A statistically predictive model for future monsoon failure in India

You may also be interested in:

Multi-decadal modulation of the El Nino–Indian monsoon relationship by Indian Oceanvariability
Caroline C Ummenhofer, Alexander Sen Gupta, Yue Li et al.

Shift in Indian summer monsoon onset during 1976/1977
A S Sahana, Subimal Ghosh, Auroop Ganguly et al.

Multi-scale drought and ocean–atmosphere variability in monsoon Asia
Manuel Hernandez, Caroline C Ummenhofer and Kevin J Anchukaitis

What caused the cool summer over northern Central Asia, East Asia and central North America during 2009?
Kyung-Ja Ha, Jung-Eun Chu, June-Yi Lee et al.

Understanding ENSO dynamics through the exploration of past climates
Steven J Phipps and Jaclyn N Brown

Improved simulation of two types of El Niño in CMIP5 models
Jong-Seong Kug, Yoo-Geun Ham, June-Yi Lee et al.

Systematic winter sea-surface temperature biases in the northern Arabian Sea in HiGEM and the CMIP3 models
D Marathayil, A G Turner, L C Shaffrey et al.

Different impacts of the two types of El Nino on Asian summer monsoon onset
Xin Wang, Xingwen Jiang, Song Yang et al.

Robust warming over East Asia during the boreal winter monsoon and its possible causes
Sun-Seon Lee, Seon-Hwa Kim, Jong-Ghap Jhun et al.

A statistically predictive model for future monsoon failure in India

Jacob Schewe and Anders Levermann

Potsdam Institute for Climate Impact Research, Telegrafenberg A62, D-14473 Potsdam, Germany
Institute of Physics, Potsdam University, Potsdam, Germany

E-mail: jacob.schewe@pik-potsdam.de

Abstract
Indian monsoon rainfall is vital for a large share of the world's population. Both reliably projecting India's future precipitation and unraveling abrupt cessations of monsoon rainfall found in paleorecords require improved understanding of its stability properties. While details of monsoon circulations and the associated rainfall are complex, full-season failure is dominated by large-scale positive feedbacks within the region. Here we find that in a comprehensive climate model, monsoon failure is possible but very rare under pre-industrial conditions, while under future warming it becomes much more frequent. We identify the fundamental intraseasonal feedbacks that are responsible for monsoon failure in the climate model, relate these to observational data, and build a statistically predictive model for such failure. This model provides a simple dynamical explanation for future changes in the frequency distribution of seasonal mean all-Indian rainfall. Forced only by global mean temperature and the strength of the Pacific Walker circulation in spring, it reproduces the trend as well as the multidecadal variability in the mean and skewness of the distribution, as found in the climate model. The approach offers an alternative perspective on large-scale monsoon variability as the result of internal instabilities modulated by pre-seasonal ambient climate conditions.

1. Introduction

Indian summer monsoon (ISM) rainfall is the major prerequisite of agricultural productivity in the region, and its variability severely affects the livelihoods of a large share of the world's population [1, 2]. While average ISM rainfall has been relatively stable during the past century of direct observations, rising trends have been observed in the annual number of extreme rain events [3]. The future evolution of the ISM, and other monsoon systems, under a combination of anthropogenic forcing factors is unclear [4]: recent projections indicate that the response to increased greenhouse gas (GHG) concentrations may differ in sign among major monsoon regions, and reveal large uncertainties about the magnitude of the response [5–7]. The effect of increased aerosol abundance is significant and may be counteracting that of GHGs [8, 9], while human-induced vegetation changes feed back on precipitation [10, 11]. Observations and modeling studies suggest a recent regime shift in Asian monsoon convection [12] and its relation to northern hemisphere thermal gradients [13, 14]. At the same time, paleodata testify to abrupt and strong shifts in both the Indian and the East Asian monsoon during the last two glacial cycles [15–18] and the Holocene [19, 20].

In this study, we show that more severe failure of ISM rainfall than ever observed in the past century is possible but unlikely under present climatic conditions, according to a realistic coupled climate model; and that such failure is projected to become much more frequent over the next 200 years under a climate change scenario. We present a simple concept of internal monsoon dynamics that is based on the fundamental driving mechanisms of continental monsoon rainfall, and demonstrate that it can explain both multidecadal variability and future trend in ISM rainfall as found in the climate model.

2. Fundamental monsoon dynamics

We use an ensemble of five millennial climate simulations performed with the MPI-ESM comprehensive climate model, which has been shown to produce a good representation of the Indian summer monsoon (see appendix A.1, for a description of the model and the simulations used in this study, references, and details of the simulation analysis). Seasonal (June–August) mean rainfall over the total of 6030 model years shows a characteristic frequency distribution which extends far towards low rainfall values, featuring very weak monsoon years with seasonal mean rainfall below 2 mm d^{-1} (figure 1(A), gray bars). Latent-classes analysis suggests that this characteristic distribution represents a superposition of two monsoon modes (supplementary figure S1 available at stacks.iop.org/ERL/7/044023/mmedia; see appendix for details).

We attempt to understand this frequency distribution in the context of the fundamental driving mechanisms of the monsoon circulation and associated continental monsoon rainfall. We consider average monsoon rainfall over the Indian subcontinent, setting aside regional differences in intensity and timing of rainfall events. The onset of monsoon precipitation in spring is initiated by the development of a tropospheric temperature contrast between land and ocean, and associated convergence of moist air over the continent. During the rainy season, the atmospheric heat budget in the monsoon region (figure S2(A) available at stacks.iop.org/ERL/7/044023/mmedia) is then dominated by the so-called moisture–advection feedback: the release of latent heat by precipitation reinforces the land–ocean tropospheric temperature contrast, thereby stabilizing the circulation that brings in more moist air from the ocean, which in turn maintains precipitation [21, 22] (figures 2(A) and (B), top).

We argue that the moisture–advection feedback stabilizes the rainy monsoon regime by inducing persistence: rainfall over a certain period within the rainy season releases latent heat, thereby reinforces the circulation and increases the probability for rainfall in subsequent days. This 'memory effect' can be seen in direct observations, where it is reflected as a quasi-proportionality between the rainfall on a given day t, P_t, and the amount of latent heating accumulated during some period prior to that day. Here we use daily rainfall data for 1951–2003 on a 1° grid from India Meteorological Department (IMD) [23], averaged over all available data within 5–30°N, 70–90°E. For the summer

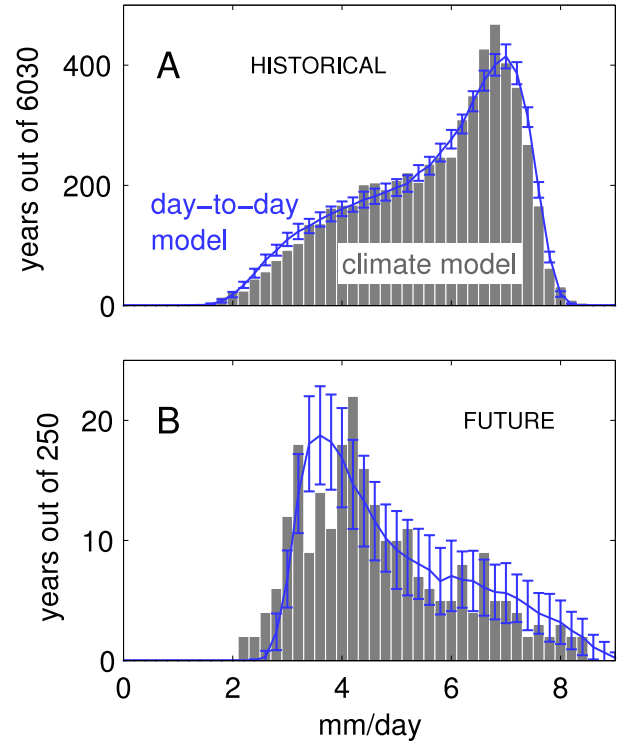**Figure 1.** ISM rainfall statistics. (A) Distribution of seasonal (JJA) mean ISM precipitation from the MPI-ESM climate simulation ensemble over the period 800–2005 CE (gray bars; 5 simulations, 1206 years each) and from 6030 runs of the stochastic day-to-day model (blue line shows mean value, and error bars show ±1 standard deviation, from 100 realizations of the model). (B) Gray bars as in (A), but over the period 2151–200 CE in a global warming scenario (see main text). Blue line and error bars as in (A), but from 250 runs and with a different set of model parameters (see supplementary table S1 available at stacks.iop.org/ERL/7/044023/mmedia).

months (May–September), we plot each day's precipitation against the normalized average rainfall $\tilde{p}$ during a memory period τ prior to that day, such that at day t,

$$\tilde{p}_t \equiv \frac{\langle P \rangle_{t-\tau}^{t-1} - P_-}{P_+ - P_-}, \qquad (1)$$

where P is the daily precipitation rate; $\langle \cdots \rangle$ indicate temporal averaging; and P_+ and P_- are maximum and minimum precipitation rates, respectively. The memory effect is reflected in the correlation between P_t and $\tilde{p}_t$ for all days within the season, shown in figure 3(A) for $\tau = 17$ days. The correlation depends on the choice of τ, and a sensitivity test reveals that the rainfall memory is effective on a timescale of 2–3 weeks (figure 3(B); see appendix A.2 for details); after that time, the effect of the rainfall history becomes undetectable. In particular, the signal-to-noise ratio of the correlation is greater than 1 for $\tau \leq 18$ days. We interpret $\tilde{p}$ as a probability for daily precipitation.

Using the same value $\tau = 17$ days as above, and averaging daily rainfall over 5–30°N, 70–90°E, we find the memory effect well represented in the climate model output (figure 3(A), inset). The atmospheric heat budget in the model is similar to the one observed in reanalysis data (figure S2(B)
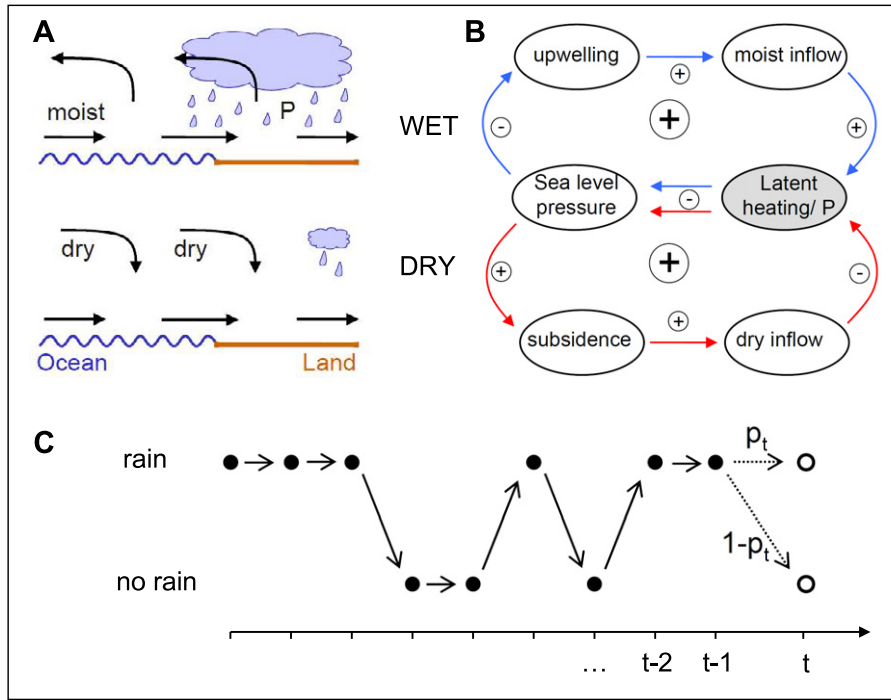
Figure 2. Illustration of idealized ISM circulation regimes (A), associated self-amplifying feedback mechanisms (B), and the simple 'day-to-day model' (C). In the wet regime, latent heating due to precipitation (P) creates a regional sea level pressure (SLP) low, which leads to atmospheric upwelling and associated inflow of moist air towards the ISM region. In turn, a lack of latent heating maintains a positive regional SLP anomaly, which leads to subsidence of dry upper troposphere air and lowers the humidity of the monsoon winds, thereby inhibiting precipitation and sustaining a dry regime. In the 'day-to-day model', the probability for a wet or dry day depends on the amount of latent heating accumulated during some previous period, except during the onset when it is determined by an initial probability p_{init}.
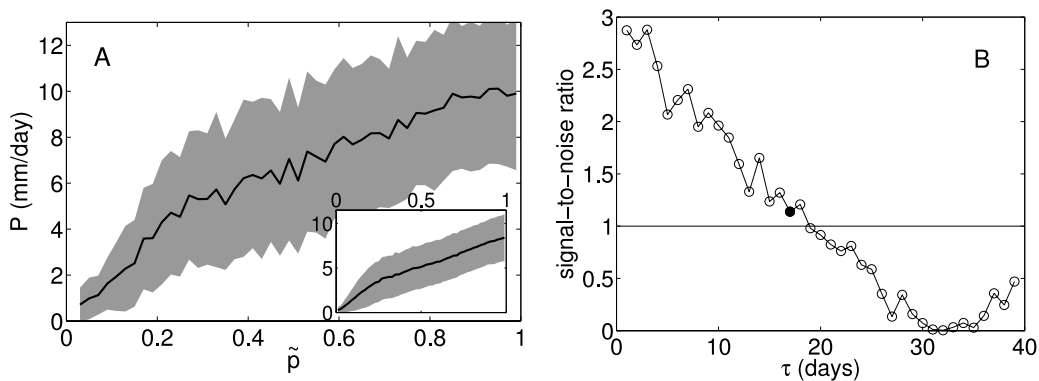

Figure 3. Memory effect in IMD daily rainfall data. (A) Daily rainfall during May–September versus precipitation probability $\tilde{p}$ (see main text), for a memory period $\tau = 17$ days. The thick line connects the mean values for each of 50 bins on the $\tilde{p}$-axis; shading denotes ± 1 standard deviation. Days with $\tilde{p} < 0.2$ stem mainly from before the monsoon onset. Inset: as main figure, but with P and $\tilde{p}$ diagnosed from one of the MPI-ESM simulations. Axes as in main figure. (B) Sensitivity of the memory effect to the memory period τ. The 'signal-to-noise ratio', i.e. the ratio of the slope of the correlation versus the standard deviation as shown in A, is greater than one for $\tau \leq 18$ d, and becomes very small for τ greater than 25 d. The filled circle marks $\tau = 17$ d, the value used for further analysis.

available at stacks.iop.org/ERL/7/044023/mmedia). Thus assuming that the moisture–advection feedback generally dominates the rainy season in the climate model as well, how does monsoon failure, as seen in figure 1(A), arise?

As an example, we investigate year 1158 of the first ensemble member as one of the driest years within the entire simulation ensemble (due to significant internal variability in this class of climate models, this year cannot be expected to coincide with any specific historical year, and will

henceforth be referred to as the 'dry year'). Average seasonal precipitation over India in this year is ∼70% below the long-term mean of 5.7 mm d^{-1} (figure 4(A), circles). The dynamical South Asian monsoon index [24] indicates an exceptionally large deviation of the zonal tropospheric wind shear from the long-term mean (figure 4(A), bars).

All-India average precipitation is erratic and largely suppressed during the entire summer (figure 4(B)), and we observe a qualitatively different circulation pattern
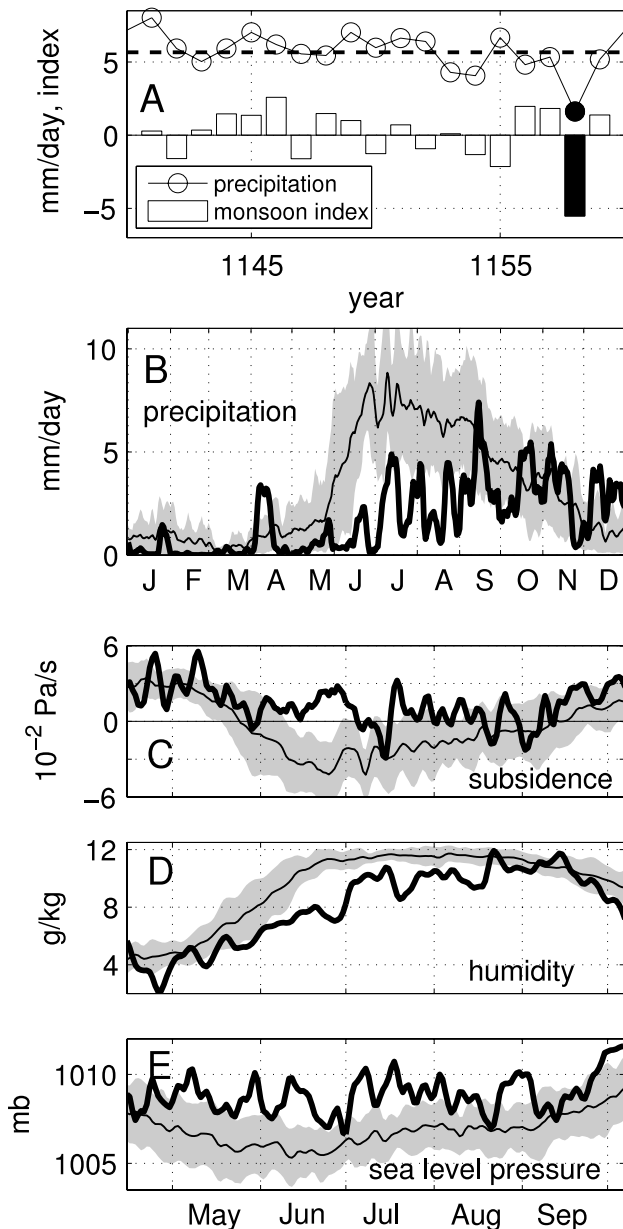
Figure 4. The dry year. (A) Seasonal mean precipitation over India (circles, in mm d^{-1}; dashed line shows long-term mean) and South Asian monsoon index [24] (bars) in a 19 yr interval of the simulation; the dry year is marked in black. (B) Annual cycle of average daily precipitation over India during the dry year (thick) and averaged over the previous 17 yr (thin; shading denotes the 16% and 84% quantiles). See figure S6 (available at stacks.iop.org/ERL/7/ 044023/mmedia) for a comparison covering the entire simulation period. (C)–(E) Characteristics of the monsoon circulation over western India and the Arabian Sea during May–September: (C) vertical velocity ω (in 10^{-2} Pa s^{-1}) at 500 hPa, with positive values indicating subsidence of air; (D) near-surface (850–1000 hPa) specific humidity, in g kg^{-1}; and (E) mean sea level pressure (MSLP, in hPa). In (B)–(E), a 3 d running mean was applied; thick lines are for the dry year; and thin lines give the average over the previous 17 yr, where the shading denotes ±1 standard deviation.

compared to normal monsoon years: instead of convective upwelling, large-scale subsidence of upper-tropospheric air occurs both over the subcontinent and over the Arabian

sea (figure 4(C) and map in figure S3 available at stacks. iop.org/ERL/7/044023/mmedia). Since specific humidity is lower in the upper troposphere, this effectively dries out the monsoon region: specific humidity in the lower troposphere (i.e. in the lower, landward branch of the ISM circulation, figure 2(A)) substantially decreases over India and the eastern Arabian Sea, and remains exceptionally low through June–August (figure 4(D) and map in figure S4 (available at stacks.iop.org/ERL/7/044023/mmedia); also note the spatial congruence of subsidence and moisture anomalies in figure S4 available at stacks.iop.org/ERL/7/ 044023/mmedia). This acts to further suppress precipitation and renders the regional-scale sea level pressure anomalously high throughout the summer (figure 4(E) and map in figure S5 available at stacks.iop.org/ERL/7/044023/mmedia). Consistent with this anomaly pattern, the direction of the monsoon flow in the upper troposphere is changed from generally westward during normal monsoon years, to northward or even eastward during this dry year, as reflected in the dynamical monsoon index (figure 4(A), bars). We suggest that this anomalous circulation pattern constitutes another self-amplifying feedback (figure 2(A), bottom) that can sustain a dry monsoon regime in a similar way as the wet regime is sustained by the moisture–advection feedback (figure 2(B)). Rather than simply an inactive mode of the moisture–advection feedback, this dry-subsidence feedback is an active dynamical counterpart because of the inertia induced by the large-scale drying-out of the monsoon winds. The hypothesis that intraseasonal dynamics are dominated by these two counteracting feedback mechanisms is consistent with the idea of two distinct modes in seasonal mean rainfall suggested by the latent-classes analysis above. In the following, we explore the potential of this hypothesis.

3. Simple stochastic model of ISM rainfall

The dry year examined above is an extreme case in the sense that the dry-subsidence feedback is dominant throughout the season. Most years in the spectrum of the climate model (figure 1(A), gray bars) are less extreme, i.e. their seasonal precipitation averages somewhere between the dry state and an upper limit of about 9 mm d^{-1}. We propose a simple statistically predictive model for seasonal mean monsoon precipitation that explains this characteristic spectrum by an interplay between the wet moisture–advection feedback and the dry-subsidence feedback on a quasi-daily timescale. This 'day-to-day' model (illustration figure 2(C)) is based only on the two feedbacks and the memory effect. It employs the idealized assumption of two discrete circulation states: one with weak precipitation (we choose $P_- \equiv 0$ mm d^{-1} in our example, supplementary table S1 available at stacks.iop.org/ ERL/7/044023/mmedia), corresponding to the dry feedback loop in figure 2(B), and one with maximum precipitation (P_+, here chosen to be 9 mm d^{-1}), corresponding to the wet feedback loop in figure 2(B). It is further assumed that the system can flip between these two states on a quasi-daily timescale, and that the probability p_t for ending up in the wet state at day t depends on the ratio of dry and rainy
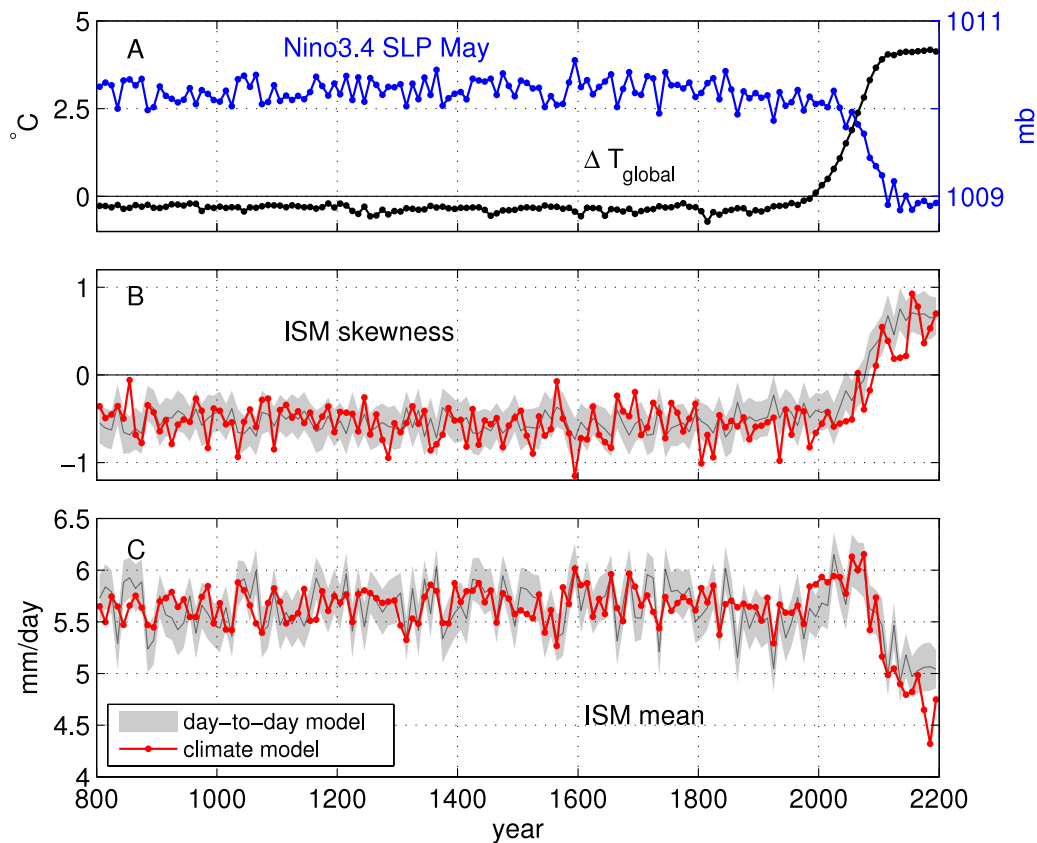
Figure 5. Application of the 'day-to-day model' to past and future ISM variability. (A) Global mean surface temperature (black, in °C relative to the 1980–99 mean) and May MSLP over the Nino3.4 region (blue, in mb) over the full period of the millennial MPI-ESM simulations and their continuations under a global warming scenario. Data have been averaged decadally (801–810, etc) and over the 5 ensemble members, such that each point represents an average over 50 model years. (B) Skewness of the decadal frequency distribution of seasonal mean ISM rainfall in MPI-ESM (red). Each point is computed from a set of 50 model years (10 yr, 5 ensemble members). The skewness computed from 50 runs each of the day-to-day model, with parameters set according to global mean temperature and May Nino3.4 MSLP as in (A), is shown in gray (thin line shows mean value, and shading shows ±1 standard deviation, from 100 realizations of the model). (C) As (B), but for the mean of the frequency distribution (in mm d^{-1}).

days during a certain period τ prior to t—this corresponds to the precipitation probability $\tilde{p}$ of equation (1): that is, the more rainy days there were in the period τ prior to day t, the more likely it is to have rainfall at day t. Thus here the memory effect is represented by an exact proportionality between rainfall and rainfall history. It acts to impede a flip between the two circulation states. Still, a flip can always be triggered by stochastic atmospheric fluctuations, e.g. regional and local weather, that perturb the large-scale dynamical flows comprised in the two feedback mechanisms. In the day-to-day model, these perturbations are represented by the fact that the actual system state at a given day is randomly drawn from the probability distribution defined by p_t.

This very basic auto-regression model of internal monsoon dynamics needs to be complemented by information on the onset of the rainy season, prior to the dominance of internal dynamics as represented by the two positive feedback loops. During the period of length τ at the beginning of the season, a constant initial probability for rainfall, p_{init}, replaces the dynamical rainfall probability $\tilde{p}$. p_{init} represents external factors that crucially influence the development of the monsoon circulation at a time when it is most sensitive to external perturbations [24]. In the case of India this could for example comprise the mean state of the tropical atmospheric circulation related to the El Niño/Southern Oscillation (ENSO), Indian Ocean sea surface temperatures, or Eurasian snow cover [25], as well as the current phase of the Madden–Julian oscillation (MJO) [26]. Finally, we introduce a maximum probability p_{m}, such that p_t cannot be greater than p_{m} or smaller than $(1 - p_{\text{m}})$. This reflects the presence of synoptic variability and prevents locking into either P_- or P_+ when p_t becomes zero or one, which otherwise could occur as an artifact of the discrete nature of the model. In this form, the day-to-day model can reproduce the spectrum of the comprehensive climate model (figure 1(A), blue line). The Matlab$^{®}$ code is provided as supplementary information (available at stacks.iop.org/ERL/7/044023/mmedia), and more detail on the model formulation and parameters is given in appendix A.3.

4. Monsoon failure under global warming

The MPI-ESM climate model simulations of the past millennium, as used above, have been complemented with IPCC SRES A1B scenario [27] simulations until 2100,
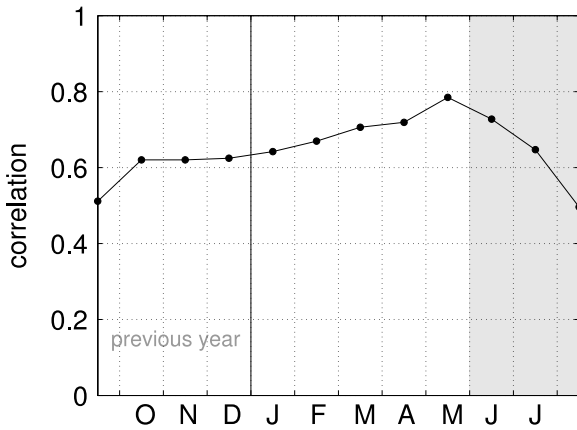
5

Figure 6. Coefficient of the correlation between seasonal (JJA) mean ISM rainfall and monthly-averaged mean sea level pressure (MSLP) over the Nino3.4 region (170–120°W, 5°S–5°N) in the MPI-ESM simulation. The correlation is strongest when MSLP is taken in May just before the monsoon season (indicated by the gray shading).

followed by constant CO_2 concentrations until 2200. This results in an increase of global mean surface air temperature of 4.6 °C relative to pre-industrial by the end of the 22nd century (figure 5(A), black line). We find that in the warm climate of 2151–200, the frequency distribution of seasonal mean ISM rainfall is inverted compared to pre-industrial conditions, in the sense that now dry years are much more frequent than wet years (figure 1(B), gray bars). This inversion occurs gradually and monotonically over the course of the warming scenario and involves a sign change in the distribution's skewness at the end of the 21st century (figure 5(B), red line). The associated change in the expected seasonal mean rainfall, however, is not monotonic: an increase throughout much of the 21st century is followed by a rapid decrease, falling short of the pre-industrial long-term mean roughly by the turn of the 22nd century (figure 5(C), red line). This finding is consistent, in terms of both sign and magnitude, with a previous study [28] reporting an increase of ISM precipitation in a higher resolution version of the MPI climate model under a $2 \times CO_2$ stabilization scenario; while at the same time it shows that qualitatively different responses can emerge when longer timescales and/or higher levels of radiative forcing are considered.

We attempt to understand these changes in the simple framework of the day-to-day model, and consider how the very few model parameters P_+, P_-, and p_{init}, previously held constant, might be persistently altered under global warming. A warmer tropical troposphere can hold more water vapor, which tends to enhance rainfall in the absence of counteracting processes [29]. As a first-order approximation, we translate this into a linear shift of the precipitation range $[P_-, P_+]$ towards higher values with increasing global mean temperature.

In addition to this thermodynamical effect, dynamical changes that affect the monsoon development during the onset period can be represented in the day-to-day model via the parameter p_{init}. The shifts in atmospheric pressure patterns associated with ENSO are known to be closely related to interannual variations in monsoon strength: El-Niño years (especially those with a maximum surface warming in the central Pacific) tend to produce drought in India through a weakening and associated displacement of the Walker circulation, which induces anomalous subsidence, and thus suppresses convection, broadly over the Indian subcontinent [30, 31]. In the climate model, the correlation between seasonal mean ISM rainfall and the strength of the Walker circulation, as measured by the mean sea level pressure (MSLP) anomaly over the Nino3.4 region in the central Pacific (170–120°W, 5°S–5°N), is strongest for May (figure 6); i.e. the state of the Walker circulation at this time of the year (the monsoon onset period) is most influential for the following monsoon season.

In the context of our stochastic day-to-day model, we are interested in the effect of long-term, decadal-scale variations in the state of the Walker circulation (as opposed to its oscillations on the ENSO timescale) on the frequency distribution of ISM rainfall. As a first-order approximation, we scale p_{init} linearly with decadally-averaged May Nino3.4 MSLP. Figure 5(A) shows the two climate model timeseries that are used to drive the simple model. In addition to some multidecadal variability, a marked decrease appears in the Walker circulation index (blue) in the future projections, indicating a persistent weakening of the Walker circulation, which is consistent with thermodynamical considerations and other modeling studies [32]. This decrease is delayed with respect to the global mean temperature (black) increase by about 30 yr.

When p_{init}, P_+ and P_- are varied according to these timeseries, the day-to-day model reproduces the long-term trends in mean and skewness of the ISM rainfall frequency distribution observed in MPI-ESM (figures 5(B) and (C), gray lines and shading). The non-trivial peak-and-decline response in ISM rainfall is captured due to the delay in the weakening of the Walker circulation compared to the global warming signal. Moreover, a significant portion of multidecadal rainfall variability is reproduced. The blue line in figure 1(B) shows the ISM rainfall frequency distribution from the day-to-day model for the future period 2151–200.

5. Discussion

Indian summer monsoon rainfall exhibits intraseasonal variability on various timescales. While some of this variability, in particular its slower components, is related to slowly varying boundary forcing and interaction with larger scale atmospheric phenomena such as the MJO [26], internally generated variability is likely an equally important factor in shaping the monsoon cycle and determining the seasonal mean rainfall [33–35]. We have presented a slim and fundamental framework to explore the potential role of an internal instability of the large-scale monsoon circulation for ISM variability on longer timescales. We have demonstrated that the non-trivial frequency distribution of seasonal mean monsoon rainfall in a complex climate model simulation can be largely explained on the basis of only two fundamental assumptions: (i) an inherent instability of the monsoon

circulation on a quasi-daily timescale, in the form of a competition between two different circulation regimes, each of which is associated with a self-amplifying feedback. (ii) A memory effect that is induced by those feedbacks, and determines, in a probabilistic manner, the transition between the wet and the dry periods. Both assumptions are backed by observations and comprehensive climate model results. The dry-subsidence feedback not only complements the picture of the atmospheric circulation in the temporary absence of monsoon rainfall, but constitutes an active antagonist of the moisture–advection feedback due to its self-amplifying nature. The memory effect is a consequence of the dominant role of these two feedbacks, and is found in observational data as well as in climate model results.

Moreover, with the above assumptions, ISM changes on decadal to centennial timescales can be explained to a large degree using a minimal amount of external information. We have shown that the 'day-to-day model' can reproduce past ISM multidecadal variability as well as projected future changes, including a substantial increase in monsoon failure, when driven by decadal changes in global mean temperature and tropical Pacific spring-time MSLP. While global warming is assumed to elevate the baseline of monsoon rainfall due to increased atmospheric moisture content, a persistent weakening of the Walker circulation, as reflected in lower MSLP in the tropical Pacific, is assumed to induce atmospheric conditions that favor more subsidence over the Indian region and thus lead to more deficient monsoon onsets. In the day-to-day model, this translates into a lower initial probability p_{init}. The delay between the global warming signal and the weakening of the Walker circulation then causes the up-and-down response of average ISM rainfall over the course of the 21st and 22nd centuries: rainfall is strengthened during much of the 21st century due to the thermodynamic effect of atmospheric warming; but then weakens, once the change of the Walker circulation mean state becomes the dominant signal and the development of a conventional monsoon circulation is more frequently suppressed via the dynamical interaction with ENSO.

The parameter p_{init} summarizes the influence of ambient climatic factors that, to some extent, predispose the monsoon system during the onset season [33]. It does not determine the rainfall amount of an individual monsoon season, but it has a strong influence on the probability distribution of seasonal mean rainfall. While in the comprehensive climate model applied here, the mean state of the Walker circulation is a good single predictor, other factors such as changing spring-time Indian Ocean sea surface temperatures might play a significant role, too [36], and could be incorporated into p_{init}. Similarly, while the day-to-day model is designed to simulate seasonal mean characteristics using a minimal set of assumptions about intraseasonal dynamics, externally forced intraseasonal variations in monsoon strength could be superimposed on the internal variability generated in the model to obtain a closer approximation of observed intraseasonal variability, and thus also improve the simulated seasonal mean rainfall.

Acknowledgments

Appendix A. Methods

A.1. Climate model analysis

The climate model simulations used in this study belong to the 'Millennium' simulation ensemble [37]. They were performed with the comprehensive MPI-M Earth System Model MPI-ESM, comprising the atmospheric general circulation model (GCM) ECHAM5 [38, 39] Version 5.4.01 at T31 (∼3.75°) horizontal resolution and 19 vertical levels; the land surface scheme JSBACH [40] at T31 resolution, one soil layer, 13 vegetation types and four vegetation tiles per grid box; and the oceanic GCM MPI-OM [41] Version 1.3.0, including the ocean biogeochemistry module HAMOCC [42], at GR3.0L40 resolution. Comparison with other GCMs and with observational data shows that ECHAM5 reproduces the Indian summer monsoon realistically [28, 43]. The five historical simulations each span the period 800-2005 CE. Time-dependent external forcing is the same for each simulation and includes the effects of volcanic stratospheric sulfate aerosols [44], variation of total solar irradiance [45], land-use change [46], and changes in orbital parameters. While they include the past century of anthropogenic GHG emissions, we consider the total of 1206 years sufficiently homogenous to be treated as belonging to the same climate state. Following the historical period, the simulations are continued with forcing according to the IPCC SRES A1B scenario [27] until 2100, and with CO_2 emissions corresponding to constant CO_2 concentrations from 2101 until 2200. Further details can be found on the MPI-M website [47].

In analyzing the climate model results, precipitation is averaged over the land area in 5–30°N, 70–90°E. The dynamical monsoon index [24] is computed as the zonal wind shear difference over 0–20°N, 60–100°E according to $(\bar{U}_{850\,\text{hPa}} - \langle \bar{U}_{850\,\text{hPa}} \rangle) - (\bar{U}_{200\,\text{hPa}} - \langle \bar{U}_{200\,\text{hPa}} \rangle)$, where $\bar{U}$ is the seasonal (May–September) average zonal wind speed, and angle brackets indicate the long-term mean over the historical simulation period. Sea level pressure is averaged over 0–25°N, 60–80°E; vertical velocity ω in pressure coordinates is averaged over 5–30°N, 55–80°E; and specific humidity over 5–25°N, 60–80°E. For demonstrating the memory effect in MPI-ESM, the same procedure is applied as for the IMD observational data, except that we now choose $P_+ = 9$ mm d^{-1}, which is closer to the maximum 17 d average daily precipitation observed in the climate model.

The latent-classes analysis is performed using the flexmix routine [48], fitting a model of one, two or three superposed Gaussian distributions to the data. According to the Bayesian Information Criterion [49] (BIC), the assumption of two Gaussians improves the model substantially (from a BIC value of 21 364 to 20 034) as compared to only one Gaussian, while a third Gaussian yields no significant improvement (BIC 20 010), and is not physically motivated.

A.2. Memory effect in observations

For demonstrating the memory effect in the IMD observational data set, we choose $P_+ = 12$ mm d^{-1} and $P_- = 0$ mm d^{-1}. P_+ is chosen to approximate the maximum observed 17 d average daily precipitation. We test the sensitivity of the memory effect to the length of the memory period τ by computing the 'signal-to-noise ratio' of the correlation for a range of τ values (figure 3(B)): a linear regression of the mean values between $\tilde{p} = 0.4$ and 0.6 is evaluated at $\tilde{p} = 0.2$ and 0.8, and the difference in P is divided by the average standard deviation over the same interval. As the maximum observed average daily precipitation is higher for shorter averaging periods and vice versa, we have varied P_+ along with τ—namely, between 14 and 10 mm d^{-1}—in producing figure 3(B). This variation however only has a minor effect on the signal-to-noise ratio, which is much more sensitive to the choice of τ.

A.3. Statistically predictive model

The precipitation probability used in the simple day-to-day model is defined as

$$p_t = \begin{cases} p_{\mathrm{m}} & \text{if } \tilde{p}_t \geq p_{\mathrm{m}}, \\ (1 - p_{\mathrm{m}}) & \text{if } \tilde{p}_t \leq (1 - p_{\mathrm{m}}), \\ \tilde{p}_t & \text{else,} \end{cases} \quad (A.1)$$

where p_{m} is a maximum probability, and $\tilde{p}_t$ is defined according to equation (1) in the main text; except that t is now the index of model time steps, not days. During an initial period of length τ in the beginning of the season, p_t is not (and cannot be) determined by equation (A.1). Instead, we introduce an initial probability p_{init} which is treated as a model parameter. This day-to-day model is integrated over l days, and the seasonal mean precipitation $\bar{P}$, $P_- \leq \bar{P} \leq P_+$, is saved. Thus, the result of the day-to-day model is determined by six parameters (see supplementary table S1 available at stacks. iop.org/ERL/7/044023/mmedia): l, p_{m}, τ, P_-, P_+, and p_{init}. The length of the model season l corresponds to a quasi-daily time step and the results are not qualitatively dependent on it as long as this timescale association is respected. The parameter p_{m} is used as a fit parameter, and its value indicates the relative importance of synoptic fluctuations that counteract the persistence of the two circulation regimes associated with the moisture–advection feedback and the dry-subsidence feedback, respectively. The value of τ is chosen according to the analysis of observed rainfall data shown in figure 3(B). All these parameters are held constant throughout all applications

of the day-to-day model. Small variations in these parameters produce only small changes in the result (i.e. the shape of the rainfall frequency distribution).

The remaining three parameters, P_-, P_+, and p_{init}, take on different values for reproducing the historical and future frequency distribution, respectively (cf figure 1) For reproducing the transient changes throughout the historical as well as the future climate simulations (figure 5), we run the day-to-day model for each decade, varying the parameter p_{init} according to $p_{\mathrm{init}} = p'(m - m_0) + p_0$, where m is the decadally-averaged May Nino3.4 MSLP in mb; $m_0 = 1008.9$ mb; $p' = 0.39$ mb^{-1}; and $p_0 = 0.2$. The parameters P_+ and P_- are varied according to $P_\pm = P'\Delta T + P_0^\pm$, where ΔT is the decadally-averaged global mean surface temperature anomaly in °C; $P' = 0.42$ mm day^{-1} °C^{-1}; $P_0^+ = 9$ mm day^{-1}; and $P_0^- = 0$ mm day^{-1}.

See supplementary table S1 (available at stacks.iop.org/ ERL/7/044023/mmedia) for all sets of parameter values used in the model application, and the attached Matlab® script for the model code.

References

[1] Parthasarathy B, Munot A and Kothawale D 1988 Regression model for estimation of Indian foodgrain production from summer monsoon rainfall *Agric. Forest Meteorol.* **42** 167–82

[2] Auffhammer M, Ramanathan V and Vincent J R 2006 Integrated model shows that atmospheric brown clouds and greenhouse gases have reduced rice harvests in India *Proc. Natl Acad. Sci.* **103** 19668–72

[3] Goswami B N, Venugopal V, Sengupta D, Madhusoodanan M S and Xavier P K 2006 Increasing trend of extreme rain events over India in a warming environment *Science* **314** 1442–5

[4] Meehl G A *et al* 2007 Global climate projections *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* ed S Solomon *et al* (Cambridge: Cambridge University Press)

[5] Kripalani R H, Oh J H and Chaudhari H S 2007 Response of the East Asian summer monsoon to doubled atmospheric CO$_2$: coupled climate model simulations and projections under IPCC AR4 *Theor. Appl. Climatol.* **87** 1–28

[6] Kripalani R H, Oh J H, Kulkarni A, Sabade S S and Chaudhari H S 2007 South Asian summer monsoon precipitation variability: coupled climate model simulations and projections under IPCC AR4 *Theor. Appl. Climatol.* **90** 133–59

[7] Cherchi A, Alessandri A, Masina S and Navarra A 2011 Effects of increased CO$_2$ levels on monsoons *Clim. Dyn.* **37** 83–101

[8] Ramanathan V *et al* 2005 Atmospheric brown clouds: impacts on South Asian climate and hydrological cycle *Proc. Natl Acad. Sci.* **102** 5326–33

[9] Lau K M and Kim K M 2006 Observational relationships between aerosol and asian monsoon rainfall, and circulation *Geophys. Res. Lett.* **33** L21810

[10] Ganopolski A, Kubatzki C, Claussen M, Brovkin V and Petoukhov V 1998 The influence of vegetation– atmosphere–ocean interaction on climate during the mid-holocene *Science* **280** 1916–9

[11] Claussen M 2009 Late quaternary vegetation-climate feedbacks *Clim. Past* **5** 203–16

[12] Turner A G and Hannachi A 2010 Is there regime behavior in monsoon convection in the late 20th century? *Geophys. Res. Lett.* **37** L16706

[13] Li J *et al* 2009 Summer monsoon moisture variability over China and Mongolia during the past four centuries *Geophys. Res. Lett.* **36** L22705

[14] Sun Y, Ding Y and Da A 2010 Changing links between South Asian summer monsoon circulation and tropospheric land–sea thermal contrasts under a warming scenario *Geophys. Res. Lett.* **37** L02704

[15] Burns S J, Fleitmann D, Matter A, Kramers J and Al-Subbary A A 2003 Indian ocean climate and an absolute chronology over Dansgaard/Oeschger events 9 to 13 *Science* **301** 1365–7

[16] Wang P *et al* 2005 Evolution and variability of the Asian monsoon system: state of the art and outstanding issues *Quat. Sci. Rev.* **24** 595–629

[17] Wang Y *et al* 2008 Millennial- and orbital-scale changes in the East Asian monsoon over the past 224,000 years *Nature* **451** 1090–3

[18] Schewe J, Levermann A and Cheng H 2012 A critical humidity threshold for monsoon transitions *Clim. Past* **8** 535–44

[19] Gupta A K, Anderson D M and Overpeck J T 2003 Abrupt changes in the Asian southwest monsoon during the Holocene and their links to the North Atlantic ocean *Nature* **421** 354–7

[20] Wang Y *et al* 2005 The Holocene Asian Monsoon: links to solar changes and North Atlantic climate *Science* **308** 854–7

[21] Webster P J *et al* 1998 Monsoons: processes, predictability, and the prospects for prediction *J. Geophys. Res.* **103** 14451–510

[22] Levermann A, Schewe J, Petoukhov V and Held H 2009 Basic mechanism for abrupt monsoon transitions *Proc. Natl Acad. Sci.* **106** 20572–7

[23] Rajeevan M, Bhate J, Kale J A and Lal B 2006 High resolution daily gridded rainfall data for the Indian region: analysis of break and active monsoon spells *Curr. Sci.* **91** 296–306

[24] Webster P J and Yang S 1992 Monsoon and ENSO: selectively interactive systems *Quart. J. R. Meteorol. Soc.* **118** 877–926

[25] Charney J G and Shukla J 1981 Predictability of monsoons *Monsoon Dynamics* (Cambridge: Cambridge University Press)

[26] Wheeler M and Hendon H 2004 An all-season real-time multivariate MJO index: development of an index for monitoring and prediction *Mon. Weather Rev.* **132** 1917–32

[27] Nakicenovic N and Swart R (ed) 2000 *IPCC Special Report on Emissions Scenarios* (Cambridge: Cambridge University Press)

[28] Annamalai H, Hamilton K and Sperber K R 2007 The South Asian summer monsoon and its relationship with ENSO in the IPCC AR4 simulations *J. Clim.* **20** 1071–92

[29] Allen M and Ingram W 2002 Constraints on future changes in climate and the hydrologic cycle *Nature* **419** 224–32

[30] Kumar K K, Rajagopalan B, Hoerling M, Bates G and Cane M 2006 Unraveling the mystery of Indian monsoon failure during El Nino *Science* **314** 115–9

[31] Ratnam J V, Behera S K, Masumoto Y, Takahashi K and Yamagata T 2010 Pacific Ocean origin for the 2009 Indian summer monsoon failure *Geophys. Res. Lett.* **37** L07807

[32] Vecchi G *et al* 2006 Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing *Nature* **441** 73–6

[33] Palmer T N 1994 Chaos and predictability in forecasting the monsoons *Proc. Natl Acad. Sci.* **60** 57–66

[34] Krishnamurthy V and Shukla J 2000 Intraseasonal and interannual variability of rainfall over India *J. Clim.* **13** 4366–77

[35] Goswami B N, Wu G and Yasunari T 2006 The annual cycle, intraseasonal oscillations, and roadblock to seasonal predictability of the Asian summer monsoon *J. Clim.* **19** 5078–99

[36] Li T, Zhang Y, Chang C and Wang B 2001 On the relationship between Indian Ocean sea surface temperature and Asian summer monsoon *Geophys. Res. Lett.* **28** 2843–6

[37] Jungclaus J H *et al* 2010 Climate and carbon-cycle variability over the last millennium *Clim. Past* **6** 723–37

[38] Roeckner E *et al* 2003 *The Atmospheric General Circulation Model ECHAM5—Part I: Model Description. Report No. 349* (Hamburg: Max Planck Institute for Meteorology)

[39] Hagemann S, Arpe K and Roeckner E 2006 Evaluation of the hydrological cycle in the ECHAM5 model *J. Clim.* **19** 3810–27

[40] Raddatz T J *et al* 2007 Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century? *Clim. Dyn.* **29** 565–74

[41] Marsland S, Haak H, Jungclaus J, Latif M and Roske F 2003 The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates *Ocean Modell.* **5** 91–127

[42] Wetzel P *et al* 2006 Effects of ocean biology on the penetrative radiation in a coupled climate model *J. Clim.* **19** 3973–87

[43] Li J and Zhang L 2009 Wind onset and withdrawal of Asian summer monsoon and their simulated performance in AMIP models *Clim. Dyn.* **32** 935–68

[44] Crowley T J *et al* 2008 Volcanism and the little ice age *PAGES Newslett.* **16** 22–3

[45] Krivova N A and Solanki S K 2008 Models of solar irradiance variations: current status *J. Astrophys. Astron.* **29** 151–8

[46] Pongratz J, Reick C, Raddatz T and Claussen M 2008 A reconstruction of global agricultural areas and land cover for the last millennium *Glob. Biogeochem. Cycles* **22** GB3018

[47] www.mpimet.mpg.de/fileadmin/ozean/drg/web_page_millennium_nov09_model_details.pdf

[48] Leisch F 2004 FlexMix: a general framework for finite mixture models and latent class regression in R *J. Stat. Softw.* **11** 1–18

[49] Schwarz G 1978 Estimating dimension of a model *Ann. Stat.* **6** 461–4