Basic mechanism for abrupt monsoon transitions

Anders Levermann\textsuperscript{a,b,1}, Jacob Schewe\textsuperscript{a,b}, Vladimir Petoukhov\textsuperscript{a}, and Hermann Held\textsuperscript{a}

\textsuperscript{a}Earth System Analysis, Potsdam Institute for Climate Impact Research, 14473 Potsdam, Germany; and \textsuperscript{b}Institute of Physics, Potsdam University, 14473 Potsdam, Germany

Edited by Hans Joachim Schellnhuber, Potsdam Institute for Climate Impact Research, Potsdam, Germany and approved August 18, 2009 (received for review February 11, 2009)

Monsoon systems influence the livelihood of hundreds of millions of people. During the Holocene and last glacial period, rainfall in India and China has undergone strong and abrupt changes. Though details of monsoon circulations are complicated, observations reveal a defining moisture-advection feedback that dominates the seasonal heat balance and might act as an internal amplifier, leading to abrupt changes in response to relatively weak external perturbations. Here we present a minimal conceptual model capturing this positive feedback. The basic equations, motivated by observed relations, yield a threshold behavior, robust with respect to addition of other physical processes. Below this threshold in net radiative influx, $R_C$, no conventional monsoon can develop; above $R_C$, two stable regimes exist. We identify a nondimensional parameter $f$ that defines the threshold and makes monsoon systems comparable with respect to the character of their abrupt transition. This framework provides a useful and potentially helpful in understanding past and future variations in monsoon circulation. Within the restrictions of the model, we compute $R_C$ for current monsoon systems in India, China, the Bay of Bengal, West Africa, North America, and Australia, where moisture advection is the main driver of the circulation.

Earth system | tipping element | abrupt climate change | atmospheric circulation | nonlinear dynamics

Monsoon rainfall shapes regional culture and the livelihoods of hundreds of millions of people (e.g., 1, 2). The future evolution of monsoon rainfall under increasing levels of atmospheric CO$_2$ and aerosol pollution is highly uncertain (3). Although greenhouse gas abundance tends to increase monsoon rainfall strength (4–6), the situation is more complex with changing aerosol distribution (7, 8). Given this large uncertainty in the future forcing of monsoons, it is crucial to understand internal monsoon dynamics, especially with respect to self-amplifying feedbacks, which might result in potentially strong responses to small perturbations. Zickfeld et al. (2005) found two stable states in a simple model of the Indian summer monsoon, which in principle allows for rapid transition between radically different monsoon circulations (9, 10) and thereby identified the Indian monsoon as a potential tipping element of the climate system (11). Evidence for such behavior is found in paleodata that show rapid and strong variations in Indian and East Asian monsoon rainfall (12, 13). These abrupt changes have been linked to climatic events in the North Atlantic for the last glacial period (14, 15) as well as for the Holocene (16, 17). Though a physical mechanism for this teleconnection has been suggested (18), relevant climatic signals of the North Atlantic events in Asia (such as temperature and moisture anomalies) are very small (19) indicating that internal feedbacks in monsoon dynamics may have amplified the weak external forcing.

Both spatial patterns and temporal evolution of monsoon rainfall are influenced by a number of physical processes (7, 18, 20–28) as well as characteristics of vegetation (29–31) and topography (32). Though these details are crucial for the specific behavior of different monsoon systems and their significance will vary from region to region, there exist defining processes fundamental to any monsoon dynamics (e.g. 33, 34). These processes are the advection of heat and moisture during monsoon season and the associated rainfall and release of latent heat. In accordance with Zickfeld et al. (9), we suggest the positive moisture-advection feedback (21) as a candidate for the main cause of abrupt changes in monsoon dynamics.

We derive a minimal conceptual model of a monsoon circulation (Fig. 1A), comprising merely conservation of heat and moisture, knowingly neglecting a large number of relevant physical processes in order to distill the fundamental nonlinearity of monsoon circulations. The resulting governing equation exhibits the necessary solution structure to explain qualitatively both strong, persistent changes in monsoon rainfall, as observed in paleorecords, and abrupt variability within one rainy season. This equation’s dynamic similitude, expressed through a single dimensionless number $f$, which defines the threshold behavior and makes different monsoon systems comparable with respect to their transition, may serve as a building block for understanding past and future abrupt changes in monsoon dynamics.

Results

Moisture-Advection Feedback in Monsoon Dynamics. The seasonal evolution of the continental heat budget for different monsoon systems (Fig. 2) shows that sensible heat flux from the land surface increases during spring and heats up the atmospheric column prior to the rainy season. The onset of heavy rainfall (red vertical lines in Fig. 2) is associated with a drop in surface temperature on land, and consequently, sensible heat flux reduces drastically. During the monsoon season, latent heat release dominates the atmospheric heat content, whereas net radiative fluxes are relatively constant throughout the year, reflecting the stabilizing long-wave radiative feedback. In response to the latent heat release, thermal energy is transported out of the region through large-scale advection and synoptic processes. The main dynamical driver of the monsoon is therefore the positive moisture-advection feedback (Fig. 1A): the release of latent heat from precipitation over land adds to the temperature difference between land and ocean, thus driving stronger winds from ocean to land and increasing in this way landward advection of moisture, which leads to enhanced precipitation and associated release of latent heat. In the following, we seek to capture this feedback in a minimal conceptual model.

Minimal Conceptual Model for Abrupt Monsoon Transitions. For this purpose, consider the heat-balance equation of the monsoon season (Fig. 2, for example at blue vertical line).

\[ \mathcal{L} \cdot P = -c_p W \cdot \Delta T + R = 0, \tag{1} \]

where latent heat release and net radiation into the atmospheric column, $R$, balance heat divergence, and the relatively weak contribution from sensible heat transport from the land surface to the atmospheric column has been neglected. $\Delta T$ is the atmospheric

Author contributions: A.L. designed research; A.L. and V.P. performed research; A.L., J.S., and H.H. analyzed data; and A.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

\(^{1}\)To whom correspondence should be addressed. E-mail: anders.levermann@pik-potsdam.de.

This article contains supporting information online at www.pnas.org/cgi/content/full/0901414106 DC supplemental.

www.pnas.org/cgi/doi/10.1073/pnas.0901414106
temperature difference between land and ocean. Latent heat of condensation is \( L = 2.6 \cdot 10^6 \text{J/kg} \) and volumetric heat capacity of air at constant pressure \( C_p = 1,295 \text{J/m}^3/\text{K} \). \( P \) is the mean precipitation over land (in kg/m\(^2\)/s). The ratio \( \epsilon = H/L \) between vertical extent \( H \) of the lower troposphere and the horizontal scale \( L \) of the region of precipitation (Fig. 1) enters because of the balance of the horizontal advective heat transport and the vertical fluxes of net radiative influx \( R \) and precipitation \( P \). A length scale for the coastline drops out. Note that no annual cycle is included in the model. Only budgets for the rainy season are considered. Consequently, this model does not capture any interseasonal or any interannual dynamics. Equations are only valid for landward winds, \( W \geq 0 \).

Assuming dominance of ageostrophic flow in low latitudes, the landward mean wind \( W \) is taken to be proportional to the temperature difference between land and ocean (33, 36, 37):

\[
W = \alpha \cdot \Delta T.
\]  

This assumption of a linear relation between the two quantities is supported by National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis data (Fig. 3) with correlation coefficients above 50% for all regions. There is significant scatter in some plots, reflecting the fact that other processes may be relevant for the monsoon dynamics in the corresponding regions. A possible offset, as observed in some regions, does not alter the model behavior qualitatively. This offset is discussed together with other possibly relevant processes in the SI Appendix. Here we seek to capture only processes relevant to the self-amplification feedback. Neglecting the effect of evaporation over land and associated soil-moisture processes in the continental moisture budget, precipitation has to be balanced by the net landward flow of moisture

\[
\epsilon W \cdot \rho (q_o - q_L) - P = 0, \tag{3}
\]

where \( q_o \) and \( q_L \) are specific humidity over ocean and land, and \( \rho = 1.3 \text{ kg/m}^3 \) is mean air density.

Note that evaporation is clearly an important process for the moisture budget (e.g. (38)) and is omitted in Eq. 3 only for the sake of clarity. Including evaporation does not change the model behavior qualitatively (see SI Appendix). It does, however, shift the value of the critical threshold, as we will show in the next section when applying our model to data. In the minimalistic spirit of this section, we omit the effect of evaporation here because it is not of first order to the problem. Consistent with reanalysis data (Fig. 4) and theoretical considerations (36, 39), continental rainfall is assumed to be proportional to the mean specific humidity within the atmospheric column

\[
P = \beta q_L. \tag{4}
\]

The effect of an offset between these quantities does not change the model behavior qualitatively (see SI Appendix). This set of assumptions (Eqs. 1–4) yields the dimensional governing equation of the model

\[
W^3 + \frac{\beta}{\epsilon \rho} W^2 - \frac{\alpha}{\epsilon C_p} (Lq_o \beta + R) \cdot W - \frac{\beta}{\epsilon^2 \rho C_p} \cdot R = 0. \tag{5}
\]

Note that through the linear relation of Eq. 2, this equation can equally be understood as an expression for the temperature difference between land and ocean \( \Delta T \), which might be more useful for some applications. Introduction of nondimensional variables \( w = \frac{W}{\epsilon \rho/\beta} \) and \( p = \frac{P}{Lq_o \beta} \) results in the nondimensional equation

\[
w^3 + w^2 - (l + r)w - r = 0, \tag{6}
\]

which depends on two parameters only: The dimensionless net radiative influx \( \epsilon \equiv R - \alpha \sigma q_o^2/(C_p \beta^2) \) and a measure for the relative role of latent and advective heat transport

\[
l \equiv (\alpha \sigma q_o^2 Lq_o)/(\epsilon C_p \beta^2), \tag{7}
\]

Large \( l \) corresponds to a strong influence of moisture advection (scaling as \( Lq_o \beta p \)) on the continental heat budget compared with heat advection by large-scale and synoptic processes (scaling as \( C_p \beta^2 \sigma^2/(\alpha \sigma^2) \)). The nondimensional precipitation is directly related to the wind through \( p = w/(1 + w) \).

Solutions \( w(r) \) of Eq. 6 are determined entirely by a choice of the only free parameter \( l \), which can be expressed in terms of a critical threshold of net radiative flux \( r_c \), below which no physical solution exists (Fig. 5). The critical point \( (r_c, w_c) \) will vary for different monsoon systems. It is directly linked to the only remaining parameter \( l \), through

\[
w_c (w_c + 1)^2 = l/2. \tag{8}
\]

and therefore uniquely defines the solution \( w(r) \) of the model. The critical radiation can be computed from

\[
r_c = -w_c^2 (2w_c + 1). \tag{9}
\]

Thus for large \( l \) (as observed in some monsoon systems) the critical threshold is well approximated by \( r_c \approx -l \). Note that \( l \) is scaling like \( q_o a/\beta \) where \( a \) and \( \beta \) have clear-cut physical meaning (39). \( a \) is essentially a function of the near-surface cross-isobar angle and thereby a function of surface roughness and static stability of the
planetary boundary layer (PBL). β is governed by the characteristic turnover (recycling) time of liquid water in the atmosphere and thereby determined by static stability and vertical velocity in the PBL. Any physical solution for ρ > r_c is characterized by landward winds w > 0 and positive precipitation p > 0.

Let us now try to understand the physical mechanism behind the threshold behavior observed in Fig. 5. In the tropics net radiative influx is negative, i.e. radiation cools the atmospheric column. During monsoon season the same is true for the advection of heat by the winds because winds blow predominantly from the colder oceanic surrounding. The release of latent heat compensates for both of these heat-loss processes. If monsoon winds get weaker, condensation and therefore latent heat release through precipitation are reduced (moisture-advection feedback, Fig. 4A). The abruptness of the transition emerges through an additional stabilizing effect of the direct heat advection which is cooling the atmospheric column and is also reduced for reduced monsoon winds. Thus both advection-related processes, evaporative warming and thermal cooling, are simultaneously reduced and partly compensate until a threshold is reached at which condensation/precipitation cannot provide the necessary latent heat to sustain a circulation. As a consequence, land-ocean temperature difference ΔT and therewith monsoon winds break down (Fig. 1B).

Estimate of Critical Threshold for Current Monsoon Systems. In order to estimate the critical threshold of different monsoon systems within the limitation of this very simple model, we use time series of precipitation P, radiation R, temperature difference ΔT, and specific humidity q_o from the NCEP/NCAR reanalysis data (35) to compute time series for α(t) = (∆P + R)/(εC_pΔT^2) and β(t) = (∆L - R - P)/(εC_pΔT), assuming applicability of the model and stationary statistics within the observational period (1948–2007). Via α(t) and β(t), the parameter f(t) is known and the system is estimated for each year. As a simple test for the model, we calculate the remaining quantity that is not used for the computation of α(t) and β(t), the specific humidity over land

\[ q_L(t) = q_o(t) - \frac{C_p \Delta T(t) P(t)}{\rho (C_p \Delta T(t) + R(t))}. \]  

The resulting model estimate of the specific humidity q_L compares reasonably well (Fig. S2 in the SI Appendix) with the independently observed q_L that was used in Fig. 4 to motivate the relation between specific humidity and precipitation (Eq. 4).
Because the processes represented in our model are fundamental to monsoon systems, we believe that the results strongly suggest the possibility of abrupt monsoon transitions. Because the dominant driving process is captured, it is not impossible that the model can provide a reasonable estimate for the critical threshold, $R_c$, once all necessary processes are incorporated. The bifurcation structure of the model is robust with respect to incorporation of other physical processes (see SI Appendix) and only changes qualitatively when either of these perturbations dominate the dynamics. Thus, the applicability of our model is based on the assumption that moisture advection is the dominant process in the heat budget of a monsoon system.

The possibility of abrupt transition is due to the competition of the main heat transport processes during the rainy season. Although latent heat release through precipitation warms the atmospheric column, direct advection of heat is cooling it. Both processes decrease with decreasing monsoon winds and thereby compensate each other with respect to the net heat injection into the atmospheric column. The threshold of this stabilizing effect is set by the radiative cooling, which is characteristic to low-latitudes and is strongly influenced by aerosol distribution in the region.

According to our model, abrupt transitions may occur in two different ways. For net radiation above the critical threshold $R > R_c$, the system is bistable. Because the model only describes the rainy season and does not capture the annual monsoon cycle, abrupt transitions in the bistable regime can only be interpreted intraseasonally, e.g., a month of heavy rain followed by a month of extraordinarily weak precipitation. An example could be the extremely weak rainfall in July and September observed in India in the year 2002, in which the rest of the season exhibited average rainfall (41).

Our model does not capture the dynamics of a decline or increase in monsoon strength over several years. Thus, paleodata in which strong variation in monsoon rainfall have been recorded cannot be explained by the bistable regime because these recordings show monsoon changes over several years, decades, or

![Fig. 3. Landward zonal wind versus temperature difference between land and ocean during monsoon season (NCEP/NCAR reanalysis data (35)). The lines show best linear regression with correlation r.](image1)

Via the definition of $I$, we compute $R_c$ from the time series $a(t)$ and $b(t)$ for each year between 1948-2007. Note that the only quantity that is not constrained by data in this computation is the parameter $\epsilon$, which defines the ratio of vertical and horizontal scale. However, the critical threshold $R_c$ is independent of $\epsilon$, and thus the calculation depends only on relatively robust averaged values of precipitation, net radiation, average temperature difference between land and ocean, specific humidity over ocean, and the natural constants $\rho$, $L$, and $C_p$. We interpret the resulting distribution of the critical threshold $R_c$ (Fig. 6, blue) as a noisy estimate of a stationary critical threshold.

Within the limitations of the model, the observed net radiation is higher than the critical threshold in the Bay of Bengal, West Africa, and China. In India, North America, and Australia, the distributions have significant overlap. Incorporating evaporation into the model shifts the distribution toward lower thresholds (Fig. 6, red), while at the same time increasing the precipitation threshold $P$. Standard bootstrapping (see SI Appendix) reveals that the estimates in Fig. 6 are already relative robust distributions, in view of the simplicity of the model approach.

**Discussion**

A minimal conceptual model for monsoon circulations that captures the moisture-advection feedback is presented. The model is unlikely to describe details of monsoon circulations quantitatively, nor is it meant to capture all dynamical processes of a monsoon circulation. Following a minimalistic philosophy, the model comprises the necessary processes for a positive feedback and thereby demonstrates the possibility of an abrupt transition of monsoon circulations from a state with strong rainfall to a weak precipitation state. All model equations are backed by relations found in NCEP/NCAR reanalysis data. For India, it has been shown that this data-set properly represents the statistics of precipitation when compared with regional observations with higher spatial resolution (40).
even centuries. Such behavior would correspond to a shift of the system across the critical threshold into the monostable regime $R < R_C$ without a conventional monsoon. If persistent, such a shift would be visible in paleorecords.

The reduction of the full set of model parameters to a single scaling number $I$, which determines the system and thereby the critical threshold, testifies to a remarkable dynamic similitude with respect to the atmospheric quantities $\alpha$, $\beta$, and $q_0$. Different monsoon systems with the same $I$ will have the same transition behavior. As illustrated in Fig. 5, $I$ provides a measure for the position and the sharpness of the transition, i.e., for the point $(r_c, w_c)$ in state space. This means, in particular, that a decrease in inflowing humidity $q_0$ associated with, e.g., colder climate conditions (which would decrease the threshold $R_c$ and shift the system closer to a collapse) could be compensated by decreasing $\beta$, representing a lower turnover (recycling) time of moisture in the atmosphere, which is influenced by, e.g., aerosols.

Fig. 5. Solution of the nondimensional governing Eq. 6. (Top) Nondimensional landward wind for two values of the only parameter $I = 0$ and $I = 1$. (Middle) Corresponding nondimensional precipitation. (Bottom) Nondimensional precipitation for higher values of $I$ as observed in some monsoon systems. The functional form of the solution is not changed qualitatively. The critical threshold $(w_c(I), p_c(I))$ is given as the black curve in each frame.

The parameter $q_0$ in our model can be interpreted in a rather broad sense as a specific humidity of the vicinity influencing a monsoon region. As an example, the years with anomalously high snow cover over the Tibetan Plateau in spring and early summer (20, 26) could be characterized by a decrease in $q_0$ during midsummer, which would shift the threshold value $R_c$ for the Indian monsoon closer to the observed precipitation over the region, thus increasing a possibility of monsoon breakdown in those years. Similarly, a colder climate with generally decreased humidity $q_0$ could be closer to the critical threshold, which might be the reason for less-stable monsoon circulations during glacial periods.

In the future, net radiation may be reduced through aerosol pollution, which will push the system qualitatively closer to the critical threshold (7). On the Indian subcontinent, in China, and in parts of sub-Saharan Africa, agricultural productivity is closely linked to and limited by monsoon rainfall. Food security in these regions is particularly sensitive to monsoon variability (42–45). Studies with comprehensive models are necessary to confirm or reject the idea of the existence of a threshold as well as its position.

ACKNOWLEDGMENTS. We thank B.N. Goswami, R. Krishnan, and J. Srinivasan for helpful hints and discussions; and T. Lenton for useful comments on the manuscript. This work was funded by the Heinrich Böll Foundation, the German National Academic Foundation, and the German Federal Ministry of Education and Research.