Hydrology and Earth System Sciences

Open Access

Continental moisture recycling as a Poisson process

H. F. Goessling[1,2,*] **and C. H. Reick**[1]

[1]Max Planck Institute for Meteorology, Hamburg, Germany
[2]International Max Planck Research School on Earth System Modelling, Hamburg, Germany
[*]now at: Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

Correspondence to: H. F. Goessling (helge.goessling@awi.de)

Abstract. On their journey over large land masses, water molecules experience a number of precipitation–evaporation cycles (recycling events). We derive analytically the frequency distributions of recycling events for the water molecules contained in a given air parcel. Given the validity of certain simplifying assumptions, the frequency distribution of recycling events is shown to develop either into a Poisson distribution or a geometric distribution.

We distinguish two cases: in case (A) recycling events are counted since the water molecules were last advected across the ocean–land boundary. In case (B) recycling events are counted since the water molecules were last evaporated from the ocean. For case B we show by means of a simple scale analysis that, given the conditions on earth, realistic frequency distributions may be regarded as a mixture of a Poisson distribution and a geometric distribution. By contrast, in case A the Poisson distribution generally appears as a reasonable approximation. This conclusion is consistent with the simulation results of an earlier study where an atmospheric general circulation model equipped with water vapor tracers was used. Our results demonstrate that continental moisture recycling can be interpreted as a Poisson process.

1 Introduction

Since the pioneering studies on the isotopic composition of precipitation and moisture recycling in the Amazon basin in the late 1970s (Salati et al., 1979; Lettau et al., 1979), it has become conventional wisdom that tropical forests maintain a substantial fraction of their precipitation by their own evaporation (including plant transpiration). That continental evaporation contributes a large fraction to continental precip-

itation has been confirmed by a large number of studies since the 1970s (e.g. Brubaker et al., 1993; Eltahir and Bras, 1994; Numaguti, 1999; van der Ent et al., 2010; Goessling and Reick, 2013). Moreover, it has been shown by means of climate model simulations that reduced continental evaporation rates, for example, due to deforestation, result in less continental precipitation (e.g. Shukla and Mintz, 1982; Shukla et al., 1990; Werth and Avissar, 2002; Goessling and Reick, 2011). Recently, Spracklen et al. (2012) found evidence also from observations that moisture recycling is an important factor for the generation of continental rainfall.

So far, most studies dealing with continental moisture recycling focussed on determining the evaporative source regions of precipitation, for example, the fraction of continental moisture in precipitation, using either atmospheric general circulation models equipped with passive water vapour tracers (Numaguti, 1999; Bosilovich et al., 2002; Goessling and Reick, 2013), or using reanalysis data together with diagnostic moisture tracing algorithms (Yoshimura et al., 2004; van der Ent et al., 2010). Numaguti (1999) (henceforth N99) distinguished moisture not only by source region (continental versus oceanic), but distinguished the continental fraction further according to the number of recycling events the water molecules have experienced since they last evaporated from the ocean, where a recycling event comprises precipitation on and subsequent re-evaporation from land. N99 noticed in one case that "from the first through the fourth generation the precipitation amount of the child generation is about 0.6 times that of the parent generation". However, N99 did not attempt to explain what determines the shapes of the simulated frequency distributions.

In this paper we provide such an explanation. We show analytically that, given the validity of certain simplifying

assumptions, the frequency distribution of continental recycling events for the water molecules contained in a given air parcel attains either a Poisson distribution or a geometric distribution, depending on the assumptions.

2   Theory

The goal of the following derivations is to find simple analytical expressions for the frequency distribution of n, where n is the number of continental recycling events the water molecules contained in an air parcel have experienced either (A) since they were last advected across an ocean–land boundary, or (B) since they last evaporated from the ocean. To achieve this goal, we first consider how in general a moisture species i is transported in the atmosphere and how it is exchanged between the atmosphere and the surface (including the subsurface).

2.1   Exact transport equations

For convenience, in the following derivations we consider only one of the two horizontal dimensions, the second one being analogous. The partial differential equation that governs the temporal evolution of the specific concentration q_i [kg kg^{-1}] of moisture species i reads

$$\frac{\partial (\rho\, q_i)}{\partial t} + \frac{\partial (\rho\, q_i\, u)}{\partial x} + \frac{\partial (\rho\, q_i\, w)}{\partial z} = S_i, \tag{1}$$

where ρ is air density [kg m^{-3}], t is time [s], u [m s^{-1}] is wind speed in the horizontal dimension x [m], w [m s^{-1}] is wind speed in the vertical dimension z [m], and S_i [kg m^{-3}s] is a non-advective source/sink term.

Following Goessling and Reick (2013), the vertical integral of Eq. (1) can be written in terms of the effective speed u_i^{eff} [m s^{-1}] at which q_i is horizontally transported. Using the notation $\widehat{\chi} = \int_0^\infty \rho\, \chi\, \mathrm{d}z$ (where χ can be any variable), with $u_i^{\text{eff}} = \widehat{q_i u}/\widehat{q_i}$ vertical integration gives

$$\frac{\partial \widehat{q_i}}{\partial t} + \frac{\partial \left(\widehat{q_i}\, u_i^{\text{eff}}\right)}{\partial x} = E_i - P_i, \tag{2}$$

where we substituted the vertical integral of S_i by the difference between surface evaporation E_i [kg m^{-2} s^{-1}] and precipitation (including dew formation) P_i [kg m^{-2} s^{-1}]. We can rewrite Eq. (2) such that it resembles a Lagrangian formulation:

$$\frac{\mathrm{d}\widehat{q_i}}{\mathrm{d}t} = E_i - P_i - \widehat{q_i}\, \frac{\partial u_i^{\text{eff}}}{\partial x} \tag{3}$$

with $\mathrm{d}\widehat{q_i}/\mathrm{d}t = \partial \widehat{q_i}/\partial t + u_i^{\text{eff}}(\partial \widehat{q_i}/\partial x)$.

2.2   The "well-mixed" assumption

While Eqs. (2) and (3) are still exact, we now make the first approximation by assuming vertically well mixed conditions,

i.e. we assume that

$$f_i(t,x,z_1) = f_i(t,x,z_2) \quad \text{for all} \ \ t, x, z_1, z_2, \tag{4}$$

where $f_i = q_i/q$ (and also, under well-mixed conditions, $f_i = \widehat{q_i}/\widehat{q}$) and $q = \sum_i q_i$ is total specific moisture. This implies two simplifications, namely that (I) the species-dependent effective transport speed becomes $u_i^{\text{eff}} = u^{\text{eff}} = \widehat{qu}/\widehat{q}$ for all moisture species, and that (II) the precipitation term becomes $P_i = f_i P$ where $P = \sum_i P_i$ is total precipitation (plus dew formation). For a comprehensive analysis of the errors that are introduced when well-mixed conditions are assumed in realistic situations, we refer the reader to Goessling and Reick (2013).

Since Eq. (3) likewise holds for q (omitting all indices i), we can use Eq. (3) to derive the Lagrangian time derivative of $f_i = \widehat{q_i}/\widehat{q}$. Under well-mixed conditions we obtain the very simple expression

$$\frac{\mathrm{d}f_i}{\mathrm{d}t} = \frac{\left(E_i - f_i\, P - \widehat{q_i}\, \frac{\partial u^{\text{eff}}}{\partial x}\right) \widehat{q} - \left(E - P - \widehat{q}\, \frac{\partial u^{\text{eff}}}{\partial x}\right) \widehat{q_i}}{\widehat{q}^2}$$

$$= \left(\widetilde{f_i} - f_i\right) \frac{E}{\widehat{q}}, \tag{5}$$

where E is total surface evaporation and $\widetilde{f_i} = E_i/E$. Note that P has dropped out here.

2.3   Species definition and the "steady-state" assumption

In the above derivations we have not specified how the moisture species i are defined, but required only that the only sources and sinks of the moisture species are those associated with surface evaporation and precipitation (including dew formation). One possible way to define different moisture species is to distinguish them according to their non-overlapping evaporative source region. In this case $\widetilde{f_i}$ would be either 1 or 0, depending on whether or not the considered air mass is located above the source region corresponding to species i. Instead, in the following we consider the case where moisture species are defined according to the number n of recycling events the water molecules have experienced on their journey over the land masses.

As adumbrated at the beginning of this section, there are two natural alternatives how one can define n:

- case A: n is the number of recycling events a water molecule has experienced since it was last advected across an ocean–land boundary.

- case B: n is the number of recycling events a water molecule has experienced since it last evaporated from the ocean.

These two cases can roughly be characterised as intra-continental (A) and inter- or pan-continental (B) moisture recycling because in case B the recycling events can take place on different continents while this is forbidden in case A.

In either case $\widetilde{f}_n$ over land depends on the composition of the surface reservoir from which the water evaporates. The composition of the surface reservoir is determined by the composition of antecedent precipitation events and the intermediate redistribution within the reservoir. However, it is possible to circumvent an explicit treatment of the surface reservoir by assuming that over land

$$\widetilde{f}_n = f_{n-1}, \quad n \in \mathbb{N} \tag{6a}$$
$$\widetilde{f}_0 = 0. \tag{6b}$$

While Eq. (6b) is by definition of n generally valid over land, Eq. (6a) constitutes an approximation that becomes exact in two limit cases: the first one is given when evaporation is fed by water that precipitated immediately before re-evaporation. In reality, evaporation from intercepted water may often be close to this situation. The second limit case is given when the system is in steady state, meaning that the atmospheric composition with respect to the moisture fractions f_n is temporally constant. When this is the case, the surface reservoir attains the same composition as the atmosphere above and, hence, the approximation defined by Eq. (6a) becomes exact. In reality, conditions that are not too far from such a steady-state situation may be given in tropical regions with weak variability of the meteorological conditions both at the daily and the seasonal timescale. In the following we denote the simplification given by Eq. (6a) the "steady-state" assumption, keeping in mind that it is exact not only in steady state but also if evaporation is fed by water that precipitated immediately before.

2.4 The Poisson distribution as solution over land

With our specification of moisture species according to the number of continental recycling events (n) and with the "steady-state" assumption, which beside the "well-mixed" assumption constitutes the second approximation we invoke, Eq. (5) becomes over land

$$\frac{\mathrm{d}f_n}{\mathrm{d}t} = (f_{n-1} - f_n)\frac{E_{\mathrm{con}}}{\widehat{q}}, \quad n \in \mathbb{N} \tag{7a}$$
$$\frac{\mathrm{d}f_0}{\mathrm{d}t} = -f_0\frac{E_{\mathrm{con}}}{\widehat{q}}, \tag{7b}$$

where E_{con} is continental evaporation. This system of ordinary differential equations has the following non-stationary solution:

$$f_n = \frac{\lambda^n}{n!}\mathrm{e}^{-\lambda}, \quad n \in \mathbb{N}_0, \tag{8}$$

where λ grows with time t:

$$\lambda(t) = \lambda(t_0) + \int_{t_0}^{t} \frac{E_{\mathrm{con}}(t')}{\widehat{q}(t')}\,\mathrm{d}t'. \tag{9}$$

Equation (8) is a Poisson distribution, characterised by the only parameter λ which also is the mean value $\langle n \rangle = \sum_{n=0}^{\infty}(n f_n)$ and the variance of the distribution. For Eq. (8) to be the actual solution of Eq. (7), it is necessary and sufficient that the initial state of the frequency distribution is a Poisson distribution, i.e. $f_n(t_0)$ must follow Eq. (8) with an arbitrary initial value $\lambda(t_0)$.

Whether $f_n(t_0)$ is Poisson distributed depends on which of the cases A and B is considered. In case A, $f_n(t_0)$ is indeed a Poisson distributed with t_0 being the moment of advection across the ocean–land boundary and $\lambda(t_0)$ being equal to zero. This simply means that at $t = t_0$ all water molecules have experienced zero recycling events ($f_0(t_0) = 1$, $f_n(t_0) = 0 \,\forall\, n > 0$, which is a special case of Eq. 8). After $t = t_0$, f_n develops as a Poisson distribution with λ growing according to Eq. (9) as long as the parcel's trajectory stays over land. Since Eq. (7) only holds over land, the solution loses validity when the parcel eventually leaves the continent. When the parcel is again advected across an ocean–land boundary, the solution regains validity starting with $\lambda = 0$, and so forth.

2.5 Extension to the ocean (case B) and the high intensity limit

In contrast to case A, in case B there is no moment in time at which the frequency distribution of n is set to an initial Poisson distribution; a parcel crossing the ocean–land boundary contains water molecules with $n > 0$. This reflects the fact that some molecules have not been newly evaporated from the ocean but experienced recycling events earlier and then traversed the ocean without being lost from the atmosphere as precipitation. The prerequisites that are necessary for the specific solution given by Eq. (8) to be the actual solution to Eq. (7) are thus not given anymore.

While Eqs. (7a) and (7b) hold only for continental evaporation (given the "steady-state" assumption), which is sufficient for case A, for case B it is necessary to extend the equations so that they can also be applied over the ocean:

$$\widetilde{f}_n = \begin{cases} f_{n-1} & \text{over land} \\ 0 & \text{over the ocean} \end{cases}, \quad n \in \mathbb{N} \tag{10a}$$

$$\widetilde{f}_0 = \begin{cases} 0 & \text{over land} \\ 1 & \text{over the ocean.} \end{cases} \tag{10b}$$

Entering this into Eq. (5) gives as extension of Eqs. (7a) and (7b) to the ocean

$$\frac{\mathrm{d}f_n}{\mathrm{d}t} = (f_{n-1} - f_n)\frac{E_{\mathrm{con}}}{\widehat{q}} - f_n\frac{E_{\mathrm{oce}}}{\widehat{q}}, \quad n \in \mathbb{N} \tag{11a}$$

$$\frac{\mathrm{d}f_0}{\mathrm{d}t} = -f_0\frac{E_{\mathrm{con}}}{\widehat{q}} + (1 - f_0)\frac{E_{\mathrm{oce}}}{\widehat{q}} \tag{11b}$$

where E_{oce} is total oceanic evaporation.

Still assuming vertically well mixed and steady-state conditions, we now show that there are two limit cases for which

also in case B the distribution of f_n can be obtained analytically. To this end we consider the integrated intensity λ^*, defined as

$$\lambda^* = \int_{t_0 - T}^{t_0} (E/\widehat{q})\, \mathrm{d}t, \tag{12}$$

where T is the typical time it takes atmospheric air to traverse either an ocean (T_oce) or a continent (T_con). λ^* relates the amount of evaporation, integrated along a trajectory from one coast to the next over an ocean or a continent, to the vertically integrated atmospheric moisture content ("precipitable water"), and thereby measures to what extent atmospheric moisture is replaced by newly evaporated moisture while the air is transported from one coast to the next.

The first limit case is the high intensity limit $\lambda^* \gg 1$. In this case, while travelling over an ocean (almost) all water molecules originally contained in the air are replaced by newly evaporated water from the ocean for which by definition $n = 0$ (note that, because the magnitude of $\widehat{q}$ typically does not change, in this limit case it also holds that precipitation and evaporation are of the same order). Accordingly, $f_0 \approx 1$ and $f_{n>0} \approx 0$ at the ocean–land boundary so that in the high intensity limit case B becomes equivalent to case A. This implies that the distribution of f_n evolves as a Poisson distribution with growing mean value $\langle n \rangle = \lambda(t)$ (Eqs. 8 and 9) until, finally, the considered air parcel leaves the continent.

2.6 The geometric distribution as stationary solution in the low intensity limit

We now consider the low intensity limit which is given if $\lambda^* \ll 1$. In this limit case, during the travel from one coast to the next over an ocean or a continent, the distribution of n is not much affected by surface evaporation, and f_n becomes largely determined at timescales larger than T. In fact, it is now possible to treat oceanic and continental evaporation as if they occurred simultaneously. In mathematical terms this is consistent with the limit $T \to 0$, which is one way how to arrive at the low intensity limit. Treating oceanic and continental evaporation simultaneously requires one to weight E_con and E_oce in Eqs. (11a) and (11b) according to T_con and T_oce. Thus replacing E_con by $\bar{E}_\mathrm{con} = (T_\mathrm{con}/(T_\mathrm{con}+T_\mathrm{oce}))E_\mathrm{con}$ and E_oce by $\bar{E}_\mathrm{oce} = (T_\mathrm{oce}/(T_\mathrm{con}+T_\mathrm{oce}))E_\mathrm{oce}$, and with $\bar{E} = \bar{E}_\mathrm{con} + \bar{E}_\mathrm{oce}$, Eqs. (11a) and (11b) become

$$\frac{\mathrm{d}f_n}{\mathrm{d}t} = (r\, f_{n-1} - f_n)\, \frac{\bar{E}}{\widehat{q}}, \quad n \in \mathbb{N} \tag{13a}$$

$$\frac{\mathrm{d}f_0}{\mathrm{d}t} = (1 - r - f_0)\, \frac{\bar{E}}{\widehat{q}}, \tag{13b}$$

where $r = \bar{E}_\mathrm{con}/\bar{E}$ is the fraction of surface evaporation entering the considered air parcel that stems from land (averaged over a path that is long enough to contain many ($\gg 1$)

alternating segments of land and ocean). Note that for $r \to 1$, which implies that $\bar{E} \to E_\mathrm{con}$, Eqs. (13a) and (13b) reduce to Eq. (7).

In contrast to Eqs. (7a) and (7b), given $\bar{E}/\widehat{q} \neq 0$, Eqs. (13a) and (13b) have a stationary solution. The unique steady state is

$$f_n = r^n\, (1 - r), \quad n \in \mathbb{N}_0 \tag{14}$$

which can easily be shown to be stable. Equation (14) is a geometric distribution with mean value $\langle n \rangle = r/(1 - r)$.

The steady-state solution given by Eq. (14) (not to be confused with the "steady-state" assumption formulated in Sect. 2.3), i.e. the geometric distribution, is attained and permanently maintained if r is globally constant. This condition is met if the low intensity limit (Eq. 12) is formulated even more stringent as $\lambda^*_\mathrm{global} = \int_{t_0 - T_\mathrm{global}}^{t_0} (E/\widehat{q})\, \mathrm{d}t \ll 1$, where T_global is the typical time it takes air to travel global distances. In this case, the distribution of f_n is a globally uniform geometric distribution with mean value $\langle n \rangle = r_\mathrm{global}$, where the latter is global land evaporation divided by total global evaporation.

It is worth mentioning that the geometric distribution obtained here bears a relation to the "recycling multiplier" defined in Savenije (1995) and van der Ent et al. (2010) as $m_c = f_0^{-1}$. van der Ent et al. (2010) argued that "when integrated over a year and all continental areas the multiplier is [. . .] the average number of times a water particle has sequentially fallen on the continent", meaning that $m_c = \langle n \rangle + 1$ and, thus, $\langle n \rangle = f_0^{-1} - 1$. This conforms precisely to the mean value of the geometric distribution (Eq. 14).

3 Stochastic interpretation

By the foregoing considerations we derived equations for the distribution of recycling events from the macroscopic transport equations of water vapour. In the present section the stochastic nature of the underlying molecular processes is elucidated.

3.1 Master equations and transition matrices

Equations (7a), (7b), (13a) and (13b) form the master equations of continuous-time Markov processes with discrete sets of states in a half-infinite range ($n \in \{0, 1, 2, \ldots, \infty\}$). These master equations are "gain–loss equation[s] for the probabilities of the separate states n" (van Kampen, 2010) obeying the general form

$$\frac{\mathrm{d}p_n(t)}{\mathrm{d}t} = \sum_{n' \neq n} \left\{ W_{nn'}\, p_{n'}(t) - W_{n'n}\, p_n(t) \right\}, \tag{15}$$

where $p_n \in [0, 1]$ is the probability that the process is in state n and $W_{nn'} \in [0, \infty)$ is the probability per unit time to jump from state n' to state n given the system is in state n' (i.e. the

transition rate). Defining further the decay rate of state n as $W_{nn} = -\sum_{n' \neq n} W_{n'n}$, which implies that $W_{nn} \in (-\infty, 0]$, Eq. (15) can be written in the compact form

$$\frac{\mathrm{d}\boldsymbol{p}(t)}{\mathrm{d}t} = \mathbf{W}\boldsymbol{p}(t),$$
(16)

where $\boldsymbol{p}(t)$ is the vector of the state probabilities $p_n(t)$ and $\mathbf{W}$ is the transition rate matrix (also known as Q-matrix) containing the elements $W_{nn'}$ with the first subscript giving the row and the second subscript giving the column (compare van Kampen, 2010).

The water contained in the atmosphere and exchanged with the surface is not a continuum but consists of single water molecules. Within the scope of the above approximations (Sects. 2.2–2.6), the fate of each individual water molecule in terms of recycling events can be interpreted as a single realisation of a stochastic process described by Eq. (16). Given the enormous number of water molecules contained in a sufficiently narrow atmospheric column (where sufficiently means very small compared to the considered horizontal scales) and the fact that relative sampling errors tend to decrease with the inverse of the square root of the sample size (a consequence of the central limit theorem; e.g. van Kampen, 2010), the probabilities $\boldsymbol{p}(t)$ in Eq. (16) can be equated with relative frequencies $\boldsymbol{f}(t)$. Hence, Eqs. (7a) and (7b) can be written in the form of Eq. (16) with the transition matrix

$$\mathbf{W}_{\mathrm{A}} = \begin{pmatrix} -I & 0 & 0 & 0 & \cdots \\ I & -I & 0 & 0 & \cdots \\ 0 & I & -I & 0 & \cdots \\ 0 & 0 & I & -I & \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix},$$
(17)

where $I = E_{\mathrm{con}}/\hat{q}$ is the intensity of the process, and Eqs. (13a) and (13b) can be written in the form of Eq. (16) with the transition matrix

$$\mathbf{W}_{\mathrm{B}} = \begin{pmatrix} -rI' & (1-r)I' & (1-r)I' & (1-r)I' & \cdots \\ rI' & -I' & 0 & 0 & \cdots \\ 0 & rI' & -I' & 0 & \cdots \\ 0 & 0 & rI' & -I' & \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix},$$
(18)

where $I' = \bar{E}/\hat{q}$ is the intensity of the process. Note that for $r \to 1$, which implies that $\bar{E} \to E_{\mathrm{con}}$ and thus $I' \to I$, $\mathbf{W}_{\mathrm{B}}$ reduces to $\mathbf{W}_{\mathrm{A}}$.

3.2 Recycling as a Poisson process

$\mathbf{W}_{\mathrm{A}}$ is known to be the transition matrix of a Poisson process (van Kampen, 2010). Thus, continental moisture recycling as described by Eqs. (7a) and (7b) can be interpreted as a nonhomogeneous Poisson process, where nonhomogeneous means that the intensity $I = E_{\mathrm{con}}/\hat{q}$ of the process varies as E_{con} and $\hat{q}$ vary with time. The Poisson process is a simple

counting process resembling a "random walk with steps to the right alone, but at random times" (van Kampen, 2010).

To illustrate how the Poisson process manifests itself in continental moisture recycling, let us consider the simple case where the rate of evaporation into the air column E_{con} and the precipitable water $\hat{q}$ remain constant, that is, the case of a homogeneous Poisson process. At the moment the air crosses the ocean–land boundary all water molecules have experienced zero recycling events – either a consequence of considering case A or, if case B is considered, a consequence of the high intensity limit. From that moment on every single water molecule undergoes a series of recycling events. Because of its Poissonian nature, this process has the following properties: the time between two subsequent recycling events (the interarrival time) is exponentially distributed. The interarrival time is independent of the number of already experienced recycling events (as can be seen from the identical diagonal elements of $\mathbf{W}_{\mathrm{A}}$). This memorylessness is a manifestation of the Markovian nature of the Poisson process. Finally, the number of recycling events in any sets of disjunct time intervals are independent and Poisson distributed, depending only on the length of the interval. These Poissonian properties (e.g. Ross, 1983) remain unchanged for the nonhomogeneous case, except that time t must be replaced by the integrated intensity $\lambda' = \int_{t_0}^{t} (E_{\mathrm{con}}/\hat{q}) \, \mathrm{d}t$ where t_0 is the moment at which the air enters the continent.

The interpretation as a Poisson process remains also valid in the presence of runoff. The latter implies that some of the realisations of the Poisson process are terminated, but since the formation of runoff is not selective with respect to the number of experienced recycling events of the water molecules (for the same reason why precipitation drops out in Eq. 5), the remaining molecules constitute an unbiased sample of the same Poisson processes.

As discussed in Sect. 2.3 our derivation of the master equation is valid either by assuming instantaneous re-evaporation after precipitation, or for a temporally constant species composition. The first assumption can be violated once the soil water reservoir is included in the description of moisture recycling. Nevertheless, this does not invalidate the interpretation as Poisson process (although the associated master equation would look different): in the presence of a soil reservoir each single water molecule still undergoes a Poisson process as described above but with the additional feature that the process is halted (i.e. λ' is not further advanced) as long as the molecule resides in the soil. Thereby, the set of molecules in an atmospheric column does generally not constitute a sample of the same Poisson processes anymore because, generally, λ' differs, i.e. the "time" (more precisely: the integrated intensity) they have been subject to a Poisson process differs between the molecules. That is, despite the Poissonian character of the process, the moisture fractions f_n for an atmospheric column are generally not Poisson distributed anymore. However, if the atmospheric composition with respect to f_n is temporally constant (the second way to arrive at

the "steady-state" assumption needed for our derivation), the Poisson distribution is restored because water molecules retained in the ground will, at the moment they evaporate back into the atmosphere, join water molecules with identical λ'.

As just described, the invalidity of the "steady-state" assumption still allows for an interpretation of continental moisture recycling as a Poisson process on a molecule by molecule basis. Strictly speaking, the same can not be said of the "well-mixed" assumption. Without the latter being valid, the water molecules have to be traced three-dimensionally through the atmosphere and precipitation-evaporation events (i.e. recycling events) become deterministic in the sense that they can not be described by probabilistic transition rates anymore. To make a Poissonian interpretation nevertheless possible, air parcels of some minimum size (allowing for the strong law of large numbers to be applicable) have to be considered as sampling units. This however leads to a strongly nonhomogeneous Poisson process where the intensity I is zero most of the time and becomes large when the parcel is intermittently part of a rain generating cloud. While this may enable the interpretation as a Poisson process on a molecule by molecule (or parcel by parcel) basis, violations of the "well-mixed" assumption will result in non-Poisson distributed recycling frequencies when whole atmospheric columns are considered because these comprise mixtures of water molecules with different λ'.

The preceding considerations can be summarised as follows. The Poissonian nature of continental moisture recycling can be discerned even if the "steady-state" assumption and/or the "well-mixed" assumption are violated. Their validity is however necessary for an unequivocal interpretation of continental moisture recycling as a Poisson process.

3.3 Recycling as an analog of a Bernoulli trials process

Part of the transition matrix $\mathbf{W}_B$ has a similar structure as $\mathbf{W}_A$ but it does not describe a Poisson process; it is the transition matrix of a more general continuous-time Markov process, still with a discrete set of states in a half-infinite range. We will show in the following that moisture recycling as described by $\mathbf{W}_B$, and thus by Eqs. (13a) and (13b), can be interpreted as a continuous-time analog of a Bernoulli trials process.

Let us for the moment assume that precipitation and evaporation of water molecules occur only in immediate pairs, both over land and over the ocean. In particular over the ocean this assumption is not realistic, but it facilitates the interpretation and is not required for the validity of the mathematical description of the process by Eqs. (13a) and (13b).

The main difference between $\mathbf{W}_B$ and $\mathbf{W}_A$ are the nonzero transition rates in the first row of $\mathbf{W}_B$. These entries imply that a water molecule with $n = i$ can transition not only to $n = i + 1$ (by experiencing a continental recycling event), but also to $n = 0$. The latter happens when a water molecule is "recycled" over the ocean. That is, the molecules

still undergo a "random walk with steps to the right at random times" (i.e. a Poisson process) as long as they hit land, but, at every transition, with a chance of $(1 - r)$ they are "reset" to $n = 0$ instead of advancing to $n = i + 1$. This ever-present probability to transition back to $n = 0$ is the reason why Eqs. (13a) and (13b) have a stable steady state, i.e. why the moisture fractions f_n run into a stationary distribution, while this is not the case for the pure Poisson process described by Eqs. (7a) and (7b).

The process described by Eqs. (13a) and (13b) and $\mathbf{W}_B$ is closely related to a Bernoulli trials process. The latter is "a sequence of chance experiments such that each experiment has two possible outcomes, success and failure, and the probabilities for success (r') and failure ($1 - r'$) are the same for each experiment" (modified from Grinstead and Snell, 1997). If one asks for the probability $p_{n'}$ that the first failure occurs after exactly n' successes, the corresponding probability distribution is

$$p_{n'} = (r')^{n'} (1 - r') \tag{19}$$

which is a geometric distribution. Obviously, with $n' = n$ (the number of continental recycling events a water molecule has experienced since it evaporated from the ocean) and with $r' = r$ (the fraction of evaporation that stems from land), Eq. (19) is identical with Eq. (14). That is, in the steady state of Eqs. (13a) and (13b) where the continuous-time process is deprived of its time dependence, moisture recycling in case B and in the low intensity limit can be considered a continuous-time analog of a Bernoulli trials process.

We point out that the Poissonian nature of moisture recycling is lost only in the practically not attainable low intensity limit. In reality, the location of an air parcel can always be assigned either to the continents or to the ocean, and for the continental parts of a parcel trajectory the process of moisture recycling remains essentially a Poisson process as described in Sect. 3.2.

4 Qualitative assessment of the involved assumptions

4.1 The intensity limits

While the intensity limits are irrelevant for case A, for case B we have shown the following: (I) in the high intensity limit the frequency distribution of n over land is a Poisson distribution with a mean value that grows as continental moisture evaporates into the considered air parcel (case B behaves like case A for the continental parts of the trajectories). (II) In the low intensity limit the frequency distribution of n is a stationary geometric distribution. Note that both of these solutions also require that the "well-mixed" assumption and the "steady-state" assumption are valid. We now investigate to what extent the two intensity limits apply to real conditions on earth by means of simple scale analysis.

Neglecting dry continental regions and ice-covered surfaces, typical values of E range between $\sim 1\,\mathrm{kg\,m^{-2}\,d^{-1}}$ in the high latitudes and $\sim 5\,\mathrm{kg\,m^{-2}\,d^{-1}}$ in the tropics (Fig. 7.26 in Peixoto and Oort, 1992). Because typical values of $\widehat{q}$ range between $\sim 10\,\mathrm{kg\,m^{-2}}$ in the high latitudes and $\sim 50\,\mathrm{kg\,m^{-2}}$ in the tropics (Fig. 12.3c in Peixoto and Oort, 1992), the intensity itself can be considered to be spatially more or less uniform at $E/\widehat{q} \approx 0.1/d$. It is thus T which largely determines the order of λ^* (Eq. 12).

We now estimate T exemplarily for air traversing the North Atlantic and for air traversing the tropical Pacific. For the easterlies above the tropical Pacific with $u^{\mathrm{eff}} \approx 5\,\mathrm{m\,s^{-1}} \approx 500\,\mathrm{km\,d^{-1}}$ (Fig. 7.20a in Peixoto and Oort, 1992) and $\Delta x \approx 15\,000\,\mathrm{km}$ one arrives at $\lambda^* \approx 3$. For the westerlies above the North Atlantic with $u^{\mathrm{eff}} \approx 10\,\mathrm{m\,s^{-1}} \approx 1000\,\mathrm{km\,d^{-1}}$ (Fig. 7.20a in Peixoto and Oort, 1992) and $\Delta x \approx 5000\,\mathrm{km}$ one arrives at $\lambda^* \approx 0.5$. This suggests that the situation on earth is somewhere between the two intensity limits.

The realism of the intensity limits can also be assessed using the results of studies where continental precipitation recycling ratios (the fraction of continental moisture in precipitation, R_{c}) were quantified by means of numerical moisture tracing (Numaguti, 1999; Bosilovich et al., 2002; Yoshimura et al., 2004; van der Ent et al., 2010; Goessling and Reick, 2011, 2013). Because by definition $R_{\mathrm{c}} = \sum_{n=1}^{\infty} f_n = 1 - f_0$, in the high intensity limit R_{c} would drop on the way over ocean basins from non-zero values at the upwind coasts to zero at the downwind coasts. By contrast, in the low intensity limit R_{c} would not vary between the continents and the ocean.

The results of the simple scale analysis above are consistent with estimated patterns of R_{c}. Considering the situation along a latitude within the northern summer extratropical westerlies, R_{c} significantly increases from west to east over the continents and decreases from west to east over the ocean basins (e.g. Fig. 6a–d in Goessling and Reick, 2013). The fact that the recycling ratio does not stay constant along the westerlies implies that the low intensity limit is not valid. On the other hand, the high intensity limit is likewise invalid because the recycling ratio does by far not drop to zero over the ocean. The situation is slightly different in the tropics, in particular south of the equator. Air masses arriving at the eastern coasts of the continents located there – South America, Africa, and Australia – are characterised by recycling ratios that are very low compared to the maxima located close to western coasts. This suggests that here the high intensity limit is a reasonable first-order approximation.

The fact that conditions on earth are somewhere between the two intensity limits suggests that, if n is defined according to case B, real frequency distributions of n can be interpreted as mixtures of Poisson distributions and geometric distributions. Moreover, depending on location and season, real distributions may be closer to either of the two analytical solutions.

4.2 The "well-mixed" assumption

In both cases A and B our analytical solution for the frequency distribution of recycling events is based on the "well-mixed" assumption (Eq. 4) as well as the "steady-state" assumption (Eq. 6a). In the following we discuss qualitatively how realistic these assumptions are, focussing on case A and the Poisson distribution. The "well-mixed" assumption is discussed in this section and the "steady-state" assumption in the next section.

The applicability of the "well-mixed" assumption was investigated in depth by Goessling and Reick (2013) in the context of 2-D moisture tracing because the latter is based on the "well-mixed" assumption. It was found that well-mixed conditions are seldom present in the atmosphere. In particular in the tropics, where horizontal winds often blow in different directions at different heights (directional vertical shear), the 2-D approximation often leads to substantial errors. By contrast, in the extratropics errors associated with 2-D moisture tracing arise mostly from the neglect of fast recycling and are rather moderate.

This suggests that in tropical regions with strong directional vertical shear of the winds, for example in western Africa during northern summer, real frequency distributions of n may significantly deviate from Poisson distributions. One can imagine that vertically sheared winds act to mix air masses of different origin horizontally. Even if one assumes that those different air masses exhibit Poisson-distributed moisture fractions f_n before horizontal mixing takes place, the horizontal mixing results in non-Poisson-distributed f_n because linear combinations of Poisson distributions (with different mean values) are not Poisson distributions. If air masses with strongly different mean values mix horizontally, the resulting frequency distributions can even be multimodal. However, even in such cases it may be admissible to interpret the resulting frequency distributions as linear combinations of Poisson distributions, corroborating that the process by which these distributions are generated is still a Poisson process.

In the extratropics horizontal mixing of air masses is weaker and, thus, horizontal advection is captured relatively well by the 2-D approximation. This is true despite the fact that vertically well mixed conditions are typically not given (Goessling and Reick, 2013). We thus expect that deviations from well-mixed conditions in the extratropics do not strongly distort the resulting frequency distributions of n. However, it still depends on the second simplification, namely the "steady-state" assumption, whether real frequency distributions of n indeed resemble Poisson distributions.

4.3 The "steady-state" assumption

First of all, it is important to recall that what we term the "steady-state" assumption, given by Eq. (6a), is valid not

only if the atmospheric composition with respect to the moisture fractions f_n is temporally constant (i.e. in steady-state), but also if evaporation is fed by precipitation that occurred immediately before.

In reality, the latter condition (fast evaporation) is approximately fulfilled in situations where evaporation is drawn from the surface skin reservoir, i.e. water that has been intercepted by leaves or retained by the uppermost millimeters of the soil. Skin evaporation rates can be substantial, in particular in warm climates with weak seasonality, and in some cases contributes up to 50 % to the total annual evaporation (Savenije, 2004). By contrast, if water from the deep soil is transpired by plants, the water can be several months (or more) old. Particularly large amounts of old water are transpired by deep-rooted forests that are located in tropical wet-dry or tropical monsoon climates during the dry season, but also by temperate or boreal forests during summer. In this case Eq. (6a) can be a poor approximation.

However, even if old water is evaporated Eq. (6a) can still be exact, namely, if the precipitation that formed the old water had the same composition with respect to the moisture fractions f_n as the current atmospheric moisture. The "steady-state" assumption is thus a reasonable approximation also for regions with weak seasonality of the frequency distribution of f_n, which may be the case for example in tropical rainforest climates.

In conclusion, these qualitative considerations suggest that (in case A) significant deviations from Poisson distributions occur globally, but for different reasons in the tropics and in the extratropics. For the tropics we expect that deviations arise mainly due to the invalidity of the "well-mixed" assumption in combination with directional vertical shear of the horizontal winds, whereas the "steady-state" assumption may often be considered acceptable. The opposite is the case for the extratropics where we expect that the invalidity of the "well-mixed" assumption plays a minor role while the "steady-state" assumption may be the main cause of deviations.

5 Comparison with Numaguti (1999)

N99 used an atmospheric general circulation model equipped with water vapor tracers to simulate explicitly frequency distributions of continental recycling events. However, in that study the shapes of the resulting distributions were not investigated in much detail. In the following we re-examine the relevant parts of the N99 results to see if our theory provides an explanation for the simulated frequency distributions.

Figure 1 shows the frequency distribution of n from N99 for July precipitation in the eastern Tibet region. Instead of treating water evaporated from all ocean basins together, only water that originally evaporated either from the North Atlantic or from the North Indian Ocean is considered separately. Accordingly these simulation results can not be used
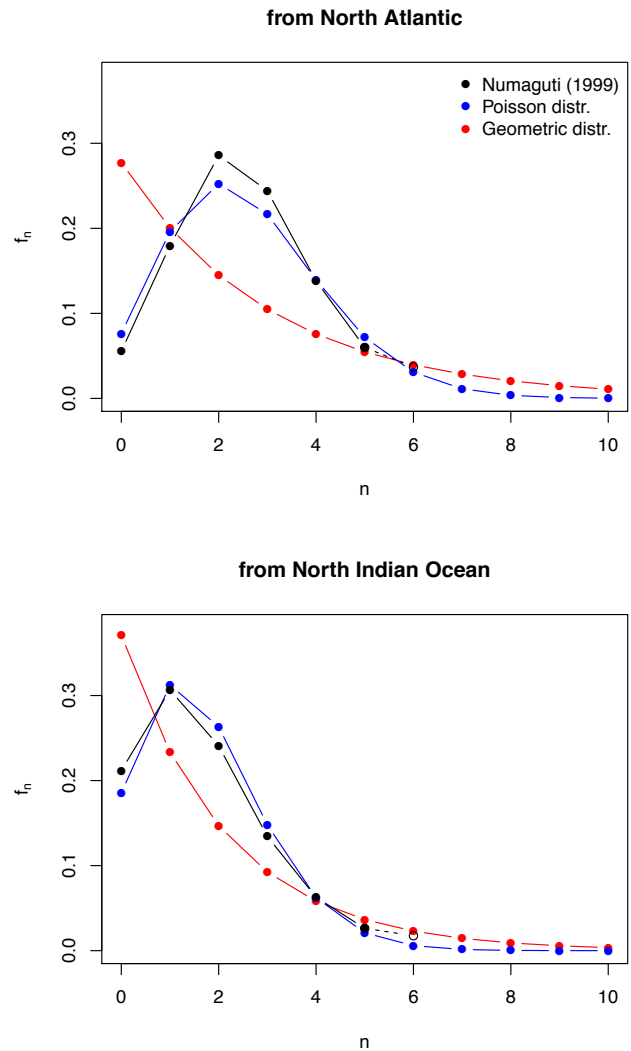

Fig. 1. Frequency distributions of the number of continental recycling events (n) for July precipitation in the eastern Tibet region. Top: precipitation stemming from the North Atlantic. Bottom: precipitation stemming from the North Indian Ocean. Black: simulated frequencies for $n = 0, 1, 2, 3, 4, 5$ and $n \geq 6$, derived from Fig. 14 in Numaguti (1999). Blue: fitted Poisson distributions with mean values $\lambda = 2.57$ (top) and $\lambda = 1.69$ (bottom). Red: fitted geometric distributions with mean values $r/(1-r) = 2.61$ (top) and $r/(1-r) = 1.70$ (bottom).

to investigate the adequacy of the intensity limits. Rather, by construction of the simulation set-up, in good approximation the data correspond to the high intensity limit, which is on the continents equivalent to case A (Sect. 2.5).

For comparison of the simulated frequency distributions with the theoretical solutions, Fig. 1 also shows fitted Poisson distributions and geometric distributions, where fitted simply means that they have the same mean values as the corresponding simulated distributions. Because in N99 water molecules with six or more recycling events are not further distinguished but only given by cumulative values for

$f_{n\geq 6}$ (see empty black circles in Fig. 1), it is not obvious how to derive mean values from the simulated frequency distributions. We decided to compute mean values separately for the Poisson distributions and the geometric distributions by distributing the unresolved part of the simulated frequencies such that they correspond to Poisson distributions and geometric distributions with parameters determined by the respective inverse cumulative probability density functions under the constraint $\sum_{n=6}^{\infty} f_n = f_{n\geq 6}$.

The Poisson distribution fitted in this way has a mean value $\langle n \rangle = 2.57$ for the moisture that has travelled the long distance from the North Atlantic, and a mean value $\langle n \rangle = 1.69$ for the moisture that has travelled the shorter way from the North Indian Ocean. The fitted mean values of the geometric distributions are slightly higher: $\langle n \rangle = 2.61$ (North Atlantic) and $\langle n \rangle = 1.70$ (North Indian Ocean). Because the results of N99 approximately correspond a priori to case A, it is not surprising that the simulated distributions are much closer to Poisson distributions than to geometric distributions (Fig. 1). Due to the construction of the simulation set-up, the comparison gives no indication of which distribution would better resemble actual distributions in case B. The similarity of the Poisson distributions with the simulated data does however suggest that violations of the "well-mixed" assumption and the "steady-state" assumption are relatively small.

For the precipitation stemming from the North Indian Ocean, N99 noticed that "from the first through the fourth generation the precipitation amount of the child generation is about 0.6 times that of the parent generation". While N99 made this statement independent of the considered month, we consider only the results for July. One may wonder how this exponential law for subsequent generations ($f_n = 0.6 \cdot f_{n-1}$) relates to our analytical solutions. And indeed one solution complies exactly with such a law, namely a geometric distribution with $\langle n \rangle = 1.5$, which is slightly less than the mean value $\langle n \rangle = 1.70$ of the fitted geometric distribution (Fig. 1, bottom). In case of the fitted Poisson distribution with $\langle n \rangle = 1.69$, the corresponding factors are dependent on n and amount to 0.84 (first to second generation), 0.56 (second to third generation), and 0.42 (third to forth generation), which considerably deviates from a constant factor of 0.6. However, we recalculated these factors from the N99 data for July and found that they are in fact closer to the ones "predicted" by the fitted Poisson distribution, namely 0.79 (first to second generation), 0.56 (second to third generation), and 0.47 (third to forth generation). This corroborates that, as expected, the Poisson distribution is a much better approximation than the geometric distribution in the considered case, which by construction is close to case A.

We conclude that our theory indeed provides an explanation for the frequency distributions simulated by N99: the distributions arise from the Poissonian nature of continental moisture recycling, the process that generates them.

6 Summary and conclusion

We have shown that under certain conditions continental moisture recycling resembles a Poisson process. This implies that, over land, the frequency distribution of the number of continental recycling events the water molecules contained in an atmospheric column have experienced, i.e. the distribution of the moisture fractions f_n, is a Poisson distribution. Two assumptions, namely, the "well-mixed" assumption and the "steady-state" assumption, are required for this solution if atmospheric moisture is a priori considered to be purely oceanic when the air crosses the ocean–land boundary (case A). If continental moisture is "allowed" to traverse the ocean (case B), Poisson distributions are attained only if a third condition is fulfilled, namely, in the high-intensity limit: all molecules of continental origin are replaced by evaporation from the ocean during the time it takes air to traverse an ocean. If the opposite is true, i.e. if the air traverses continents and ocean basins much faster than it would take to replace the atmospheric moisture by surface evaporation, the moisture fractions f_n attain a geometric distribution. In this (practically not attainable) case moisture recycling can be considered an analog of a Bernoulli trials process.

Simple scale analysis reveals that neither of the two intensity limits is adequate in the extratropics, suggesting that (in case B) real frequency distributions of n in the extratropics can be interpreted as mixtures of Poisson distributions and geometric distributions. In the tropics, however, the high intensity limit may be a reasonable first-oder approximation, leading to Poisson distributions of f_n if the "well-mixed" assumption and the "steady-state" assumption are valid. Moreover, we have elucidated that even with partial invalidity of these assumptions, continental moisture recycling can still be considered a Poisson process on a molecule by molecule basis. At the bottom line, it appears that continental moisture recycling is primarily Poissonian.

The comparison of our theoretical solutions with the simulation data of Numaguti (1999) corroborates our finding that continental moisture recycling can be interpreted as a Poisson process. However, the N99 results are not comprehensive enough to sufficiently assess the transferability of our theoretical solutions to real conditions on earth. Instead, our results should be confronted with new, more specific simulations in the spirit of N99. It would then also be possible to check the adequacy of the involved assumptions one by one, and to investigate how the solutions depend on the climatic conditions, i.e. the geographical location and the season.

Edited by: D. Gerten

References

Bosilovich, M. G., Sud, Y., Schubert, S. D., and Walker, G. K.: GEWEX CSE sources of precipitation using GCM water vapor tracers, Global Energy and Water Cycle Experiment NEWS, 3, 6–7, 2002.

Brubaker, K. L., Entekhabi, D., and Eagleson, P. S.: Estimation of continental precipitation recycling, J. Climate, 6, 1077–1089, 1993.

Eltahir, E. A. B. and Bras, R. L.: Precipitation recycling in the Amazon Basin, Q. J. Roy. Meteor. Soc., 120, 861–880, 1994.

Goessling, H. F. and Reick, C. H.: What do moisture recycling estimates tell us? Exploring the extreme case of non-evaporating continents, Hydrol. Earth Syst. Sci., 15, 3217–3235, doi:10.5194/hess-15-3217-2011, 2011.

Goessling, H. F. and Reick, C. H.: On the "well-mixed" assumption and numerical 2-D tracing of atmospheric moisture, Atmos. Chem. Phys., 13, 5567–5585, doi:10.5194/acp-13-5567-2013, 2013.

Grinstead, C. C. M. and Snell, J. L.: Introduction to Probability, 2nd Edn., American Mathematical Society, 1997.

Lettau, H., Lettau, K., and Molion, L. C. B.: Amazonia's hydrologic cycle and the role of atmospheric recycling in assessing deforestation effects, Mon. Weather Rev., 107, 227–238, 1979.

Numaguti, A.: Origin and recycling processes of precipitating water over the Eurasian continent: Experiments using an atmospheric general circulation model, J. Geophys. Res., 104, 1957–1972, 1999.

Peixoto, J. P. and Oort, A. H.: Physics of Climate, American Institute of Physics, New York, 1992.

Ross, S.: Stochastic Processes, Vol. 23, John Wiley & Sons, New York, 1983.

Salati, E., Dall'Olio, A., Matsui, E., and Gat, J. R.: Recycling of water in the Amazon basin: an isotopic study, Water Resour. Res., 15, 1250–1258, 1979.

Savenije, H.: New definitions for moisture recycling and the relationship with land-use changes in the Sahel, J. Hydrol., 167, 57–78, doi:10.1016/0022-1694(94)02632-L, 1995.

Savenije, H.: The importance of interception and why we should delete the term evapotranspiration from our vocabulary, Hydrol. Process., 18, 1507–1511, doi:10.1002/hyp.5563, 2004.

Shukla, J. and Mintz, Y.: Influence of land-surface evapotranspiration on the Earth's climate, Science, 215, 1498–1501, 1982.

Shukla, J., Nobre, C., and Sellers, P.: Amazon deforestation and climate change, Science, 247, 1322–1325, 1990.

Spracklen, D. V., Arnold, S. R., and Taylor, C. M.: Observations of increased tropical rainfall preceded by air passage over forests, Nature, 489, 282–285, doi:10.1038/nature11390, 2012.

van der Ent, R. J., Savenije, H. H. G., Schaefli, B., and Steele-Dunne, S. C.: Origin and fate of atmospheric moisture over continents, Water Resour. Res., 46, W09525, doi:10.1029/2010WR009127, 2010.

van Kampen, N.: Stochastic Processes in Physics and Chemistry, 3rd Edn., Elsevier, Singapore, 2010.

Werth, D. and Avissar, R.: The local and global effects of Amazon deforestation, J. Geophys. Res., 107, 8087, doi:10.1029/2001JD000717, 2002.

Yoshimura, K., Oki, T., Ohte, N., and Kanae, S.: Colored moisture analysis estimates of variations in 1998 Asian monsoon water sources, J. Meteorol. Soc. Jpn., 82, 1315–1329, doi:10.2151/jmsj.2004.1315, 2004.