, it is not easy to identify a harmonic oscillator by a lag–correlation analysis in practice.

3. This quantity is also often referred as *column precipitable water*. However, we avoid to use this term herein due to an ambiguity behind the notion *precipitable*. Note that the water vapor in an atmospheric column is never *precipitable* in practice in its totality.

4. A separate analysis limited to a summer period does not change our overall conclusions.

Acknowledgments

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Ambaum, M. H. P. 2010. Significance tests in climate science. *J. Clim.* **23**, 5927–5932. doi:10.1175/2010JCLI3746.1

Ambaum, M. H. P. and Novak, L. 2014. A nonlinear oscillator describing storm track variability. *Q. J. R. Meteorol. Soc.* **140**, 2680–2684. doi:10.1002/qj.2352

Arakawa, A. and Schubert, W. H. 1974. Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *J. Atmos. Sci.* 31, 674–701. doi:10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2

Barkiđija, S. and Fuchs, Ž. 2013. Precipitation correlation between convective available potential energy, convective inhibition and saturation fraction in middle latitudes. *Atmos. Res.* **124**, 170–180. doi:10.1016/j.atmosres.2012.12.010

Brown, R. G. and Zhang, C. 1997. Variability of midtropospheric moisture and its effect on cloud–top height distribution during TOAG COARE. *J. Atmos. Sci.* **54**, 2760–2774. doi:10.1175/1520-0469(1997)054<2760:VOMMAI>2.0.CO;2

Ciesielski, P. E., Hartten, L. and Johnson, R. H. 1997. Impacts of merging profiler and rawinsonde winds on TOGA COARE analyses. *J. Atmos. Oceanic Technol.* **14**, 1264–1279. doi:10.1175/1520-0426(1997)014<1264:IOMPAR>2.0.CO;2

Ciesielski, P. E., Johnson, R. H., Haertel, P. T. and Wang, J. 2003. Corrected TOGA COARE sounding humidity data: impact on diagnosed properties of convection and climate over the warm pool. *J. Clim.* **16**, 2370–2384. doi:10.1175/2790.1

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P. and co-authors. 2011. The ERA–interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* 137,554–597.

Emanuel, K. A., Neelin, J. D. and Bretherton, C. S. 1994. On large–scale circulation in convective atmospheres. *Q. J. R. Meteorol. Soc.* **120**, 1111–1143. doi:10.1002/qj.49712051902

Feudale, L. and Manzato, A. 2014. Cloud-to-ground lightning distribution and its relationship with orography and anthropogenic emissions in the Po valley. *J. Appl. Meteorol. Climatol.* **53**, 2651–2670. doi:10.1175/JAMC-D-14-0037.1

Glinton, M. R., Gray, S. L., Chagnon, J. M. and Morcrette, C. J. 2017. Modulation of precipitation by conditional symmetric instability release. *Atmos. Res.* **185**, 186–201. doi:10.1016/j.atmosres.2016.10.013

Gray, S. L., Martínez–Alvarado, O., Baker, L. H. and Clark, P. A. 2011. Conditional symmetric instability in stingjet storms. *Q. J. R. Meteorol. Soc.* **137**, 1482–1500. doi:10.1002/qj.859

Hawcroft, M. K., Shaffrey, L. C., Hodges, K. I. and Dacre, H. F. 2012. How much Northern Hemisphere precipitation is associated with extratropical cyclones? *Geophys. Res. Lett.* **39**, L24809. doi:10.1029/2012GL053866.

Hawcroft, M. K., Shaffrey, L. C., Hodges, K. I. and Dacre, H. F. 2016. Can climate models represent the precipitation associated with extratropical cyclones? *Clim. Dyn.* **47**, 679–695. doi:10.1007/s00382-015-2863-z

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Hornyi, A. and co-authors. 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. *730*, 1999–2049. doi.org/10.1002/qj.3803

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P. and co-authors. 2014. The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data. *Int. J. Climatol.* **34**, 1657–1675. doi:10.1002/joc.3794

Johnson, R. H., and P. E. Ciesielski. 2000. Rainfall and radiative heating rates from TOGA–COARE atmospheric budgets. *J. Atmos. Sci.* **57**, 1497–1514. doi:10.1175/1520-0469(2000)057<1497:RARHRF>2.0.CO;2

Kållberg, P. 2011. *Forecast drift in ERA–Interim*. ECMWF, Reading, UK.

Manzato, A. 2003. A climatology of instability indices derived from Friuli Venezia Giulia soundings, using three different methods. *Atmos. Res.* **67-68**, 417–454. doi:10.1016/S0169-8095(03)00058-9

de Leeuw, J., Methven, J. and Blackburn, M. 2015. Evaluation of ERA–Interim reanalysis precipitation products using England and Wales observations. *Q. J. R. Meteorol. Soc.* **141**, 798–806. doi:10.1002/qj.2395

Manzato, A. 2005. The use of sounding-derived indices for a neural network short-term thunderstorm forecast. *Wea. Forecast.* **20**, 896–917. doi:10.1175/WAF898.1

Manzato, A. 2007. Sounding-derived indices for neural network based short-term thunderstorm and rainfall forecasts. *Atmos. Res.* **83**, 349–365. doi:10.1016/j.atmosres.2005.10.021

Manzato, A. 2013. Hail in NE Italy: a neural network ensemble forecast using sounding–derived indices. *Wea. Forecast.* **28**, 3–28. doi:10.1175/WAF-D-12-00034.1

Manzato, A., Cicogna, A. and Pucillo, A. 2016. 6-hour maximum rain in Friuli Venezia Giulia: climatology and ECMWF-based forecasts. *Atmos. Res.* **169**, 465–484. doi:10.1016/j.atmosres.2015.07.013

Manzato, A. and Morgan, G. Jr., 2003. Evaluating the sounding instability with the lifted parcel theory. *Atmos. Res.* **67-68**, 455–473. doi:10.1016/S0169-8095(03)00059-0

Mapes, B. E., Ciesielski, P. E. and Johnson, R. H. 2003. Sampling errors in rawinsondes–array budgets. *J. Atmos. Sci.* **60**, 2697–2714. doi:10.1175/1520-0469(2003)060<2697:SEIRB>2.0.CO;2

Mapes, B. E., Milliff, R. and Morzel, J. 2009. Composite life cycle of maritime tropical mesoscale convective systems in scatterometer and microwave satellite observations. *J. Atmos. Sci.* **66**, 199–208. doi:10.1175/2008JAS2746.1

Mapes, B. E., Tulich, S., Lin, J.-L. and Zuidema, P. 2006. The mesoscale convection life cycle: building block or prototype for large-scale tropical waves? *Dyn. Atmos. Oceans* **42**, 3–29. doi:10.1016/j.dynatmoce.2006.03.003

Molinari, J., Romps, D. M., Vollaro, D. and Nguyen, L. 2012. CAPE in tropical cyclones. *J. Atmos. Sci.* **69**, 2452–2463. doi:10.1175/JAS-D-11-0254.1

Novak, L., Ambaum, M. H. P. and Tailleux, R. 2017. Marginal stability and predator-prey behaviour within storm tracks. *Q. J. R. Meteorol. Soc.* **143**, 1421–1433. doi:10.1002/qj.3014

Pavan, V., Antolini, G., Barbiero, R., Berni, N., Brunier, F. and co-authors. 2019. High resolution climate precipitation analysis for north-central Italy, 1961–2015. *Clim. Dyn.* **52**, 3435–3453. doi:10.1007/s00382-018-4337-6

Pavliotis, G. A. and Stuart, A. M. 2007. *Multiscale Methods: Averaging and Homogenization*. Springer, Berlin, 307 pp.

Poelman, D. R., Schulz, W., Diendorfer, G. and Bernardi, M. 2016. The European lightning location system EUCLID - Part 2: observations. *Nat. Hazards Earth Syst. Sci.* **16**, 607–616. doi:10.5194/nhess-16-607-2016

Raymond, D. J. 1995. Regulation of moist convection over the warm tropical oceans. *J. Atmos. Sci.* **52**, 3945–3959. doi:10.1175/1520-0469(1995)052<3945:ROMCOT>2.0.CO;2

Rennó, N. O. and Ingersoll, A. P. 1996. Natural convection as a heat engine: a theory for CAPE. *J. Atmos. Sci.* **53**, 572–585. doi:10.1175/1520-0469(1996)053<0572:NCAAHE>2.0.CO;2

Rickenbach, T. M. and Rutledge, S. A. 1998. Convection in TOGA COARE: horizontal scale, morphology, and rainfall production. *J. Atmos. Sci.* **55**, 2715–2729. doi:10.1175/1520-0469(1998)055<2715:CITCHS>2.0.CO;2

Sherwood, S. C. 1999. Convective precursors and predictability in the Tropical Western Pacific. *Mon. Wea. Rev.* **127**, 2977–2991. doi:10.1175/1520-0493(1999)127<2977:CPAPIT>2.0.CO;2

Sherwood, S. C. and Wahrlich, R. 1999. Observed evolution of tropical deep convection event and their environment. *Mon. Wea. Rev.* **127**, 1777–1795. doi:10.1175/1520-0493(1999)127<1777:OEOTDC>2.0.CO;2

Simmons, A. J., Willett, K. M., Jones, P. D., Thorne, P. W. and Dee, D. P. 2010. Low–frequency variations in surface atmospheric humidity, temperature, and precipitation: inferences from reanalyses and monthly gridded observational data sets. *J. Geophys. Res.* **115**, D01110. doc:10.1029/2009JD012442.

Smith, R. K. and Montgomery, M. T. 2012. Observations of the convective environment in developing and non-developing tropical distrubances. *Q. J. R. Meteorol. Soc.* **138**, 1721–1739. doi:10.1002/qj.1910

Sobel, A. H., Yuter, S. E., Bretherton, C. S. and Kiladis, G. N. 2004. Large–scale meteorology of shallow cumulus convection. *J. Atmos. Sci.* **60**, 1201–1219.

Sueki, K. and Niino, H. 2016. Toward better assessment of tornado potential in typhoons: significance of considering entrainment effects for CAPE. *Geophys. Res. Lett.* **43**, 12,597–604. doi:10.1002/2016GL070349

Yano, J.-I. and Ambaum, M. H. P. 2017. Moist static energy: Definition, reference constants, a conservation law, and effects on buoyancy. *Q. J. R. Meteorol. Soc.* **143**, 2727–2734. doi:10.1002/qj.3121

Yano, J.-I. and Mukougawa, H. 1992. The attractor dimension of a quasi-geostrophic two-layer system. *Geophys. Astrophys. Fluid Dyn.* **65**, 77–91. doi:10.1080/03091929208225240

Yano, J.-I. and Plant, R. S. 2012a. Finite departure from convective quasi-equilibrium: Periodic cycle and discharge-recharge mechanism. *Q. J. R. Meteorol. Soc.* **138**, 626–637. doi:10.1002/qj.957

Yano, J.-I. and Plant, R. S. 2012b. Interactions between shallow and deep convection under a finite departure from convective quasi-equilibrium. *J. Atmos. Sci.* **69**, 3463–3470. doi:10.1175/JAS-D-12-0108.1

Yano, J. I., Bister, M., Fuchs, Z., Gerard, L., Phillips, V. and co-authors. 2013. Phenomenology of convection-parameterization closure. *Atmos. Chem. Phys.* **13**, 4111–4131. doi:10.5194/acp-13-4111-2013 doi:10.5194/acp-13-4111-2013

Yano, J.-I., Chaboureau, J.-P. and Guichard, F. 2005. A generalization of CAPE into potential–energy convertibility. *Q. J. R. Meteorol. Soc.* **131**, 861–875. doi:10.1256/qj.03.188

Zhang, Y. and Klein, S. A. 2010. Mechanism affecting the transition from shallow to deep convection over land: inferences form observations of the diurnal cycle collected at the ARM Southern Great Plains site. *J. Atmos. Sci.* **67**, 2943–2959. doi:10.1175/2010JAS3366.1

Appendix
A.1. Derivation of an elliptical oscillator system

An elliptical oscillator introduced by Eq. (4.2) in the main text with $\mu_1 = \mu_2$ is derived in the following manner. First consider a simple circular oscillator around the origin, described by $x = A \cos \omega t$ and $y = A \sin \omega t$ with A the radius of an orbit. This system is described by a dynamical system:

$$\dot{x} = -\omega y,$$
$$\dot{y} = \omega x.$$

As a simple extension, when the orbit is elongated by a factor, α, in the x-direction, the above system is modified into:

$$\dot{x} = -\alpha \omega y$$
$$\dot{y} = \frac{\omega}{\alpha} x$$

In this case, the orbit satisfies $(x/A)^2 + (y/B)^2 = 1$ with $\alpha = A/B$. When this elongated axis is tilted by an angle, θ, against the x-axis, we designate the rotated coordinates, (X, Y), and perform a linear coordinate transform

$$x = X\cos\theta - Y\sin\theta,$$
$$y = X\sin\theta + Y\cos\theta.$$

The final result is Eq. (4.2) with the coefficients defined by Eq. (4.3) in the main text.

A.2.   *Eigenvalue problem*

Setting σ as the eigenvalue of Eq. (4.2), its eigenvalue problem is:

$$\begin{vmatrix} \mu_1 - \sigma & -\lambda_1 \\ \lambda_2 & -\mu_2 - \sigma \end{vmatrix} = 0$$

It leads to

$$(\mu_1 - \sigma)(\mu_2 + \sigma) - \lambda_1\lambda_2 = 0$$

Solving it for σ, we obtain

$$\sigma = \sigma_r \pm i\omega$$

where σ_r and ω are defined by Eqs. (4.4b) and (4.4a), respectively, in the main text.