



# Thermodynamic limits of hydrologic cycling within the Earth system: concepts, estimates and implications

A. Kleidon and M. Renner

Max-Planck-Institut für Biogeochemie, Jena, Germany

Correspondence to: A. Kleidon (akleidon@bgc-jena.mpg.de)

Received: 14 February 2013 – Published in Hydrol. Earth Syst. Sci. Discuss.: 12 March 2013

Revised: 18 June 2013 – Accepted: 22 June 2013 – Published: 23 July 2013

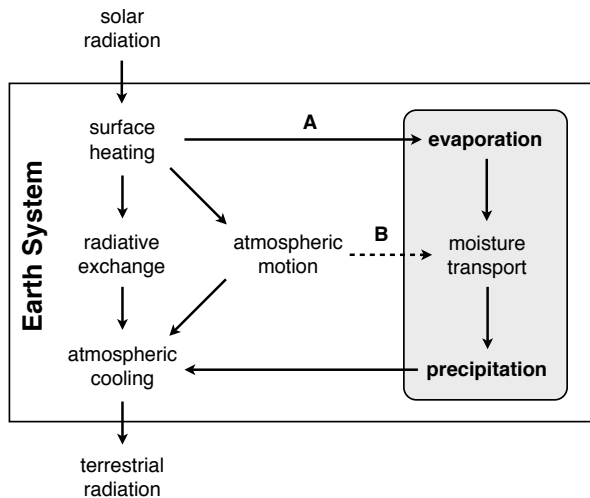
**Abstract.** The hydrologic cycle results from the combination of energy conversions and atmospheric transport, and the laws of thermodynamics set limits to both. Here, we apply thermodynamics to derive the limits of the strength of hydrologic cycling within the Earth system and about the properties and processes that shape these limits. We set up simple models to derive analytical expressions of the limits of evaporation and precipitation in relation to vertical and horizontal differences in solar radiative forcing. These limits result from a fundamental trade-off by which a greater evaporation rate reduces the temperature gradient and thus the driver for atmospheric motion that exchanges moistened air from the surface with the drier air aloft. The limits on hydrologic cycling thus reflect the strong interaction between the hydrologic flux, motion, and the driving gradient. Despite the simplicity of the models, they yield estimates for the limits of hydrologic cycling that are within the observed magnitude, suggesting that the global hydrologic cycle operates near its maximum strength. We close with a discussion of how thermodynamic limits can provide a better characterization of the interaction of vegetation and human activity with hydrologic cycling.

generated. Atmospheric motion, in turn, plays a vital role in transporting moisture from the places at which water evaporates to where it condenses. Hence, the hydrologic cycle shapes energy balances and interacts strongly with atmospheric motion and transport.

In this study we ask about the fundamental factors that limit the strength of hydrologic cycling within the Earth system. To establish such limits, we combine thermodynamics with a holistic view of how hydrologic cycling takes place within the Earth system. Thermodynamics describes general rules for the constraints, directions, and limits associated with energy conversions. A general result of thermodynamics is the Carnot limit, which describes the maximum rate by which a heating difference can be converted into mechanical work. This mechanical work is needed to maintain atmospheric motion, and motion is needed to maintain the ability to transport moisture. Several studies indicate that the atmospheric circulation operates near its thermodynamic limit (Paltridge, 1975, 1978; Lorenz et al., 2001; Kleidon et al., 2003, 2006), so that we can infer the strength of the associated heat and moisture transport by atmospheric motion from the assumption of maximum thermodynamic efficiency. The limit of hydrologic cycling that we explore here results from the general trade-off by which stronger hydrologic cycling involves more energy associated with phase transitions, so that less energy is available to drive atmospheric transport. By doing so, we do not need to look at the thermodynamics of atmospheric processes involved in hydrologic cycling in detail, as done by several previous studies on the thermodynamics of hydrologic cycling (Renno and Ingersoll, 1996; Pauluis and Held, 2002a,b; Konings et al., 2012). We rather start with the assumption of the Carnot limit in the context of the Earth system and estimate the maximum strength of

## 1 Introduction

The global hydrologic cycle plays a critical role in the Earth system (Chahine, 1992). The phase transitions of water between its liquid to gaseous form involve substantial amounts of energy, so that the fluxes of evaporation and precipitation strongly affect heating and cooling rates at the surface and within the atmosphere. These differences in heating and cooling form the basic drivers from which atmospheric motion is



**Fig. 1.** Schematic diagram to illustrate the primary factors that limit the strength of hydrologic cycling within the Earth system. The arrows indicate the flow of energy, from the incoming solar radiation at the top of the diagram which results mostly in the heating of the earth's surface, to the emission of terrestrial radiation, mostly from the atmosphere, to space at the bottom. Hydrologic cycling is embedded in these energy transfers (shown by the grey box) as it removes heat where water evaporates (arrow "A") and releases that heat when it condenses and results in precipitation. The differences in heating between the surface and the atmosphere generates atmospheric motion, which plays a critical role in transporting moisture from the locations at which water evaporates to the locations at which it condenses (as indicated by the dashed arrow, "B"). The limit to hydrologic cycling results from the tradeoff between paths "A" and "B", because the associated flows of energy are subject to energy conservation.

the hydrologic cycle as a result of the trade-off between energy being used for phase transitions and for the generation of atmospheric transport.

To explore the interplay between phase transitions and moisture transport in more detail, we consider a thought experiment in which we vary the magnitude of evaporation and consider its consequences in the context of the processes shown in Fig. 1. In this thought experiment, the surface is heated by absorption of solar radiation, terrestrial radiation is exchanged by absorption of solar radiation, terrestrial radiation is exchanged between the surface and the atmosphere, and the atmosphere cools by the emission of terrestrial radiation into space. The evaporation of water removes latent heat from the surface (arrow "A"), and this heat is released within the atmosphere when it condenses. From the differential heating between the surface and the atmosphere, atmospheric motion is generated and transports sensible as well as latent heat (arrow "B") from the heated surface to the cooler atmosphere.

When we now consider a case in which more water is evaporated at the surface (i.e., we strengthen arrow "A"), this would in steady state be accompanied by a greater rate of precipitation, which at first sight would seem to imply a greater

strength of hydrologic cycling. This greater strength, however, is associated with more evaporative cooling at the surface and more condensational heating in the atmosphere, so that the heating difference between the surface and the atmosphere would be reduced. Since motion is generated out of heating differences, this reduced heating difference should result in less motion and a lower ability to transport moisture (i.e., a weaker arrow "B") from the locations at which water evaporates to the locations at which it condenses. This transport is essential to maintain hydrologic cycling, as it is through atmospheric transport that the moistened air near the surface is replaced by drier air from above, and the condensed moisture that is removed by precipitation from the atmosphere is replaced by the moistened air from the surface. Hence, the increased transport of latent heat with a stronger hydrologic cycle and the reduced ability to transport due to the reduced heating difference should result in a trade-off that sets an upper limit on the strength of the hydrologic cycle. The limit to transport is then determined by thermodynamics, specifically the Carnot limit. The combination of this limit with the energy and transport requirements for hydrologic cycling then yields the maximum possible strength of hydrologic cycling within the Earth system.

The goal of this paper is to establish this maximum strength of hydrologic cycling from the thermodynamic constraints by which motion, and hence moisture transport, can be generated within the Earth system from its external forcing. We derive this limit from very simple energy balance models similar to those that were used to explore a related, proposed thermodynamic principle of Maximum Entropy Production (Lorenz et al., 2001; Kleidon, 2004), although we do not invoke this principle here. By using observations, we can then show that the current hydrologic cycle appears to operate near this limit.

In the following section, we first provide a brief introduction to thermodynamics, thermodynamic limits, and describe an overview of how we apply thermodynamics to the combined roles of energy balances and atmospheric transport for hydrologic fluxes. In Sect. 3, we set up two simple models to derive analytic expressions for the maximum strength of hydrologic cycling that are related to vertical convection and large-scale horizontal atmospheric transport. We identify the main sensitivities of these limits, derive global estimates, and compare these to observations. Our estimates are certainly strong simplifications and only consider first-order effects, so that we discuss the limitations of our models and the estimates, and compare our approach to previous work in Sect. 5. We then describe how this approach can provide a framework to better understand and characterize the effects of vegetation activity and human alterations on hydrologic cycling. We close with a brief summary and conclusion.

## 2 Thermodynamics of hydrologic cycling within the Earth system

The derivation of the maximum strength of hydrologic cycling requires some background on the relevant constraints on the processes that are shown in Fig. 1, particularly by the conservation of energy, and how thermodynamics sets limits to these processes. In the following, we first briefly describe the laws of thermodynamics and how thermodynamic limits emerge from these laws. We describe how heating gradients are generated within the Earth system from the solar radiative forcing that can be used to derive work, how this work is used to generate atmospheric motion, and how motion results in the transport of sensible heat. We then describe how hydrologic cycling is embedded in these processes. The application of thermodynamic limits to this description is then performed in the following section.

A summary of the variables used in the formulations is shown in Table 1. The variables use mostly standard notation, with the exception that the symbol  $J$  is used for radiative fluxes,  $J^s$  is used for entropy fluxes, and  $G$  for power (to avoid confusion with precipitation,  $P$ ). The subscripts used are summarized in Table 2. Unless otherwise noted, all properties are expressed per unit area. In the derivations below, the main assumptions that are being made are that the relevant energy-, mass-, and momentum balances are in steady state, that the atmosphere is fully absorbing so that the emission to space originates only from the top of the atmosphere, and that the non-linear relationships regarding emission of radiation and the saturation vapor pressure curve can be linearly approximated.

### 2.1 The laws of thermodynamics

Thermodynamics is a physical theory that describes the rules for energy conversions. For a complete accounting of these energy conversions, we consider a system with a well-defined boundary through which heat, mass and other physical properties are exchanged. The thermodynamic state of the system is characterized at least by its internal energy,  $U$ , and its entropy,  $S$ , which measures the dispersal of energy within the system.

The first law of thermodynamics describes that the change in internal energy of the system equals the net addition or removal of heat by exchange with the surroundings, and the net work done by (or on) the system. If we express these changes through time, the first law is expressed as

$$\frac{dU}{dt} = J_{\text{net}} - G + D, \quad (1)$$

where  $J_{\text{net}}$  is the net heat flux through the system boundary,  $G = dW/dt$  is the rate by which work is done by the system on its surroundings (or the physical power) to generate other forms of energy (e.g. potential or kinetic energy), and  $D$  is the dissipative heating that results from the conversions of

**Table 1.** Variables used in this study.

Symbol	Variable
$c_p$	specific heat capacity of air
$C_d$	drag coefficient
$D$	dissipation
$\Delta z$	height at which condensation occurs
$F$	force
$\Gamma$	adiabatic lapse rate
$g$	gravitational acceleration
$J$	heat flux
$J_{\text{net}}^s$	net entropy exchange
$\lambda E$	latent heat flux
$H$	sensible heat flux
$E$	evaporation
$P$	precipitation
$\lambda$	latent heat of vaporization
$G$	power, or generation rate of free energy
$q$	specific humidity
$\rho$	air density
$s$	slope of water vapor saturation curve
$S$	entropy
$\sigma$	entropy production
$T$	temperature
$T_h$	temperature of a hot reservoir
$T_c$	temperature of a cold reservoir
$U$	internal energy
$v$	velocity
$w$	exchange velocity

**Table 2.** Subscripts used for the main variables in this study.

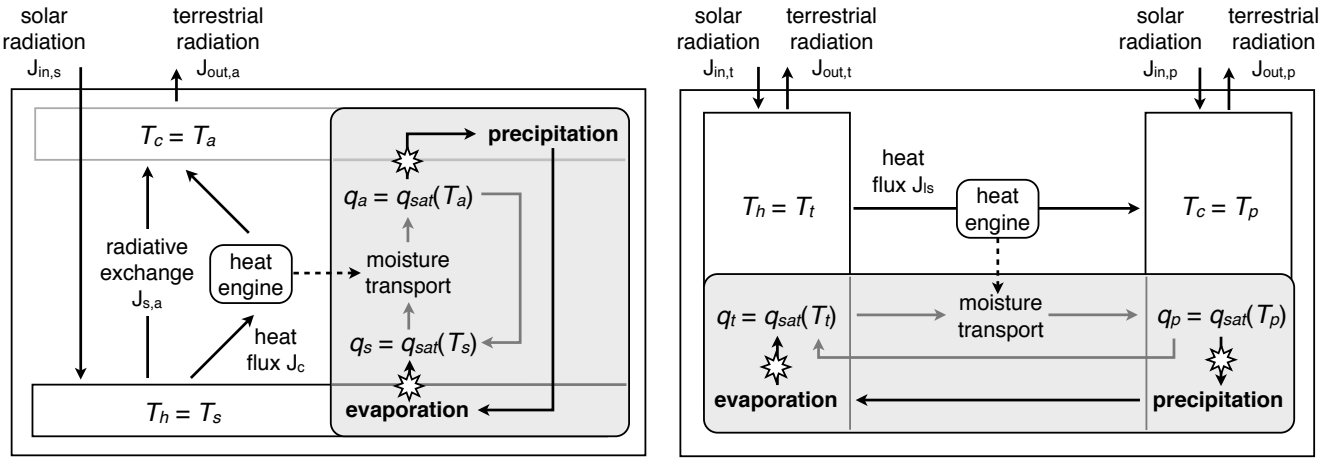
Subscript	Description	Examples
c	convection (vertical gradient)	$\sigma_c, E_c$
ls	large-scale (horizontal gradient)	$\sigma_{ls}, E_{ls}$
a	atmosphere	$J_{\text{out,a}}, T_a$
p	extratropical regions	$J_{\text{in,p}}, T_p$
s	surface	$J_{\text{in,s}}, T_s$
t	tropical regions	$J_{\text{in,t}}, T_t$
sh	sensible heat	$\sigma_{\text{sh}}, G_{\text{sh}}$
lh	latent heat	$\sigma_{\text{lh}}, G_{\text{lh}}$

other forms of energy back into heat within the system, for instance by friction.

The second law of thermodynamics requires that during energy conversions, energy is increasingly dispersed. This requirement is accounted for by the entropy budget of the system, which balances the changes in entropy of the system with the entropy produced by processes within the system,  $\sigma$ , with the entropy that is exchanged when energy (or mass) is exchanged with the surroundings,  $J_{\text{net}}^s$ :

$$\frac{dS}{dt} = \sigma - J_{\text{net}}^s. \quad (2)$$

The second law requires that  $\sigma \geq 0$ .



**Fig. 2.** Simple representation of hydrologic cycling resulting from (a, left) vertical differences and (right) large-scale horizontal differences in solar radiative heating. The first heating difference shown in (b, left) results from surface heating due to the absorption of solar radiation (lower box, with temperature  $T_s$ ) and radiative cooling of the atmosphere (upper box, with temperature  $T_a$ ). The second heating difference (right) results from the greater insolation in the tropical regions (left box, with temperature  $T_t$ ) compared to the extratropics (right box, with temperature  $T_p$ ). These heating differences serve as the hot and cold reservoirs (with temperatures  $T_h$  and  $T_c$ ) for heat engines from which work can be derived to fuel atmospheric motion. The grey areas show the hydrologic cycle and how it is associated with the energy balances. The stars (★) symbolize the phase transitions of water between its liquid and vapor state. Grey arrows indicate mass fluxes of water vapor, while black arrows indicate fluxes of liquid water.

A system is in a steady state when its internal energy as well as its entropy do not change in the mean, i.e.,  $dU/dt = 0$  and  $dS/dt = 0$ . It is in a state of thermodynamic equilibrium when the entropy of the system has reached its maximum value. In this state,  $\sigma = J_{net}^s = 0$ . A state of disequilibrium is maintained if the system maintains a net exchange of entropy,  $J_{net}^s > 0$ , which allows for entropy production within the system. Hence, such a disequilibrium state is maintained by entropy exchange. The most relevant form of entropy exchange in this study is the exchange of heat that takes place at different temperatures. In steady state, the heating,  $J_{in}$ , of a hot reservoir at temperature  $T_h$  balances the cooling,  $J_{out}$ , of a cold reservoir at temperature  $T_c$ , so that  $J_{out} = J_{in}$ . This heat exchange is, however, associated with a net entropy exchange,  $J_{net}^s$ , of

$$J_{net}^s = J_{in} \left( \frac{1}{T_c} - \frac{1}{T_h} \right) = \sigma \tag{3}$$

which balances the total entropy production,  $\sigma$ , by dissipative processes within the system (see Appendix A1 for a detailed derivation).

Work can be performed by the system from this net entropy exchange to generate, for instance, kinetic energy. When work is performed through time, it is described by the physical term power. The maximum rate at which work can be performed is known as the Carnot limit, which follows directly from the combination of the first and second law. If we consider a system in steady state with heat exchange fluxes  $J_{in}$  and  $J_{out}$  and that performs work at a rate  $G$ , then the first law states  $J_{in} = J_{out} + G$ . When  $J_{in}$  is added at a

temperature  $T_h$  and  $J_{out}$  is removed at a temperature  $T_c$ , then the second law requires that  $J_{net}^s = J_{in}/T_h - J_{out}/T_c \geq 0$ . The Carnot limit follows from the best case where the entropy exchange is zero,  $J_{net}^s = 0$ , yielding the expression:

$$G_{max} = J_{in} \cdot \frac{T_h - T_c}{T_h} \tag{4}$$

In the case of the atmosphere, this work is performed within the system, so that in steady state the work performed equals the free energy dissipated,  $G_{max} = D$ . If this dissipation occurs at the temperature of the cold reservoir, this results in entropy production within the system of

$$\sigma = \frac{D}{T_c} = \frac{G_{max}}{T_c} = J_{in} \frac{T_h - T_c}{T_h T_c} = J_{net}^s \tag{5}$$

Note that the Carnot limit does not require any particular knowledge about the mechanisms and steps by which this work is actually being derived from the heating gradient.

### 2.2 Energy and entropy balances of the Earth system

Heating gradients in the Earth system are generated from differential absorption of solar radiation. We consider the two dominant gradients that are formed from solar radiation, which are shown schematically in Fig. 2. Solar radiation is predominantly absorbed at the surface, which results in a vertical heating gradient between the surface and the atmosphere (Fig. 2a). Since more solar radiation is absorbed in the tropics than at the poles due to the orientation of surfaces

towards the incoming radiation, a large-scale horizontal gradient is formed between the tropics and the poles (Fig. 2b). These heating gradients are described by the respective energy balances, which we formulate in the following, based on the simple models of Kleidon (2004). These balances are considered in simplified form and in a steady state for simplicity, and set the basis for the following derivation of thermodynamic limits to hydrologic cycling.

### 2.2.1 Vertical heating gradient

Solar radiation is mostly absorbed at the surface, resulting in surface heating, while this heat is emitted mostly from the atmosphere in form of terrestrial radiation. This forms a vertical gradient in heating between the surface and the atmosphere (Fig. 2a) which forms the driving gradient for convective motion and local, vertical transport of moisture.

The difference in heating and cooling between the surface and the atmosphere is described by the energy balances of the surface and the atmosphere respectively. The following variables use the indices “s” and “a” to refer to the surface and atmosphere. The surface energy balance results from the absorption of solar radiation,  $J_{in,s}$ , the net radiative exchange between the surface and the atmosphere,  $J_{s,a}$ , and a convective flux of sensible and latent heat that is associated with atmospheric motion,  $J_c = H + \lambda E$ :

$$0 = J_{in,s} - J_{s,a} - J_c. \quad (6)$$

The energy balance for the atmosphere consists of the cooling by the emission of radiation into space,  $J_{out,a}$ , the heating due to radiative exchange with the surface,  $J_{s,a}$ , and the convective heat flux,  $J_c$ :

$$0 = -J_{out,a} + J_{s,a} + J_c. \quad (7)$$

Taken together, these energy balances yield the global radiative energy balance of the system:

$$0 = J_{in,s} - J_{out,a}. \quad (8)$$

The entropy balance of the system in steady state is composed of the import of heat  $J_{in,s}$  by absorption of solar radiation at the temperature  $T_s$  of the surface, the export of  $J_{out,a} = J_{in,s}$  by emission of terrestrial radiation from the atmosphere to space at a temperature  $T_a$ , and the total production of entropy  $\sigma_c$  within the system:

$$0 = \sigma_c + \frac{J_{in,s}}{T_s} - \frac{J_{out,a}}{T_a}. \quad (9)$$

Hence, the total entropy production is given by

$$\sigma_c = J_{in,s} \cdot \left( \frac{1}{T_a} - \frac{1}{T_s} \right). \quad (10)$$

This entropy production is composed of radiative exchange,  $\sigma_{rad}$ , the frictional dissipation associated by motion

and the sensible heat flux,  $\sigma_{sh}$ , and of irreversible processes associated with hydrologic cycling,  $\sigma_{lh}$ , so that

$$\sigma_c = \sigma_{rad} + \sigma_{sh} + \sigma_{lh} \quad (11)$$

as is detailed further below.

### 2.2.2 Horizontal heating gradient

Insolation is higher in the tropics than in the extratropics due to the uneven orientation of the surface of the earth to the incoming solar radiation. This results in a large-scale heating gradient between the tropics and the extratropics (Fig. 2b) which forms the heating gradient for large-scale motion and transport of moisture.

The large-scale temperature gradient in the horizontal is described by the energy balances of the tropical and extratropical (or polar) regions. The indices “t” and “p” in the following variables refer to the tropical and polar regions. The energy balance of the tropical regions is described by the absorption of solar radiation,  $J_{in,t}$ , the emission of terrestrial radiation into space,  $J_{out,t}$ , and the large-scale transport of sensible and latent heat from the tropics to the polar regions,  $J_{ls}$ :

$$0 = J_{in,t} - J_{out,t} - J_{ls}. \quad (12)$$

The energy balance of the extratropical regions is described equivalently by absorption of solar radiation,  $J_{in,p}$  (with  $J_{in,p} < J_{in,t}$ ), the emission of terrestrial radiation,  $J_{out,p}$ , and the heating by the heat transport from the tropical regions,  $J_{ls}$ :

$$0 = J_{in,p} - J_{out,p} + J_{ls}. \quad (13)$$

Taken together, both energy balances yield the global radiative energy balance:

$$0 = J_{in,t} + J_{in,p} - J_{out,t} - J_{out,p}. \quad (14)$$

The entropy balance of the system is composed of the import of heat when solar radiation,  $J_{in,t}$  and  $J_{in,p}$ , is absorbed at the respective temperatures, the export of heat when it is emitted as terrestrial radiation,  $J_{out,t}$  and  $J_{out,p}$ , and the entropy production within the system,  $\sigma_{ls}$ :

$$0 = \sigma_{ls} + \frac{J_{in,t} - J_{out,t}}{T_t} + \frac{J_{in,p} - J_{out,p}}{T_p} \quad (15)$$

In steady state, the entropy production within the system is given by

$$\sigma_{ls} = J_{ls} \cdot \left( \frac{1}{T_p} - \frac{1}{T_t} \right) \quad (16)$$

using Eqs. (12) and (13). As above (Eq. 10), the entropy production consists of changes in radiative processes,  $\sigma_{rad}$ , frictional dissipation of large-scale motion,  $\sigma_{sh}$ , and hydrologic processes,  $\sigma_{lh}$ :

$$\sigma_{ls} = \sigma_{rad} + \sigma_{sh} + \sigma_{lh} \quad (17)$$

with the contributions,  $\sigma_{sh}$  and  $\sigma_{lh}$ , described further below.

### 2.3 Atmospheric motion

Mechanical power is derived from the two heating gradients formulated above. This power results in the generation of kinetic energy and the formation of a momentum gradient that is associated with atmospheric motion. This power results from a buoyancy force that is generated from radiative heating of air, which lowers its density, but as we will see in the following, we do not need to consider a specific form of this buoyancy force. We describe the generation of motion in general, using the temperature  $T_h$  to refer to the heated region (surface or tropics) and  $T_c$  for the temperature of the cooled region (atmosphere or extratropics). For the two heating gradients discussed above,  $T_h = T_s$  and  $T_c = T_a$  for the vertical heating gradient, and  $T_h = T_t$  and  $T_c = T_p$  for the large-scale horizontal gradient.

The momentum associated with atmospheric motion reflects the balance of momentum generation by buoyancy,  $F_b$ , and dissipation by friction,  $F_f$ . This is expressed by the momentum balance, which expresses the time change of momentum as the balance of forces acting on the fluid. This balance is expressed here in steady state, so that the momentum change in time is zero:

$$0 = F_b - F_f = F_b - C_d \rho v^2. \quad (18)$$

Here,  $C_d$  is a drag coefficient that characterizes the strength of friction, which is typically turbulent. The balance equation for the kinetic energy relates the change in kinetic energy in time with the processes that generate and dissipate kinetic energy. Kinetic energy is generated by the buoyancy force,  $F_b$  at a rate,  $G$ , which expresses the work done by this force, i.e.  $F_b v$ . The dissipation of kinetic energy,  $D$ , results from the friction force. Hence, we obtain the balance equation for the kinetic energy associated with motion:

$$0 = G - D = G - C_d \rho v^3. \quad (19)$$

The power involved to generate motion that is reflected in  $G$  is drawn from heating differences. We consider the buoyancy forces that result from the differential heating as the result of a perfect heat engine within the atmosphere, for which the maximum rate at which kinetic energy is generated is given by the Carnot limit,  $G = G_{\max}$  (Eq. 4).

An important factor in these balances is friction, which determines the rate at which momentum, and hence kinetic energy, is dissipated. If we express the friction force as  $F_f = C_d \rho v^2 = (\rho v)(C_d v) = (\rho v)w$ , then we can view friction as a flux of momentum,  $\rho v$ , that is being exchanged between the fluid and the surface at rest with an effective exchange velocity of  $w = C_d v$ . Neither the momentum balance nor the kinetic energy balance provide the constraints to quantify the magnitude of  $w$ . We will use the maximum power limit to determine the limit associated with this exchange, and hence the value of  $w$ , in Sect. 3 below.

This exchange velocity,  $w$ , not only exchanges momentum, but also describes the exchange of sensible heat,  $H$ ,

between the heated reservoir with temperature  $T_h$  and the cooled reservoir with temperature  $T_c$ :

$$H = c_p \rho w (T_h - T_c), \quad (20)$$

where  $c_p$  is the specific heat capacity of air, and  $\rho$  is the air density.

The motion that is generated from the heating gradient and the associated heat flux follow the direction imposed by the second law. This heat flux is directed to transport heat from the warm reservoir to the cold reservoir. The extent by which this increases the entropy within the system is characterized by the rate at which entropy is being produced by atmospheric motion,  $\sigma_{\text{sh}}$ :

$$\sigma_{\text{sh}} = H \cdot \left( \frac{1}{T_c} - \frac{1}{T_h} \right) = \frac{G_{\max}}{T_c} \quad (21)$$

since power equals frictional dissipation,  $G = D$ , in steady state. Hence, atmospheric motion at the maximum power limit is approximately equivalent with producing entropy at the maximum possible rate (with an equivalence given if the kinetic energy is entirely dissipated at the cold temperature,  $T_c$ ).

### 2.4 Hydrologic cycling

Hydrologic cycling involves the phase transitions from liquid to gaseous during evaporation and the reverse during condensation. A state of maximum entropy and thermodynamic equilibrium is reached for an open water surface when the overlying air reaches saturation. The distance to thermodynamic equilibrium can thus be characterized by the relative humidity,  $r$ , of the air. To maintain evaporation and condensation in steady state, and thus hydrologic cycling, air needs to be continuously dehumidified by the atmosphere. This is mostly achieved by lifting of vapor by vertical motion.

We consider the formulation of hydrologic cycling in a general way, with cycling of moisture from a heated reservoir which evaporates at a temperature  $T_h$  to a cooled reservoir at which moisture condenses at a temperature  $T_c$  (as above, with the application to the two gradients given by  $T_h = T_s$  and  $T_c = T_a$  for the vertical heating gradient and  $T_h = T_t$  and  $T_c = T_p$  for the large-scale horizontal gradient). The water cycle is assumed to operate in a steady state within the system, in which evaporation  $E$  balances precipitation  $P$  so that the total amount of water vapor within the atmosphere does not change:

$$0 = E - P. \quad (22)$$

The evaporative flux balances the overall moisture transport between the two reservoirs, which transports the moistened air from the heated reservoir with a specific humidity of  $q_h$  and replaces it with the drier air from the cooled reservoir with a specific humidity of  $q_c$ . Hence, the latent heat flux,  $\lambda E$ , can be written as

$$\lambda E = \lambda \rho w (q_h - q_c). \quad (23)$$

If we further assume that air had sufficient time to reach saturation when water evaporates from the heated reservoir (or when it condenses at the cooled reservoir), we can approximate  $q_h \approx q_{\text{sat}}(T_h)$  and  $q_c \approx q_{\text{sat}}(T_c)$ . With the linearization of the saturation vapor pressure (see Appendix A2), we obtain for the latent heat flux:

$$\lambda E = c_p \rho w \frac{s}{\gamma} (T_h - T_c), \quad (24)$$

where  $\gamma = c_p p / (0.622 \lambda) \approx 65 \text{ Pa K}^{-1}$  is the psychrometric constant, and  $s$  is the slope of the saturation vapor pressure curve, as described by Eq. (A12).

The process of converting heat during evaporation at the heated reservoir and condensation at the cooled reservoir results in entropy production,  $\sigma_{\text{lh}}$ , associated with hydrologic cycling:

$$\sigma_{\text{lh}} = \lambda E \cdot \left( \frac{1}{T_c} - \frac{1}{T_h} \right). \quad (25)$$

This entropy production results from various dissipative processes related to hydrologic cycling. It includes all terms related to work and subsequent dissipation as well as diffusive processes that are directly related to hydrologic fluxes. This is because in the setup chosen here, the entropy produced by these processes is exchanged by  $\lambda E$  at the temperatures  $T_h$  and  $T_c$  at the system boundary, respectively. It includes the frictional dissipation associated with the additional generation of atmospheric motion due to the greater buoyancy of moistened air compared to dry air and due to the release of latent heat at an atmospheric temperature that is typically above the radiative temperature of the atmosphere (both of these effects are not explicitly considered here further for sake of simplicity). This entropy production also includes the frictional dissipation of falling raindrops (Pauluis et al., 2000) resulting from the work done in lifting water from the surface to the level at which it precipitates and it includes diffusive losses of water vapor within the atmosphere (Goody, 2000). Hence, the entropy production  $\sigma_{\text{lh}}$  should consist of at least three terms, relating to the frictional dissipation due to moist convection, frictional dissipation of falling raindrops, and diffusive losses.

In the case of vertical convection, the entropy production,  $\sigma_{\text{lh}}$ , would in the ideal case be associated entirely with the dissipation resulting from the work that converts energy into additional kinetic energy (thus contribute a term  $G_{\text{lh}}$  to the  $G$  in Eq. (19), with the power  $G_{\text{lh}}$  then being related to generate moist convection), and into the work  $G_{\text{lift}}$  that is associated with the lifting of water to the height at which it precipitates. The entropy would then be generated when the generated additional kinetic energy and potential energy is dissipated. We would then have

$$G_{\text{lh}} + G_{\text{lift}} = \lambda E \frac{T_h - T_c}{T_h} = D_{\text{lh}} = \sigma_{\text{lh}} T_c. \quad (26)$$

The power that is associated with lifting water at a rate  $E$  to the height  $\Delta z$  at which it precipitates is given by

$$G_{\text{lift}} = E g \Delta z, \quad (27)$$

where  $g$  is the gravitational acceleration, and the height of lifting is approximated by

$$\Delta z = (T_h - T_c) / \Gamma, \quad (28)$$

where  $\Gamma$  is the dry adiabatic lapse rate. We use the dry adiabatic lapse rate because we assume that condensation takes place at the temperature  $T_c$  so that the lifting would occur in the absence of condensational heating.

In the case of large-scale horizontal transport, moisture is transported horizontally and does not involve lifting. Hence, in the ideal case the power associated with the latent heat flux would be converted into kinetic energy and  $G_{\text{lift}} = 0$ .

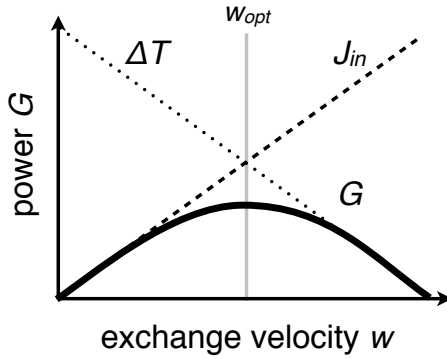
### 3 Maximum strength of hydrologic cycling

In this section we derive expressions and values for the maximum strength of hydrologic cycling from the set of energy- and entropy balances of the previous section. While these formulations are highly simplified, the description of the two dominant radiative gradients is “complete” in the sense that energy and mass is being conserved and the second law of thermodynamics is being fulfilled.

#### 3.1 Maximum power limit

The energy balances provide the basis to determine the temperatures and the respective temperature differences. For the net emission of terrestrial radiation, we use a linearized approximation for terrestrial radiation as detailed in the Appendix A2. For given fluxes of insolation ( $J_{\text{in},s}$ ,  $J_{\text{in},t}$ , and  $J_{\text{in},p}$ ) and a given radiative “conductivity”  $k_r$ , this set of energy balances is in principle determined, except for the strength of convective motion. Motion is unconstrained because the value of the strength of turbulent friction,  $C_d$ , or, equivalently, the exchange velocity  $w$  is unknown. In this context it is important to note that the Carnot limit (Eq. 4) does not determine the strength of motion because the convective heat flux that drives the “engine”,  $J_{\text{in}}$ , while constrained, is not determined by the energy balances. The case of no motion ( $J_{\text{in}} = 0$  in Eq. 4) is a solution to the energy balance in steady state, but the energy balances are also solved by values greater than zero ( $J_{\text{in}} > 0$ ). The second law further constrains the value of  $J_{\text{in}}$  to values that maintain a temperature gradient that directs the heat fluxes from the heated region to the cooled region.

The additional constraint to motion comes from the contrasting effects of the heat flux  $J_{\text{in}}$  on the power that can be derived from this heat flux which results in a maximum power state. These contrasting effects on power are illustrated qualitatively in Fig. 3. A small heat flux is associated



**Fig. 3.** Illustration of the maximum power limit to convective exchange resulting from the depletion of the driving temperature difference by the resulting exchange of heat. The Carnot limit to the power,  $G$ , of a heat engine (solid line) is given by the product of the heat flux,  $J_{in}$  (dashed line), and the driving temperature difference,  $\Delta T$  (dotted line). In the atmosphere, this work is used to generate motion and is reflected in a certain velocity  $w$  by which heat is exchanged by convection. The resulting heat flux increases with  $w$ , but depletes the driving temperature difference. These contrasting effects result in a state of maximum power, which is associated with an optimum exchange velocity  $w_{opt}$ .

with a small rate of heat exchange,  $w$ , and at this end, the power,  $G$  (solid line), first increases with higher rates of convective exchange,  $w$ . However, with greater values of  $w$ , heat is transported from the heated to the cooled reservoir at a faster rate, so that the difference in temperatures must decrease. This reduction in the temperature difference reduces the efficiency of generating power out of the convective heat flux, so that a state of maximum power is achieved at an intermediate value of the heat flux associated with an optimum exchange velocity,  $w_{opt}$ , and temperature difference,  $T_h - T_c$ . We use this maximum power state to derive the exchange velocity,  $w_{opt}$ , which then represents the limit by which motion can be generated in steady state. Since a maximum in power corresponds to a maximum in turbulent dissipation, the turbulent exchange through the surface boundary layer should also be maximized in this state. Hence, the maximum strength of hydrologic cycling is derived from the assumption that the vapor transport by atmospheric motion limits hydrologic cycling.

### 3.2 Maximum strength of hydrologic cycling by vertical convection

To derive an analytic expression for the maximum strength of hydrologic cycling that is driven by vertical convection, we start at the Carnot limit for convective motion. Here, the sensible heat flux acts as the heat flux from which motion is generated ( $J_{in} = H_c$  in Eq. 4), and the temperature difference  $T_h - T_c$  is described by the temperature difference between the surface and the atmosphere,  $T_s - T_a$ :

$$G_c = H_c \cdot \frac{T_s - T_a}{T_s}. \quad (29)$$

We use the subscript  $c$  in the following to refer to the variables associated with vertical convection. With the linearized expression for the radiative exchange flux  $J_{s,a}$ , we can express the temperature difference by the energy balances, Eqs. (6) and (7), and the expressions for the turbulent heat fluxes, Eqs. (20) and (23), as

$$T_s - T_a = \frac{J_{in,s}}{k_r + c_p \rho w_c (1 + s/\gamma)}. \quad (30)$$

With this expression for the temperature difference and the expression for the sensible heat flux, Eq. (20), the power can be expressed as

$$G_c = \frac{c_p \rho w_c}{T_s (k_r + c_p \rho w_c (1 + s/\gamma))^2} J_{in,s}^2. \quad (31)$$

This expression has a maximum for an optimum value for the vertical exchange velocity, which results from the trade-off between heat flux and driving gradient, as shown in Fig. 4. For this figure, values of  $J_{in,s} = 165 \text{ W m}^{-2}$ ,  $k_r = 5.4 \text{ W m}^{-2} \text{ K}^{-1}$ ,  $c_p \rho = 1200 \text{ J K}^{-1} \text{ m}^{-3}$  and  $s = 124 \text{ Pa K}^{-1}$  were used (with a mean temperature of  $T = 288 \text{ K}$ ). It shows how the driving temperature gradient is the largest with no convective exchange. In such a state, the difference in heating would be entirely dissipated by radiative exchange, as shown by the ratio  $J_{s,a} \approx J_{in,s}$  in Fig. 4b. When more of the solar radiative heating is transported by convective heat transport to the cold reservoir (i.e. a greater ratio of  $(H_c + \lambda E_c)/J_{in,s}$ ), then this is associated with a lower value of the temperature difference. Hence, this trade-off between the convective heat flux and the temperature difference shape the maximum power state shown in Fig. 4b.

The analytic expressions for the optimum characteristics that lead to the maximum in power are derived from  $\partial G_c / \partial w_c = 0$ . We neglect the dependence of  $T_s$  on the heat fluxes in the denominator (because the variations in  $1/T_s$  with a mean value of  $T_s \approx 300 \text{ K}$  are small), and obtain

$$w_{c,opt} = \frac{\gamma}{s + \gamma} \frac{k_r}{c_p \rho}. \quad (32)$$

The associated optimum turbulent heat fluxes,  $H_{c,opt}$  and  $\lambda E_{c,max}$ , are given by

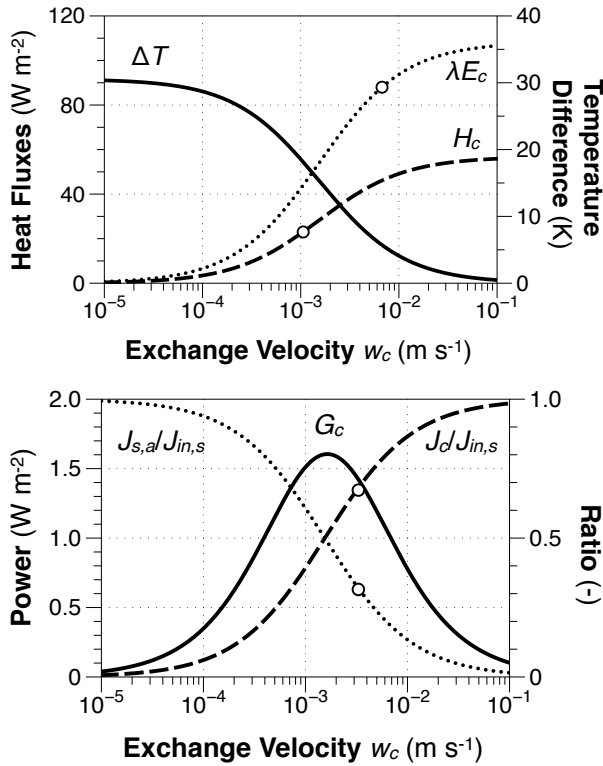
$$H_{c,opt} = \frac{\gamma}{s + \gamma} \frac{J_{in,s}}{2} \quad (33)$$

and

$$\lambda E_{c,max} = \frac{s}{s + \gamma} \frac{J_{in,s}}{2}. \quad (34)$$

The optimum latent heat flux,  $\lambda E_{c,max}$ , corresponds to the evaporative mass flux of water at the surface, and the flux





**Fig. 4.** Trade-off involved in setting the maximum strength in hydrologic cycling due to vertical solar heating differences. **(a)** Sensible (dashed line,  $H_c$ ) and latent (dotted line,  $\lambda E_c$ ) heat flux as well as the temperature difference (solid line,  $\Delta T = T_s - T_a$ ) as a function of exchange velocity  $w_c$ . **(b)** Power (solid line,  $G_c$ ) as well as the ratio of radiative cooling (dotted line,  $J_{s,a}$ ) and cooling by convective heat fluxes (dashed line,  $J_c = H_c + \lambda E_c$ ) in relation to the rate of heating by absorption of solar radiation,  $J_{in,s}$ . The circles (o) on the lines  $\lambda E_c$  and  $H_c$  in **(a)** and on  $J_{s,a}/J_{in,s}$  and  $J_c/J_{in,s}$  in **(b)** represent the observed estimates of Stephens et al. (2012).

of precipitation of vapor within the atmosphere. Hence, we derive the maximum strength of evaporation,  $E_{c,max}$ , that is permitted by convective motion from Eq. (34) by

$$E_{c,max} = \frac{s}{s + \gamma} \frac{J_{in,s}}{2\lambda} = P_{c,max}. \quad (35)$$

Noting that the net emission of radiation from the surface,  $J_{s,a}$ , is also  $J_{in,s}/2$  at this maximum power state, we can express  $E_{c,max}$  in terms of the net surface radiation,  $J_{net} = J_{in,s} - J_{s,a} = J_{in,s}/2$ :

$$E_{c,max} = \frac{s}{s + \gamma} \frac{J_{net}}{\lambda}. \quad (36)$$

This expression is identical with the equilibrium evaporation rate (Slayter and McIlroy, 1961; Priestley and Taylor, 1972), a concept that is well established in estimating evaporation rates at the surface.

Furthermore, we can characterize the state of the atmospheric vapor content in terms of its distance to

thermodynamic equilibrium. The relative humidity of the surface air,  $r_c$ , that drives surface evaporation is in the context of the formulations here given by

$$r_c = \frac{e_{sat}(T_a)}{e_{sat}(T_s)}. \quad (37)$$

Using the linearizations for  $e_{sat}$  and the approximation  $1/(1+x) \approx 1-x$ , this expression can be reformulated as

$$r_c \approx 1 - \frac{s}{e_{sat}(T_a)} \frac{J_{in,s}}{2k_r}. \quad (38)$$

Since the emission of terrestrial radiation is constrained by the global energy balance,  $T_a$  as well as  $q_{sat}(T_a)$  should be relatively fixed properties within the system, so that the dominant effect on  $r$  is the decrease in  $r$  with an increase in  $J_{in,s}$ .

In summary, we note that the maximum strength of hydrologic cycling, in terms of its fluxes (cf. Eq. 35) as well as its state of disequilibrium (cf. Eq. 38), is directly related to the absorption of solar radiation,  $J_{in,s}$ , which represents the heat input that generates the driving temperature gradient between the earth's surface and the atmosphere.

### 3.3 Maximum strength of hydrologic cycling by the large-scale circulation

The maximum strength of hydrologic cycling by the large-scale circulation is derived equivalently. We start with the expression of the Carnot limit, applied to the large-scale horizontal difference in insolation:

$$G_{ls} = H_{ls} \cdot \frac{T_t - T_p}{T_t}. \quad (39)$$

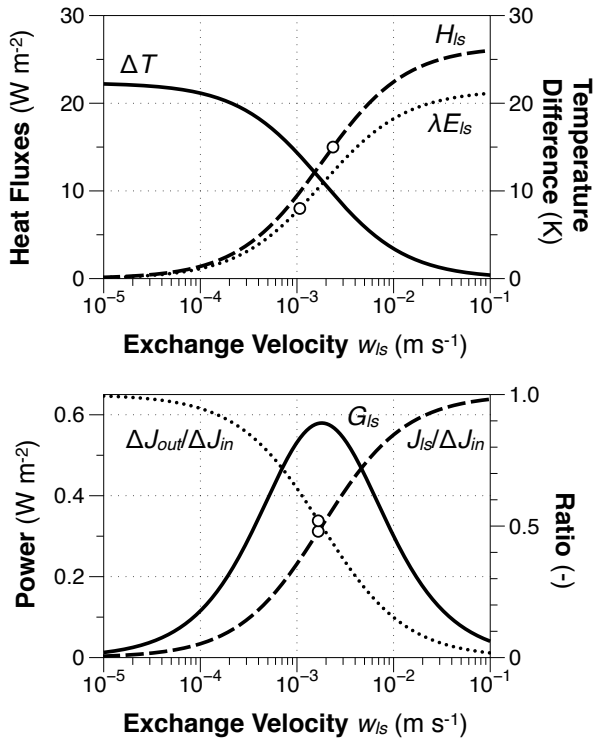
We solve for the temperature difference,  $T_t - T_p$ , using Eqs. (12) and (13), use the expressions for the turbulent heat fluxes, Eqs. (20) and (23), and obtain:

$$T_t - T_p = \frac{J_{in,t} - J_{in,p}}{k_r + 2c_p \rho w_{ls} (1 + s/\gamma)}. \quad (40)$$

The power  $G_{ls}$  associated with large-scale generation of motion is then given by

$$G_{ls} = \frac{c_p \rho w_{ls}}{T_t (k_r + 2c_p \rho w_{ls} (1 + s/\gamma))^2} (J_{in,t} - J_{in,p})^2. \quad (41)$$

As in the case of vertical convection, this expression has a maximum for an optimum value for the horizontal exchange velocity,  $w_{ls}$ . This maximum is illustrated in Fig. 5. We used values of  $J_{in,t} = 288 \text{ W m}^{-2}$ ,  $J_{in,p} = 192 \text{ W m}^{-2}$ , which roughly correspond to the mean solar absorption in the tropics (latitudes  $< 30^\circ$ ), which is about 20% above the global mean absorption of solar radiation of  $240 \text{ W m}^{-2}$ , and the extratropics, which is about 20% below the global



**Fig. 5.** Same as Fig. 4, but for the exchange fluxes caused by the large-scale horizontal difference in solar radiative heating. **(a)** Sensible (dashed line,  $H_{ls}$ ) and latent (dotted line,  $\lambda E_{ls}$ ) heat flux as well as the temperature difference (solid line,  $\Delta T = T_t - T_p$ ) as a function of exchange velocity  $w_{ls}$ . **(b)** Power (solid line,  $G_{ls}$ ) as well as the ratio of the difference in radiative cooling (dotted line,  $\Delta J_{out} = (J_{out,t} - J_{out,p})/2$ ) and cooling by convective heat fluxes (dashed line,  $J_{ls} = H_{ls} + \lambda E_{ls}$ ) in relation to the difference in solar radiation,  $\Delta J_{in} = (J_{in,t} - J_{in,p})/2$ . The circles (o) on the lines  $\lambda E_{ls}$  and  $H_{ls}$  in **(a)** and on  $\Delta J_{out}/\Delta J_{in}$  and  $J_{ls}/\Delta J_{in}$  in **(b)** represent the observed estimates of Masuda (1988) and Oki et al. (1995).

mean. For the other parameters, we used values representative of the mean mid-troposphere, with a temperature of  $T = 266$  K and pressure of 500 hPa, yielding  $k_r = 4.3$  W m<sup>-2</sup> K<sup>-1</sup>,  $c_p \rho = 657$  J K<sup>-1</sup> m<sup>-3</sup>,  $\gamma = 32$  Pa K<sup>-1</sup> and  $s = 26$  Pa K<sup>-1</sup>. The shape of the relationships are practically identical with those shown in Fig. 4, yet the interpretation is somewhat different. Here, it is not the absorbed solar radiation from the surface,  $J_{in,s}$  that is transported by convective motion to the atmospheric cold reservoir, but rather the difference in absorbed solar radiation,  $(J_{in,t} - J_{in,p})/2$ , which is transported by horizontal motion to the cold, extratropical reservoir. Also, the relative proportion of transported sensible and latent heat are different, which is due to the lower value of  $s$  due to the lower, assumed temperature at which the value of  $s$  was calculated.

The analytic expressions associated with the maximum power state,  $G_{ls,max}$ , are derived as above from the condition  $\partial G_{ls}/\partial w_{ls} = 0$ . This yields an expression for the optimum

exchange velocity,  $w_{ls,opt}$ :

$$w_{ls,opt} = \frac{\gamma}{s + \gamma} \frac{k_r}{2c_p \rho}. \quad (42)$$

The associated optimum large-scale transport of sensible and latent heat,  $H_{ls,opt}$  and  $\lambda E_{ls,max}$ , are given by

$$H_{ls,opt} = \frac{\gamma}{s + \gamma} \frac{(J_{in,t} - J_{in,p})}{4} \quad (43)$$

and

$$\lambda E_{ls,max} = \frac{s}{s + \gamma} \frac{(J_{in,t} - J_{in,p})}{4}. \quad (44)$$

Again, we convert the latent heat flux to a mass flux associated with evaporation and precipitation due to large-scale motion:

$$E_{ls,max} = \frac{s}{s + \gamma} \frac{(J_{in,t} - J_{in,p})}{4\lambda} = P_{ls,max}. \quad (45)$$

The relative humidity,  $r_{ls}$ , of the air from the extratropics to the tropics that would drive evaporation in the tropics is given by

$$r_{ls} = \frac{e_{sat}(T_p)}{e_{sat}(T_t)} \quad (46)$$

or, using the equivalent approximations as above,

$$r_{ls} \approx 1 - \frac{s}{e_{sat}(T_p)} \frac{J_{in,t} - J_{in,p}}{4k_r}. \quad (47)$$

As in the case of vertical convection, we note that the maximum strength of large-scale hydrologic cycling is directly related to the heat input that generates the driving gradient, which in this case is given by the magnitude of differential solar radiative heating,  $J_{in,t} - J_{in,p}$ . Since  $J_{in,s}$  is generally greater than  $(J_{in,t} - J_{in,p})/2$ , we would expect convective cycling to play a more dominant role in cycling water than large-scale transport.

### 3.4 Coupling effects

So far, we treated the two driving gradients separately. However, the horizontal gradient in radiative heating provides additional means to generate motion, and this motion was not accounted for in the model of vertical exchange. In the following we explore the effect of horizontal motion on the partitioning within the surface energy balance in a simplified way.

To do so, we extend the parameterizations of the turbulent heat fluxes for convection by adding a contribution to vertical exchange (Eqs. 20 and 23):

$$H_c = c_p \rho (w_c + w_{ls})(T_s - T_a) \quad (48)$$

and

$$\lambda E_c = c_p \rho (w_c + w_{ls}) \frac{s}{\gamma} (T_s - T_a). \quad (49)$$

For simplicity, we express the additional contribution by large-scale motion as being proportional to the vertical exchange, so that

$$w_{ls} = f_{ls} w_c. \quad (50)$$

With this modification, the surface energy balance yields a somewhat different expression for the temperature difference,  $T_s - T_a$  (cf. Eq. 30):

$$T_s - T_a = \frac{J_{in,s}}{k_r + c_p \rho w_c (1 + f_{ls})(1 + s/\gamma)}. \quad (51)$$

Using the expression for  $w_{c,opt}$  (Eq. 32), we then obtain for the partitioning of heat fluxes

$$J_{s,a} = \frac{2}{2 + f_{ls}} \frac{J_{in,s}}{2} \quad (52)$$

$$H_{c,opt} = \frac{2 + 2f_{ls}}{2 + f_{ls}} \frac{\gamma}{\gamma + s} \frac{J_{in,s}}{2} \quad (53)$$

and

$$\lambda E_{c,opt} = \frac{2 + 2f_{ls}}{2 + f_{ls}} \frac{s}{\gamma + s} \frac{J_{in,s}}{2}. \quad (54)$$

These expressions are essentially the same as before (cf. Eqs. 33 and 34), except for a shift from radiative to convective cooling since the net radiative cooling,  $J_{s,a}$ , is reduced by a factor of  $2/(2 + f_h)$ , while the turbulent fluxes are enhanced by a factor of  $(2 + 2f_h)/(2 + f_h)$ .

Given that the power for large-scale motion is maximized at an exchange velocity that is essentially of the same magnitude as the vertical exchange velocity (cf. Figs. 4 and 5), the contribution of large-scale motion is of similar magnitude as the locally generated motion, so that  $f_{ls} \approx 1$ . With this value for  $f_h$ ,  $2/(2 + f_{ls}) = 2/3 = 0.67$  and  $(2 + 2f_{ls})/(2 + f_{ls}) = 4/3 = 1.33$ . This enhancement factor of the turbulent heat fluxes is very close in value to the empirically derived Priestley–Taylor coefficient of 1.26 (Priestley and Taylor, 1972), which is typically used in empirical estimates of potential evaporation. This coefficient reflects the effect of large-scale motion which enhances turbulent exchange at the surface at the expense of net radiative cooling. This effect shifts the partitioning away from a  $J_c = J_{in,s}/2$  that would result if turbulent exchange was generated by local surface heating only. This interpretation is consistent with the general interpretation of potential evaporation (Penman, 1948) in which potential evaporation is seen as the combined contribution of local radiative heating and a dryness term associated with atmospheric motion.

**Table 3.** Global estimates of the maximum strength of the hydrologic cycle due to the vertical solar heating gradient (“convection”,  $J_{in,s}$ ) and due to the large-scale horizontal solar heating gradient (“large-scale”,  $(J_{in,t} - J_{in,p})/2$ ). The estimates represent the values associated with the maximum power states shown in Figs. 4 and 5. The optimum latent heat flux associated with a maximum power state is then converted to a corresponding water flux per unit area and to a global estimate.

Property	Convective	Large-scale
driving gradient (observed)	165 W m <sup>-2</sup>	48 W m <sup>-2</sup>
max. power (dry convection)	1.5 W m <sup>-2</sup>	0.6 W m <sup>-2</sup>
max. power (moist convection)	2.5 W m <sup>-2</sup>	0.5 W m <sup>-2</sup>
max. power (lifting)	0.4 W m <sup>-2</sup>	–
exchange velocity	1.5 mm s <sup>-1</sup>	1.8 mm s <sup>-1</sup>
terrestrial radiation	83 W m <sup>-2</sup>	24 W m <sup>-2</sup>
sensible heat	28 W m <sup>-2</sup>	13 W m <sup>-2</sup>
latent heat	54 W m <sup>-2</sup>	11 W m <sup>-2</sup>
max. evaporation	2.0 mm d <sup>-1</sup>	0.4 mm d <sup>-1</sup>

### 3.5 Global estimates

The expressions for optimum evaporative fluxes associated with maximum power are now used to derive the maximum strength of the hydrologic cycle that is associated with the two dominant gradients in solar radiative heating and associated properties. The numerical values of these estimates should be seen as estimates of the order of magnitude because, obviously, quite a number of simplifications were made and many specifics are not included in the estimates. Some analysis on the importance of different factors that shape these estimates, in particular to those that shape the values of the involved power, are further explored in the context of sensitivities in the next section.

We used the above values of the mean absorption of solar radiation at the surface,  $J_{in,s} = 165 \text{ W m}^{-2}$  (Stephens et al., 2012), and the difference in absorption of solar radiation between the tropics and the extratropics,  $J_{in,t} = 288 \text{ W m}^{-2}$  and  $J_{in,p} = 192 \text{ W m}^{-2}$ . The separation between tropics and extratropics is done at a latitude of 30°, so that the associated areas are of the same size. Due to the difference in inclination of the surfaces, the tropics absorb approximately 20 % more solar radiation than the mean of 240 W m<sup>-2</sup>, while the extratropics receive 20 % less than the mean. These values were already used in the examples shown in Figs. 4 and 5. The values associated with the maximum power states are shown in Table 3 and these values were then combined with a value of  $f_{ls} \approx 1$  to compare these estimates with the corresponding estimates based on observations in Table 4.

**Table 4.** Comparison of the maximum strength estimates of hydrologic cycling derived here from the assumption of maximum power associated with atmospheric transport (and with  $f_{1s} = 1$ ) with observation-based estimates.

Property	Estimated	Observed	Sources
net terrestrial radiation, $J_{s,a}$	55 W m <sup>-2</sup>	52 W m <sup>-2</sup>	Stephens et al. (2012)
sensible heat flux, $H$	37 W m <sup>-2</sup>	24 W m <sup>-2</sup>	Stephens et al. (2012)
latent heat flux, $\lambda E$	72 W m <sup>-2</sup>	88 W m <sup>-2</sup>	Stephens et al. (2012)
evaporation/precipitation, $E$	2.7 mm d <sup>-1</sup>	2.7 mm d <sup>-1</sup>	Oki and Kanae (2006)
poleward heat transport, $J_{1s}$	24 W m <sup>-2</sup>	23 W m <sup>-2</sup>	Masuda (1988)
large-scale moisture transport, $\lambda E_{1s}$	11 W m <sup>-2</sup>	8 W m <sup>-2</sup>	Oki et al. (1995)
large-scale generation of kinetic energy, $G_{1s}$	1.1 W m <sup>-2</sup>	2 W m <sup>-2</sup>	Peixoto and Oort (1992)
frictional dissipation of raindrops, $G_{\text{lif}}$	0.4 W m <sup>-2</sup>	< 2–4 W m <sup>-2</sup>	Pauluis et al. (2000)

We first compare the estimated partitioning of fluxes in the surface energy balance to the estimates of Stephens et al. (2012). Our estimates are of similar magnitude as observed. The magnitude of the net flux of terrestrial radiation of 55 W m<sup>-2</sup> is very close to the observed value of 52 W m<sup>-2</sup>. This very close agreement is due to the correction made to account for the effect of large-scale motion, using Eq. (52), which reduces the estimate shown in Table 3 by a factor of 2/3. The partitioning between sensible and latent heat is biased towards the sensible heat flux, overestimating it by about 13 W m<sup>-2</sup>, while the latent heat flux is underestimated by 16 W m<sup>-2</sup>. When we compare the corresponding mean strength of the hydrologic cycle in terms of mean global evaporation to estimates of Oki and Kanae (2006), our estimate is nevertheless about the same. The optimum value of the exchange velocity of about  $w_{\text{opt}} \approx 1.5 \text{ mm s}^{-1}$  is in the order of magnitude of observed updraft velocities (Peixoto and Oort, 1992), although the observed latent heat flux would imply a stronger exchange of about 6.7 mm s<sup>-1</sup> (see circle marked in Fig. 4a). In our global estimate, a higher surface exchange is partially accounted for by accounting for the contribution of large-scale transport, as described in Sect. 3.4.

The estimated heat fluxes associated with the large-scale horizontal difference in solar irradiation are compared to values of poleward heat and moisture transport. Masuda (1988) estimates from satellite radiative data that the peak hemispheric heat transport is about 5.9 PW (1 PW = 10<sup>15</sup> W). We use the total heat transport rather than just the part associated with atmospheric transport, because we do not resolve the oceanic contribution. However, Trenberth and Caron (2001) estimate that the atmosphere transports the vast majority of this heat. When the value of 5.9 PW is divided by surface area, this value corresponds to a mean heat transport per unit surface area of 23 W m<sup>-2</sup>. Our estimate of 24 W m<sup>-2</sup> is very close to this number. The contribution to this heat transport by water vapor was estimated by Oki et al. (1995). They estimated the peak in large-scale hemispheric moisture transport to be about  $25 \times 10^{12} \text{ m}^3 \text{ a}^{-1}$ . This flux corresponds to a transport of latent heat of about 2 PW per hemisphere, or about

8 W m<sup>-2</sup>. Our estimate of 11 W m<sup>-2</sup> is in the same range, but somewhat higher.

The estimated values for the power associated with large-scale motion is compared to the generation rate of kinetic energy, which is typically in the order of 2 W m<sup>-2</sup> (Lorenz, 1955; Peixoto and Oort, 1992). Our estimate of 1.1 W m<sup>-2</sup> is about half of the observed value. This low value can be explained by the comparatively large value of  $k_r$  used here, which was derived from the linearization of blackbody radiation. A comparable model which used an empirical parameterization of net terrestrial radiation (Kleidon, 2010) with a smaller slope and yielded a generation rate comparable to the observed magnitude. Since  $k_r$  does not enter the estimates for the optimum heat flux directly, the optimum heat flux, the partitioning into sensible and latent heat transport, and the strength of large-scale hydrologic cycling should be relatively unaffected by this bias. The power involved in lifting moisture which is subsequently dissipated by falling raindrops has been estimated by Pauluis et al. (2000) to be up to 2–4 W m<sup>-2</sup> in the tropics. Our global estimate of 0.4 W m<sup>-2</sup> is quite a bit lower, but as it represents a global average, it does not correspond to an upper bound on this number so that these numbers are not directly comparable.

To compare this effect to spatial differences in radiative forcing, another estimate was performed in which the flux partitioning at the surface were computed separately for tropical and extratropical regions and then averaged (Table 5). This estimate shows that the spatial differences in radiative forcing mostly average out and show little effect on the partitioning of absorbed solar radiation into radiative vs. turbulent cooling at the global scale. This insensitivity is reasonable, because the fluxes are mostly proportional to the absorption of solar radiation,  $J_{\text{in},s}$ . This aspect would, however, need to be further explored in more spatial and temporal detail.

In summary, our extremely simple estimates that are based on the assumption of maximum power associated with atmospheric transport estimate the observed magnitudes of heat and moisture transport within the climate system very well. This suggests that the global hydrologic cycle can indeed be

**Table 5.** Evaluation of the effect of spatial differences in absorption of solar radiation in the tropics and extratropics on the global average. For the estimate, it was assumed that the surface temperatures in the tropics and the extratropics were 303 K and 273 K, respectively.

Heat flux	$f_{\text{ls}} = 0$			$f_{\text{ls}} = 1$		
	Tropics	Extratropics	Global	Tropics	Extratropics	Global
absorption of solar radiation, $J_{\text{in,s}}$	198	132	165	198	132	165
net emission of terrestrial radiation, $J_{\text{s,a}}$	99	66	83	66	44	55
sensible heat flux, $H_{\text{c}}$	20	39	30	27	53	40
latent heat flux, $\lambda E_{\text{c}}$	79	27	53	105	35	70

understood as a cycle that operates close to the limit of maximum strength.

#### 4 Sensitivities

The maximum strength of hydrologic cycling is inferred here from the assumption of maximum power regarding moisture transport. The comparison with observations shows that these estimates are quite reasonable in their magnitude, but are sometimes lower than what is observed. The tradeoff that results in the maximum power limit is a fundamental consequence of the second law in combination with energy balance constraints in steady state, so that the existence of this limit should not be a concern. The particular optimum state that is associated with this limit depends, however, on the particular formulations that describe the relationship between heat flux, the energy balance constraints, on the assumption of steady state conditions, and on the use of global mean averages rather than spatially and temporally explicit values. Hence, it is important to understand the factors that affect these maximum power states and what these imply for the estimated maximum strength of hydrologic cycling.

The most important factor in these estimates is the magnitude of the solar radiative forcing. The magnitude of the forcing enters directly into the estimates for the maximum power (Eqs. 31 and 41) as well as the expressions for the optimum evaporation rate (Eqs. 35 and 45). The magnitude of these forcings vary spatially and temporally on earth, so we first look at the sensitivity of these expressions to the two solar radiative heating differences,  $J_{\text{in,s}}$  and  $(J_{\text{in,t}} - J_{\text{in,p}})$  (Fig. 6).

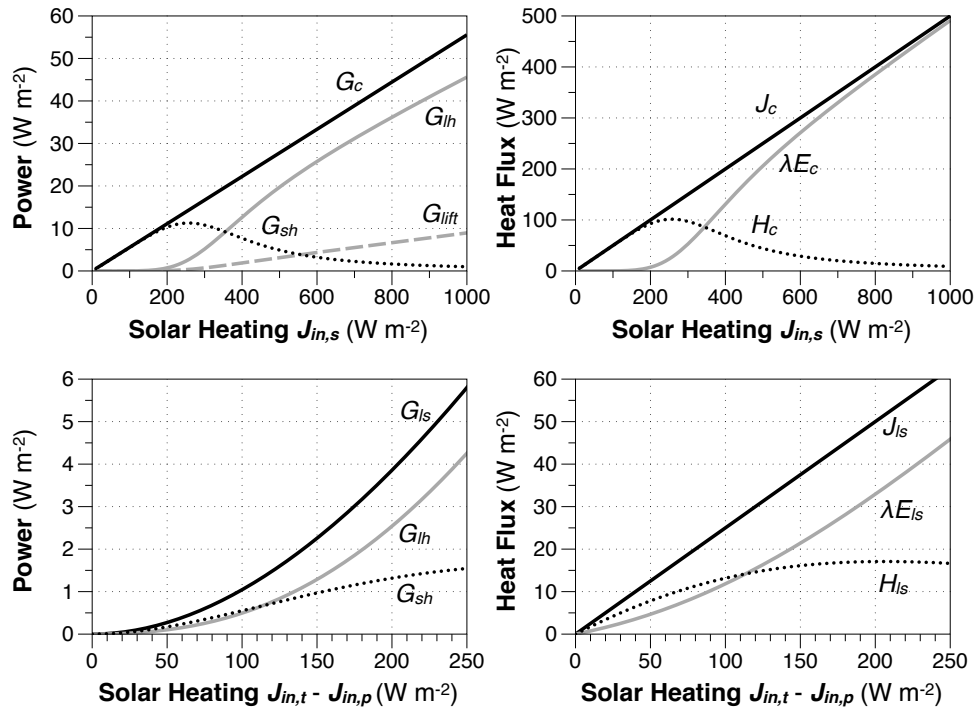
The sensitivity to the vertical solar radiative heating difference,  $J_{\text{in,s}}$ , is shown in Fig. 6a and b. The sensitivity was calculated in a slightly different way as in the maximum power examples of the previous section, because for the wide range of variation in  $J_{\text{in,s}}$  it can no longer be assumed that  $k_{\text{r}}$  and  $s$  are unaffected. These values were calculated for the radiative temperature that is associated with the particular radiative forcing. We find that at the limit of low and high values of  $J_{\text{in,s}}$ , the power as well as the heat fluxes vary roughly linearly with the forcing. At these two extremes, variations in the forcing in either space or time would average out when using mean values, because of this linearity.

It is rather in the middle range of the sensitivity at values of  $J_{\text{in,s}} \approx 200 - 400 \text{ W m}^{-2}$  where the relationships are non-linear, and variations do not average out. In this range, an averaging of variations would likely result in an underestimation of the latent heat flux,  $J_{\text{lh}}$ , and the associated values of  $G_{\text{lh}}$  and  $G_{\text{lif}}$ .

The sensitivity to the magnitude of the large-scale horizontal forcing is shown in Fig. 6c and d. In these sensitivities, the mean solar radiation was fixed to a value of  $240 \text{ W m}^{-2}$ , the value of  $k_{\text{r}}$  was derived from the mean forcing, and the value of  $s$  from the temperature  $T_{\text{t}}$ . The sensitivities for this forcing are much more linear, although we note that  $G_{\text{ls}}$  shows a slight quadratic dependence on  $J_{\text{in,t}} - J_{\text{in,p}}$ , and  $H_{\text{ls}}$  saturates at high values of  $J_{\text{in,t}} - J_{\text{in,p}}$ . When variations are averaged, we would expect that the power is underestimated, and we may also expect a slight underestimation of the latent heat flux.

In the estimates, we assumed that the atmosphere is fully absorbing. In the natural atmosphere, this is often not the case, and the absorptivity of the atmosphere depends on many factors, in particular on cloud cover and the concentration of greenhouse gases (such as water vapor and carbon dioxide). The effect of the atmospheric greenhouse can be explored by a sensitivity to  $k_{\text{r}}$ , with a smaller value of  $k_{\text{r}}$  corresponding to a stronger greenhouse forcing since a greater temperature difference would be needed to accomplish a certain net radiative exchange. This sensitivity to  $k_{\text{r}}$  is shown in Fig. 7.

In the sensitivity of the vertical heat fluxes (Fig. 7a and b), we note that the value of the optimum convective heat flux  $J_{\text{c}}$  is not affected by the value of  $k_{\text{r}}$ , but that the power as well as the partitioning between sensible and latent heat are strongly affected. A low value of  $k_{\text{r}}$  results in higher surface temperatures, a greater temperature difference between the surface and the atmosphere, and hence in a greater overall value of the maximum power,  $G_{\text{c}}$ , as well as a shift in the partitioning towards a greater value of the latent heat flux. The same line of reasoning applies to the sensitivity of the horizontal heat fluxes that is shown in Fig. 7c and d. What this implies for the averaging of variations is that the overall magnitude of the optimum convective heat fluxes is quite robust, but that the estimates for the associated evaporative flux as well as the



**Fig. 6.** Sensitivity of the maximum power state associated with the exchange fluxes caused by the vertical ( $J_{in,s}$ , top row) and the large-scale horizontal ( $J_{in,t} - J_{in,p}$ , bottom row) solar radiation difference to the magnitude of the solar radiative forcing. The left plots show the sensitivity of the power associated with the sensible heat flux ( $G_{sh}$ , dotted line), the latent heat flux ( $G_{lh}$ , grey line), the power involved in lifting ( $G_{lift}$ , dashed grey line), and the sum of all ( $G_c$  and  $G_{ls}$ , solid line). The plots on the right show the sensible heat flux ( $H_c$  and  $H_{ls}$ , dotted line), the latent heat flux ( $\lambda E_c$  and  $\lambda E_{ls}$ , grey line), and the total turbulent heat fluxes ( $J_c$  and  $J_{ls}$ , solid line). The global mean values of the present day are about  $J_{s,in} = 165 \text{ W m}^{-2}$  and  $J_{in,t} - J_{in,p} = 96 \text{ W m}^{-2}$  (although different values for  $k_r$  and  $s$  were used in Figs. 4 and 5.)

power involved is rather sensitive to the radiative properties of the atmosphere.

These sensitivities are computed assuming steady state conditions in which the simulated temperatures are in steady state with the energy fluxes. When the solar radiative forcing varies during the day (or through seasons), then the energy balances are typically not in a steady state. What this averaging over such variations in solar radiation imply for our estimates, we may consider two extreme cases for the sake of simplicity. In the first case, we assume a heat reservoir which absorbs solar radiation with a large heat capacity, such as a water surface. For this case we can assume a large thermal inertia so that the averaging should not play a role. For the second case we consider a heat reservoir which absorbs solar radiation with a very small heat capacity, such as a forest canopy. In this case, we can assume that the steady state between radiative fluxes and energy balances is established fast. Hence, the above considerations of the biases due to temporal variations in the solar radiative forcing would only apply when the thermal inertia of the absorbing surface is relatively small compared to the timescale by which the solar forcing varies.

To sum up, the sensitivities we performed here suggest that spatial and temporal averaging could result in an underestimation of the latent heat flux. These biases are, however, not of a substantial magnitude, because the deviations from linear relationships are not that large. Nevertheless, such averaging could result in an underestimation of the strength of hydrologic cycling. Furthermore, we found that radiative properties of the atmosphere that would affect the strength of radiative exchange play an important role in shaping the strength of hydrologic cycling as well.

## 5 Discussion

### 5.1 Limitations

The estimates derived here on the maximum strength of the hydrologic cycle are, of course, subject to several limitations. Some of these limitations regarding the use of global mean values have already been discussed above in the context of the sensitivities, where we specifically identified the importance of the interaction of large-scale flow and surface exchange fluxes and of radiative characteristics. Other notable

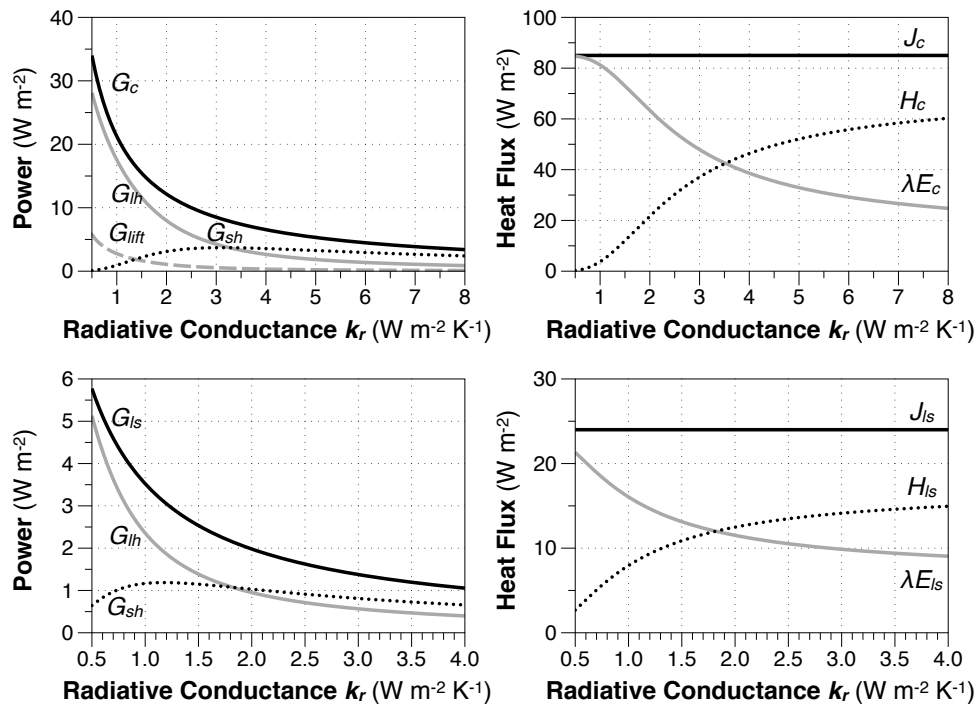


Fig. 7. Same as Fig. 6, but showing the sensitivity to the radiative conductance,  $k_r$ .

limitations are that we did not consider land explicitly, where water availability reduces evaporation rates in some regions, we did not include interactions with cloud cover into the model which would affect the radiative properties of the atmosphere, and the two forcing gradients were treated in isolation. The sensitivity we performed regarding the coupling of these forcing gradients suggests that the maximization of power should really be applied to the whole system that includes both gradients, rather than in separation, as it overall generates more motion and thus, more convective exchange. We also did not link the turbulent dissipation of kinetic energy with the turbulent transport of moisture at the interface between the surface and the atmosphere. Since dissipation of kinetic energy equals the generation of kinetic energy in steady state, a state of maximum power associated with atmospheric motion would likely also be associated with maximum dissipation near the surface and hence maximum turbulent moisture transport within the boundary layer. We did not consider this link explicitly here.

These aspects could be addressed in future work. The addition of extra details would make the models necessarily more complicated and would likely require some empirical parameters. It would also potentially provide a wider range of predictions that could be tested with observations. The estimates developed here can in fact be seen as the simplest possible estimate of the maximum strength of the hydrologic cycle. It contains a minimum representation of the physics, in terms of the energy balances that describe the driving gradients and

which explicitly consider the first and second law of thermodynamics. This formulation has practically no empirical parameters, except for the use of observed values for the solar forcing and typical mean temperatures used to determine  $s$  and  $k_r$ . From this perspective it is quite remarkable how close our estimates are to the observed strength of the hydrologic cycle. The sensitivities also revealed that the associated optimum heat fluxes are quite robust, although the maximum power limit as well as the partitioning between sensible and latent heat can be quite sensitive.

Overall, there are clearly many aspects that could be improved and that could refine our estimates of the maximum strength. The fact that our simple estimates are, nevertheless, in the same order of magnitude as the observed hydrologic cycle suggests that we captured the relevant, first-order processes very well. Furthermore, it implies that the natural hydrologic cycle on earth would seem to operate near its maximum strength.

## 5.2 Relation to previous work

Our results build on and extend earlier work on the thermodynamics of hydrologic cycling. Thermodynamics has been used, for instance, to explain the intensity of hurricanes (Emanuel, 1999), or to describe hydrologic cycling in the context of vertical convection as a dehumidifier (Pauluis and Held, 2002a,b; Pauluis, 2011). Konings et al. (2012) recently explored the thermodynamics of the diurnal growth of

a moist atmospheric boundary layer. These studies also used the Carnot limit to describe the strength of cycling, but they provided a more detailed treatment of the different steps involved in the thermodynamic cycle that generates the work within the atmosphere. Here, we did not consider these individual steps, but rather started with the Carnot limit and treated atmospheric dynamics like a “black box” engine. We then looked at the combination of this limit with the energy balances that shape the driving gradient for the heat engine. By doing so, we do not need to describe the details of the “black box” that is constrained by the Carnot limit. What we show here is that the magnitude of the heat flux through the “engine” is substantial, it affects the magnitude of the driving gradient in steady state, and thereby sets the maximum power limit. This is in contrast to previous studies, who used a prescribed temperature difference as forcing. In this sense, our work complements previous approaches on the thermodynamics of the hydrologic cycle.

The strong interaction between the heat flux and the driving gradient that we found here relates very closely to studies that employed the proposed principle of Maximum Entropy Production (MEP, Paltridge, 1975; Ozawa et al., 2003; Kleidon et al., 2010). The MEP principle states that thermodynamic systems organize into states in which they produce entropy at the maximum possible rate. This principle was previously applied to poleward atmospheric heat transport and could reproduce temperature profiles on earth (Paltridge, 1975, 1978), empirical values for turbulent dissipation (Kleidon et al., 2003, 2006), and the climate characteristics of other planets (Lorenz et al., 2001). It was applied to vertical exchange of radiation and heat within the atmosphere (Ozawa and Ohmura, 1997; Lorenz and McKay, 2003; Kleidon, 2004) to determine vertical fluxes and resulting temperature profiles. The models developed here are very similar to those used in these studies, although the latent heat flux was not explicitly considered and the interpretation of the limit is different. Here, we merely applied the Carnot limit to the heat flux and considered the depleted temperature gradient that is associated with a greater heat flux. This combination of the Carnot limit with the trade-off between flux and gradient resulted in a maximum power state. In steady state, power equals dissipation, and a maximum power state corresponds to a state of maximum dissipation. As already mentioned at the end of Sect. 2.3, if this dissipation occurs at the temperature of the cold sink, the maximum power state used here corresponds to a state of maximum entropy production. Since kinetic energy is typically not dissipated exclusively at the cold sink, it implies that the maximum power state would be close, but somewhat below the MEP state in terms of its entropy production. While the values associated with both states are very similar, the assumption of a maximum power is easier to interpret in classical mechanical terms. Hence, the results obtained from MEP may be more appropriately phrased as a Carnot limit in which the assumption of a fixed, unaffected temperature difference is being relaxed.

At a more general level, our work demonstrates how tightly linked the hydrologic cycle is to the functioning of the whole Earth system and its forcing. The maximum strength of hydrologic cycling is in first order directly proportional to the magnitude of differences in absorption of solar radiation. This is reflected in the estimates for maximum evaporation rates (Eqs. 35 and 45), which are proportional to the surface absorption of solar radiation and the difference in solar radiation between the tropics and the poles. These differences in solar radiative heating result directly from the planetary forcing and are of critical importance because these differences act as the thermodynamic driver that maintains a system away from a state of thermodynamic equilibrium. This disequilibrium can be expressed by the relative humidity. Our approximations for the relative humidity shows that the extent of disequilibrium is directly proportional to the magnitude of this forcing (Eqs. 38 and 47).

The tight linkage of the maximum strength of hydrologic cycling to the solar forcing represents the one side of the interplay between the hydrologic cycle and the Earth system. The other side is represented by the strong effect that the latent heat flux has on the magnitude of its driving gradient. It is this interaction between the heat flux and the driver which shapes the maximum power limit. The maximum power limit, in turn, results practically in an insensitivity of the total heat fluxes to changes in radiative properties of the atmosphere, as shown by the sensitivities to  $k_r$  shown in Fig. 7. This insensitivity of heat fluxes to changes in  $k_r$  is consistent with the observation that atmospheric heat transport is indeed insensitive to such changes and depends primarily on solar radiative forcing, albedo, and planetary geometry (Stone, 1978). Here, the optimum heat fluxes depend solely on the difference in solar absorption,  $J_{in,t} - J_{in,p}$ , which reflect the same properties as those stated by Stone (1978), but in addition it explicitly makes use of the assumption that atmospheric motion operates at its thermodynamic limit. This interplay between a state of thermodynamic disequilibrium associated with hydrologic cycling, the maximum power limit, and strong interactions is consistent with the broad picture of how disequilibrium is generated within the Earth system in general and how Earth system processes follow and accelerate the direction of the second law (Kleidon, 2012).

The insight that the hydrologic cycle appears to operate near its maximum strength is non-trivial. Any engineer would agree that reaching the Carnot limit for an engine is a formidable, if not nearly impossible challenge. This necessarily raises the question of how the natural processes of the hydrologic cycle are organized such that they are able to operate close to this limit. We may get a clue to the answer from a recent approach to understand thermodynamic limits associated with river systems (Kleidon et al., 2013). In this paper we looked at the energetics of water and sediment transport in catchments and suggested that it is through structure formation that river systems can deplete their driving gradient of

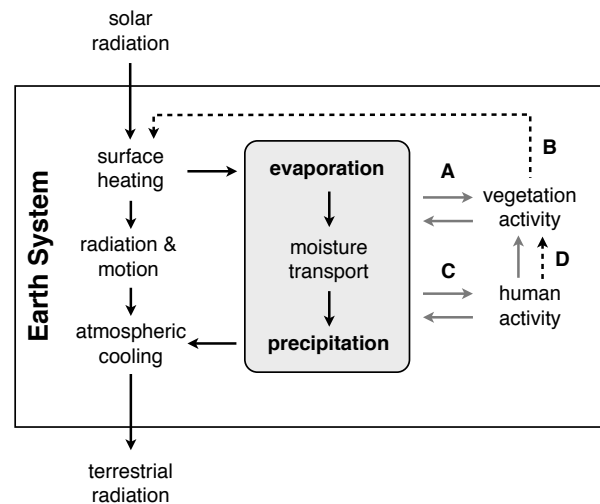


continental topography at the fastest possible rate. Applied to moisture transport within the atmosphere, this would suggest that the maximum power limit is achieved by the spatial, and possibly temporal, organization of moist convection. Clearly, these aspects would need to be explored further in future work.

### 5.3 Implications

These insights emphasize the need to take a thermodynamic as well as a holistic Earth system perspective when we want to understand the functioning of the hydrologic cycle, how it is affected by biotic and human activity, and how the hydrologic cycle changes in response to altered forcing. We first note that an aspect of change can result from two different aspects: it can either lead to a hydrologic flux to operate closer to (or further from) the maximum strength, or it can change the maximum strength itself. To illustrate this distinction with common terminology, the first aspect corresponds to a change that would affect the actual evaporation rate, while the latter aspect corresponds to a change in the potential evaporation rate (which corresponds to the maximum strength, as shown above).

We illustrate these different aspects of change in terms of vegetation effects on hydrologic cycling. We first note that vegetation strongly affects the fluxes of water from the soil into the atmosphere over land. Soil water is taken up by the root system of the vegetation, it is then transported by the vascular system to the canopy, where it is transpired into the atmosphere. These flows of water are represented in Fig. 8 by the arrows “A”. The effect of this biotic “plumbing system” is typically to enhance evapotranspiration rates on land because the ability of a bare soil is typically very low to sustain high evaporation rates during dry episodes. This enhanced ability of the vegetated surface to evaporate water would seem to represent an aspect that would make the evaporative flux on land to operate closer to the maximum strength. An example of vegetation that alters the maximum strength is the effect of vegetation on the radiative properties of the surface (arrow “B” in Fig. 8). Vegetated surfaces generally have a lower surface albedo, which results in a higher absorption of incident solar radiation. Because absorbed solar radiation was a key factor in shaping the maximum strength estimates, the lowering of the surface albedo corresponds to a greater capacity to evaporate water. The effect of this albedo change can be illustrated by sensitivity simulations with a coupled climate-vegetation model (Kleidon, 2006), in which the conditions of a “Desert World” with no vegetation present were simulated. Compared to this “Desert World” state, the mean latent heat flux on land increased by  $26.8 \text{ W m}^{-2}$  while net solar radiation increased by  $9.6 \text{ W m}^{-2}$ . What this seems to imply is that the resulting change in hydrologic cycling can be attributed to a  $9.6/26.8 \approx 36\%$  increase in the capacity to cycle water while the remaining increase is an enhancement of the actual evaporation rate without a change in the



**Fig. 8.** Conceptual diagram to illustrate interactions between vegetation and human activity with the strength of hydrologic cycling within an Earth system context, which is based on Fig. 1. Grey arrows “A” and “C” reflect the fluxes of water associated with vegetation and human activity, which affect how close hydrologic fluxes operate to the maximum that is set mostly by the radiative forcing. The dashed arrow “B” describes the effects of vegetation on the radiative forcing which alters the maximum strength limit and thereby the capacity of the Earth system to cycle water. Human activity can also affect this limit through vegetation changes, e.g. by land cover change (arrow “D”).

maximum strength. At the same time it is important to note that this increase by vegetation activity comes at an energetic cost of creating and maintaining biomass, which is derived from photosynthesis. It is this vegetation activity that builds and maintains the “plumbing system” as well as the other effects described above that shape the effects on hydrologic cycling. Ultimately, one would also need to consider whether these effects would enhance, or reduce, the ability of vegetation to perform photosynthesis as it is through photosynthesis that vegetation activity is being maintained.

We can also apply this line of reasoning to human activity, which requires freshwater for a variety of reasons, including drinking water, sanitation, irrigation, and industrial uses. This freshwater is taken out, and eventually returned, to the natural hydrologic cycle (represented by arrow “C” in Fig. 8). The largest fraction of freshwater use is associated with irrigation of croplands (Oki and Kanae, 2006), with different sources of where the water from irrigation comes from (diversion of river flow, groundwater or desalination of seawater, with the latter two sources associated with energy needs in the process). In the context here, irrigated croplands result in two different types of change (represented by arrow “D”). First, more water is made available for evaporation, so that evaporation could operate closer to its limit. Second, cropland is associated with land cover change from its natural state. The change in land cover would likely impact

the maximum strength for water cycling. If the natural state was forested, the change to cropland would likely be associated with an increase in the surface albedo, which would reduce the maximum strength limit. If the irrigated cropland is located in a desert region, the surface albedo is likely to decrease, which would potentially increase the maximum strength limit. Hence, it is not directly clear whether human activities would tend to reduce or enhance the strength of hydrologic cycling as well as its limit. In this context it is important to recognize that hydrologic fluxes do not operate in the context of a static, natural limit (which the concept of a “planetary boundary” (Rockström et al., 2009) would suggest), but that the limits are in principle affected by human activity as well.

## 6 Summary and conclusions

We derived the thermodynamic limits of global hydrologic cycling from a simple representation of the hydrologic cycle within the Earth system that can be seen to represent the minimum amount of required physics. This minimum representation includes the energy balances that form the driving gradient for motion, which is needed to transport moisture, and serve as the source (or sink) of energy for the phase transitions associated with hydrologic cycling as well as the first and second law of thermodynamics, which yield the limits to motion and transport. The combined need for energy for evaporation and for atmospheric transport then established the limit to the strength of the hydrologic cycle. This limit results from the strong interaction between hydrologic fluxes with the heating gradient that drives atmospheric transport. When we used observed values for the forcing, we obtained estimates for the maximum strength of hydrologic cycling that are close to the observed state of the hydrologic cycle. We conclude that the natural hydrologic cycle within the Earth system appears to operate near its maximum, thermodynamic strength.

Our work is only a first demonstration of the utility of a thermodynamic view, as well as a holistic view on the hydrologic cycle within the Earth system. The thermodynamic view emphasizes the role of a thermodynamic driver that is needed to generate dynamics associated with hydrologic cycling and its associated disequilibrium. The holistic view includes the interactions of hydrologic cycling with other Earth system processes and which links hydrologic processes to the ultimate drivers of planetary dynamics. In future work, this view should be developed further. The use of spatially and temporally explicit datasets of climatic forcing would allow for a much more detailed comparison of the maximum strength predictions with observations. The effects of vegetation activity and, more generally, global change on the hydrologic cycle could be analyzed in terms of their effects on the maximum strength, as outlined in the discussion. The establishment of such maximum strength limits of hydrologic

processes and their sensitivity to change should help us to better understand and predict how the hydrologic cycle has changed in the past, and how it is likely to change in the future.

## Appendix A

### A1 Entropy exchange in a steady state system

To derive the entropy exchange by heat exchange of a system in steady state, let us consider a system of two heat reservoirs, one being heated and maintained at a temperature  $T_h$ , and the other being cooled and maintained at a temperature  $T_c$ . Both reservoirs are considered in steady state, so that both temperatures are constant, i.e.  $dT_h = 0$  and  $dT_c = 0$ . The steady state implies that the heat added to the hot reservoir of the system,  $dQ_{in}$ , balances the heat transfer within the system from the hot to the cold reservoir,  $dQ_{h,c} = dQ_{in}$ , and the heat transferred to the cold reservoir within the system balances the removal of heat from the system,

$$dQ_{out} = dQ_{h,c} = dQ_{in}. \quad (A1)$$

To evaluate the entropy exchange associated with these heat transfer processes, we use  $dT_h = 0$  and  $dT_c = 0$  because of the steady state assumption. Then, the change in heat content,  $dQ$ ,

$$dQ = d(TS) = TdS + SdT \quad (A2)$$

simplifies to  $dQ = TdS$ . The heat transfer process between the two reservoirs removes heat from the hot reservoir,  $dQ_{h,c} = T_h dS_h$ , and adds this heat to the cold reservoir,  $dQ_{h,c} = T_c dS_c$ . The increase of entropy,  $dS_{h,c}$ , caused by this transfer of heat from hot to cold within the system is hence

$$dS_{h,c} = dS_c - dS_h = dQ_{h,c} \left( \frac{1}{T_c} - \frac{1}{T_h} \right) \quad (A3)$$

or, with  $dQ_{h,c} = dQ_{in}$  in steady state,

$$dS_{h,c} = dQ_{in} \left( \frac{1}{T_c} - \frac{1}{T_h} \right) > 0 \quad (A4)$$

which is greater than zero because  $T_c < T_h$ .

When we consider the entropy that is exchanged by this system with the surroundings, then the addition of heat to the system,  $dQ_{in}$ , adds entropy,  $dS_{in}$ , to the system. From  $dQ_{in} = T_h dS_{in}$  we obtain  $dS_{in} = dQ_{in}/T_h$ . The removal of heat removes entropy from the system, which we obtain from  $dQ_{out} = T_c dS_{out}$  so that  $dS_{out} = dQ_{out}/T_c$ . Overall, the net entropy exported by the system due to the heat exchange with the surroundings is

$$dS_{ex} = dS_{out} - dS_{in} = dQ_{in} \left( \frac{1}{T_c} - \frac{1}{T_h} \right) > 0 \quad (A5)$$

since  $dQ_{\text{out}} = dQ_{\text{in}}$ , and  $T_h > T_c$ . Furthermore, we note that

$$dS_{\text{ex}} = dS_{h,c}, \quad (\text{A6})$$

i.e., the entropy produced within the system balances the net entropy exported by the system to the surroundings.

When these changes are considered during a small time interval  $dt$ , then we use  $J_{\text{in}} = dQ_{\text{in}}/dt$  for heat fluxes and  $J_{\text{net}}^s = dS_{\text{ex}}/dt$  for the entropy exchange. The entropy production,  $\sigma = dS_{h,c}/dt$  within the system then balances the net entropy exchange and is simply given by

$$\sigma = J_{\text{net}}^s = J_{\text{in}} \left( \frac{1}{T_c} - \frac{1}{T_h} \right). \quad (\text{A7})$$

## A2 Linear approximations

We derive simple, linearized approximations for terrestrial radiation and for the saturation vapor pressure that are used in the models. These approximations allow us to derive very simple, analytical solutions for the limits and this facilitates the evaluation of the first order dynamics that shape these limits. We should nevertheless recognize that these models cannot be highly precise, and serve the purpose to illustrate the limits as well as to provide a first order estimate of the magnitude of the limits.

To obtain a linearized expression for net emission of long-wave radiation, we do a Taylor expansion to first order of the Stefan–Boltzmann law:

$$J_l(T) \approx \sigma T_0^4 + \left( \frac{d}{dT} (\sigma T^4) \right)_{T=T_0} \cdot (T - T_0) \quad (\text{A8})$$

with a reference temperature  $T_0$ ,  $\sigma = 5.67 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$  being the Stefan–Boltzmann constant, and a slope that will be abbreviated by  $k_r$  in the following:

$$k_r = \left( \frac{d}{dT} (\sigma T^4) \right)_{T=T_0} = 4\sigma \cdot T_0^3. \quad (\text{A9})$$

With this linearization, the net radiative exchange between a warm reservoir at temperature  $T_h$  and a cool reservoir at a temperature  $T_c$  is then given by

$$J_{h,c} = J_l(T_h) - J_l(T_c) = k_r \cdot (T_h - T_c) \quad (\text{A10})$$

assuming that both emit like black bodies, i.e., that their emissivities are  $\epsilon = 1$ .

The saturation vapor pressure  $e_{\text{sat}}$  is a near-exponential function of air temperature,  $T$  (Bohren and Albrecht, 1998):

$$e_{\text{sat}}(T) = e_0 \cdot \exp[19.83 - 5417 \text{ K}/T] \quad (\text{A11})$$

with a value of  $e_0 = 6.11 \text{ hPa}$  and  $T$  in units of K. We obtain a linearized expression by a Taylor expansion around a reference temperature  $T_0$ :

$$e_{\text{sat}}(T) \approx e_{\text{sat}}(T_0) + \left( \frac{de_{\text{sat}}}{dT} \right)_{T=T_0} \cdot (T - T_0). \quad (\text{A12})$$

Here, the slope of the vapor pressure curve, evaluated at the reference temperature  $T_0$ , will be abbreviated by  $s$ :

$$s = \left( \frac{de_{\text{sat}}}{dT} \right)_{T=T_0} = \frac{s_0}{T_0^2} \cdot \exp[19.83 - 5417/T_0] \quad (\text{A13})$$

with  $s_0 = e_0 \cdot 5417 \text{ K}$ . With this linearization, the difference in saturation vapor pressure,  $\Delta e_{\text{sat}}$  between two air masses with temperatures  $T_h$  and  $T_c$  is then simply linear in the associated temperature difference:

$$\Delta e_{\text{sat}} = s \cdot (T_h - T_c). \quad (\text{A14})$$

*Acknowledgements.* The ideas for this paper originated from the preparation for the “Biogeophysics and Earth System Science” (BESS) summer school. We thank one anonymous reviewer and Hubert Savenije for constructive reviews that improved the manuscript. A. Kleidon acknowledges financial support from the Helmholtz Alliance “Planetary Evolution and Life”. This research contributes to the “Catchments As Organized Systems (CAOS)” research group funded by the German Science Foundation (DFG).

The service charges for this open access publication have been covered by the Max Planck Society.

Edited by: M. Sivapalan

## References

- Bohren, C. F. and Albrecht, B. A.: Atmospheric Thermodynamics, Oxford Univ. Press, New York, 1998.
- Chahine, M. T.: The hydrologic cycle and its influence on climate, *Nature*, 359, 373–380, 1992.
- Emanuel, K. A.: Thermodynamic control of hurricane intensity, *Nature*, 401, 665–669, 1999.
- Goody, R.: Sources and sinks of climate entropy, *Q. J. R. Meteorol. Soc.*, 126, 1953–1970, 2000.
- Kleidon, A.: Beyond Gaia: Thermodynamics of life and Earth system functioning, *Clim. Ch.*, 66, 271–319, 2004.
- Kleidon, A.: The climate sensitivity to human appropriation of vegetation productivity and its thermodynamic characterization, *Glob. Planet. Ch.*, 54, 109–127, 2006.
- Kleidon, A.: Life, Hierarchy, and the Thermodynamic Machinery of Planet Earth, *Phys. Life Rev.*, 7, 424–460, 2010.
- Kleidon, A.: How does the Earth system generate and maintain thermodynamic disequilibrium and what does it imply for the future of the planet?, *Phil. Trans. R. Soc. A*, 370, 1012–1040, 2012.
- Kleidon, A., Fraedrich, K., Kunz, T., and Lunkeit, F.: The atmospheric circulation and states of maximum entropy production, *Geophys. Res. Lett.*, 30, 2223, doi:10.1029/2003GL018363, 2003.
- Kleidon, A., Fraedrich, K., Kirk, E., and Lunkeit, F.: Maximum Entropy Production and the Strength of Boundary Layer Exchange in an Atmospheric General Circulation Model, *Geophys. Res. Lett.*, 33, L06706, doi:doi:10.1029/2005GL025373, 2006.
- Kleidon, A., Malhi, Y., and Cox, P. M.: Maximum entropy production in environmental and ecological systems, *Phil. Trans. R. Soc. B*, 365, 1297–1302, 2010.

- Kleidon, A., Zehe, E., Ehret, U., and Scherer, U.: Thermodynamics, maximum power, and the dynamics of preferential river flow structures at the continental scale, *Hydrol. Earth Syst. Sci.*, 17, 225–251, doi:10.5194/hess-17-225-2013, 2013.
- Konings, A. G., Feng, X., Molini, A., Manzoni, S., Vico, G., and Porporato, A.: Thermodynamics of an idealized hydrologic cycle, *Water Resour. Res.*, 48, W05527, doi:doi:10.1029/2002RG000113, 2012.
- Lorenz, E. N.: Available potential energy and the maintenance of the general circulation, *Tellus*, 7, 157–167, 1955.
- Lorenz, R. D. and McKay, C. P.: A simple expression for vertical convective fluxes in planetary atmospheres, *Icarus*, 165, 407–413, 2003.
- Lorenz, R. D., Lunine, J. I., Withers, P. G., and McKay, C. P.: Titan, Mars and Earth: Entropy production by latitudinal heat transport, *Geophys. Res. Lett.*, 28, 415–418, 2001.
- Masuda, K.: Meridional heat transport by the atmosphere and the ocean: analysis of FGGE data, *Tellus A*, 40, 285–302, 1988.
- Oki, T. and Kanae, S.: Global hydrological cycle and world water resources, *Science*, 313, 1068–1072, 2006.
- Oki, T., Musiak, K., Matsuyama, H., and Masuda, K.: Global atmospheric water balance and runoff from large river basins, *Hydrol. Process.*, 9, 655–678, 1995.
- Ozawa, H. and Ohmura, A.: Thermodynamics of a global-mean state of the atmosphere – A state of maximum entropy increase, *J. Clim.*, 10, 441–445, 1997.
- Ozawa, H., Ohmura, A., Lorenz, R. D., and Pujol, T.: The second law of thermodynamics and the global climate system – A review of the Maximum Entropy Production principle, *Rev. Geophys.*, 41, 1018, doi:doi:10.1029/2011WR011264, 2003.
- Paltridge, G. W.: Global dynamics and climate – a system of minimum entropy exchange, *Q. J. Roy. Meteorol. Soc.*, 101, 475–484, 1975.
- Paltridge, G. W.: The steady-state format of global climate, *Q. J. Roy. Meteorol. Soc.*, 104, 927–945, 1978.
- Pauluis, O.: Water vapor and mechanical work: a comparison of Carnot and steam cycles, *J. Atmos. Sci.*, 68, 91–102, 2011.
- Pauluis, O. and Held, I. M.: Entropy budget of an atmosphere in radiative convective equilibrium, Part I: Maximum work and frictional dissipation, *J. Atmos. Sci.*, 59, 126–139, 2002a.
- Pauluis, O. and Held, I. M.: Entropy budget of an atmosphere in radiative convective equilibrium. Part II: Latent heat transport and moist processes, *J. Atmos. Sci.*, 59, 140–149, 2002b.
- Pauluis, O., Balaji, V., and Held, I. M.: Frictional dissipation in a precipitating atmosphere, *J. Atmos. Sci.*, 57, 987–994, 2000.
- Peixoto, J. P. and Oort, A. H.: *Physics of Climate*, American Institute of Physics, New York, NY, 1992.
- Penman, H. L.: Natural evaporation from open water, bare soil and grass, *Proc. R. Soc. Lond. A*, 193, 120–146, 1948.
- Priestley, C. H. B. and Taylor, R. J.: On the assessment of surface heat flux and evaporation using large-scale parameters, *Mon. Weather Rev.*, 100, 81–92, 1972.
- Renno, N. O. and Ingersoll, A. P.: Natural convection as a heat engine: A theory for CAPE, *J. Atmos. Sci.*, 53, 572–585, 1996.
- Rockström, J., Steffen, W., Noone, K., Persson, A., Chapin, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. S., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R. W., Fabry, V. J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., and Foley, J. A.: A safe operating space for humanity, *Nature*, 461, 472–475, 2009.
- Slayter, R. O. and McIlroy, I. C.: *Practical Micrometeorology*, CSIRO, Melbourne, Australia, 310 pp., 1961.
- Stephens, G. L., Li, J., Wild, M., Clayson, C. A., Loeb, N., Kato, S., L'Ecuyer, T., Stackhouse, P. W., Lebsock, M., and Andrews, T.: An update on Earth's energy balance in light of the latest global observations, *Nature Geosci.*, 5, 691–696, 2012.
- Stone, P. H.: Constraints on dynamical transports of energy on a spherical planet, *Dynam. Atmos. Ocean.*, 2, 123–139, 1978.
- Trenberth, K. E. and Caron, J. M.: Estimates of meridional atmosphere and ocean heat transports, *J. Climate*, 14, 3433–3443, 2001.