Bjerknes compensation in a coupled global box model

Jiaqi Shi[1] · Haijun Yang[2]

Abstract

The Earth climate system has an intrinsic mechanism to maintain its energy conservation by impelling opposite changes in meridional ocean and atmosphere heat transports. This mechanism is briefed as the Bjerknes compensation (BJC). We set up a global coupled two-hemisphere box model in this study, and obtain an analytical solution to the BJC of this system. In the two-hemisphere model, the thermohaline circulation is interhemispheric and parameterized by the density difference between two polar boxes. The symmetric poleward atmosphere heat and moisture transports are considered and parameterized by the temperature gradient between tropical and polar boxes. Different from the BJC in the one-hemisphere box model that depends only on the local climate feedback, the BJC here is determined by both local climate feedback and temperature change. The asymmetric thermohaline circulation leads to a better BJC in the Northern Hemisphere than in the Southern Hemisphere. Furthermore, an analytical solution to the probability of a valid BJC (i.e., negative BJC) is derived, which is determined only by the local climate feedback. The probability of a valid BJC is very high under reasonable climate feedback. Based on observational data, the climate feedback can be estimated and the BJC in real world can be calculated using the analytical formulae. It is found that at the decadal and longer timescales the BJC for both hemispheres are robust in reality.

Keywords Bjerknes compensation · Box model · Climate system · Heat transport

1 Introduction

Meridional heat transport (MHT) of the Earth climate system plays a critical role in maintaining the balance of global climate. In the mid-latitudes around 35°N/S, up to 5.5 PW (1 PW = 10^{15} W) energy can be transported poleward from the tropics through the joint efforts of atmosphere and ocean (Trenberth and Caron 2001). There are two fundamental questions regarding the MHT: one is the partitioning of poleward energy transport between the atmosphere and ocean in terms of mean climatology, and the other is the relationship between the changes of atmosphere heat transport (AHT) and ocean heat transport (OHT) in terms of climate change and climate variability. The first question has been widely

explored in the past two decades (e.g., Held 2001; Czaja and Marshall 2006; Vallis and Farneti 2009; Farneti and Vallis 2013). The second question was first proposed by Bjerknes (1964), which stated that as long as the net radiation at the top of the atmosphere (TOA) and ocean heat content do not vary apparently, the changes in AHT and OHT should be of the same magnitude with opposite signs, that is, the so-called Bjerknes compensation (BJC) occurs. Under the perfect BJC, the total MHT of the Earth system should remain unchanged. Though the idea of BJC was proposed in the 1960 s, the intrinsic mechanism controlling the BJC is only disclosed in recent years (Liu et al. 2016, 2018; Yang et al. 2016; Zhao et al. 2016).

Although it remains to be validated by observational data, the BJC has been well recognized in various models, ranging from earlier simple energy balance models (EBMs) (e.g., Lindzen and Farrell 1977; Stone 1978; North 1984) to the state of the art coupled general circulation models (GCMs) (e.g., Zhang and Delworth 2005; Shaffrey and Sutton 2006; Swaluw et al. 2007; Kang et al. 2008, 2009; Vellinga and Wu 2008; Vallis and Farneti 2009; Frierson and Huang 2012; Donohoe et al. 2013; Farneti and Vallis 2013; Rose and Ferreira 2013). The BJC is usually valid on decadal and longer

✉ Haijun Yang
yanghj@fudan.edu.cn

1 Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China

2 Department of Atmospheric and Oceanic Sciences, Institute of Atmospheric Science and CMA- FDU Joint Laboratory of Marine Meteorology, Fudan University, 2005 Songhu Road, Shanghai 200438, China

timescales (Zhao et al. 2016). To quantify the extent of the compensation, the BJC rate is defined as the ratio of the changes in AHT and OHT. Regional climate feedback is the most critical factor in determining the BJC (Stone 1978; Langen and Alexeev 2007; Enderton and Marshall 2009; Rose and Ferreira 2013). In Liu et al. (2016) and Yang et al. (2016), we derived an analytical solution to the BJC rate, showing explicitly that the BJC is only determined by the local climate feedback, and independent of temperatures, circulations and heat transports of atmosphere and ocean themselves. We concluded that the constrain of the energy conservation of the climate system is the intrinsic mechanism to the occurrence of BJC, whose magnitude is determined by the climate feedback.

Previous simple model's studies on the BJC use a single hemisphere model that was first proposed by Stommel (1961). In Stommel's model, the thermohaline circulation (THC) is parameterized to be proportional to the meridional density contrast between tropical and polar boxes (e.g., Nakamura et al. 1994; Marotzke and Stone 1995; Yang et al. 2016), transporting mass and heat from the tropics to the Northern Hemisphere (NH) subpolar ocean. In reality, the THC spans the two hemispheres, transporting mass and heat all the way from the Southern Ocean to the NH subpolar ocean. This structure of the THC leads to asymmetric heat transports between the two hemispheres, and it may result in different BJC in the two hemispheres. Therefore, in this work a two-hemisphere box model is used to study the BJC. The two-hemisphere model was first proposed by Rooth (1982), which is thought to be dynamically superior to

the Stommel's model by many researchers (e.g., Scott et al. 1999; Longworth et al. 2005) for theoretical studies of the multi-equilibrium and stability of the THC.

In the two-hemisphere model, the THC is parameterized by the density difference between two polar boxes (Scott et al. 1999). What would the BJC be in the two-hemisphere box model? What are the similarities and differences of BJC between the two hemispheres, considering the asymmetric THC between the NH and Southern Hemisphere (SH)? We attempt to untangle these questions in this paper. Based on Rooth' model, we design a 9-box coupled model spanning both hemispheres in this study (Fig. 1). Different from the 4-box model in Yang et al. (2016), the latitude range of the two-hemisphere box model in this study extends from 70°S to 75°N, and is divided into three regions as the NH extratropics, tropics and SH extratropics.

Two BJC rates are derived for the NH and SH, respectively. Different from that in the one-hemisphere model, the BJC here depends on both local climate feedback and surface temperature pattern. Yet, the BJC in the NH is still consistent with that in the one-hemisphere box model (Liu et al. 2016; Yang et al. 2016), that is, a positive (negative) feedback in the NH causes overcompensation (undercompensation) and the BJC rate can be reasonably estimated based merely on the climate feedback. The BJC in the SH, however, tends to be always overcompensated regardless the sign of the local climate feedback. The asymmetric THC leads to different BJC behaviors between the two hemispheres. We also derive a formula for the probability of a valid BJC. It is found that the probability of a valid BJC depends only on the local climate feedback,
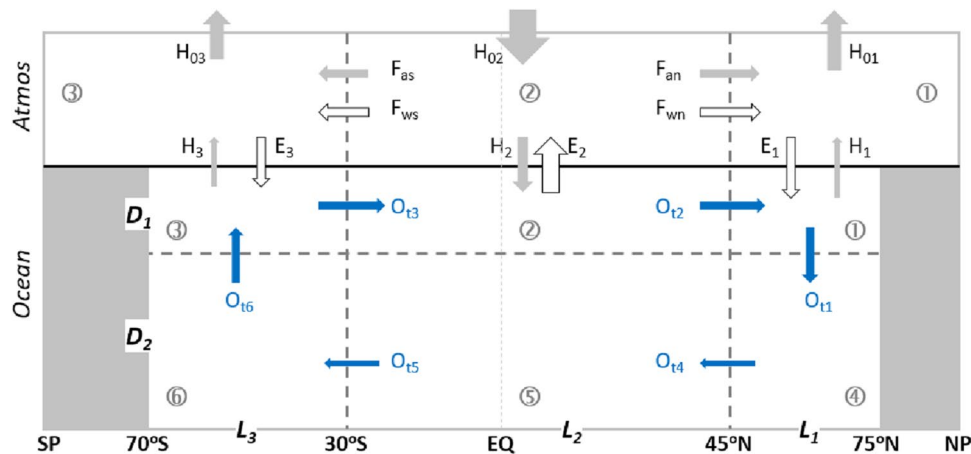

Fig. 1 Schematic diagram of the coupled box model. The model consists of three atmospheric boxes and six ocean boxes that are denoted by ①, ②,…,⑥. Boxes 1 and 4 represent the upper and lower layers of the northern extratropical ocean, respectively; boxes 2 and 5, of the tropical ocean, respectively; and boxes 3 and 6, of the southern extratropical ocean, respectively. D_1 and D_2 are the depths of upper and lower ocean layers, respectively. L_1, L_2 and L_3 are the meridional scales of ocean boxes. H_1, H_2 and H_3 are ocean heat gains through the surface, and H_{01}, H_{02} and H_{03} are the net heat radiation at the

top of atmosphere (TOA). E_1 and E_3 are the net freshwater gains in the northern and southern extratropics, respectively, and E_2 is the net freshwater loss in the tropics. O_{t1} - O_{t6} illustrate the heat transports through thermohaline circulation among different ocean boxes, denoted by solid blue arrows. F_{an} and F_{as} are the meridional atmosphere energy transports; F_{wn} and F_{ws} are the meridional atmosphere moisture transport; the subscripts n and s denote northward and southward, respectively

and independent of temperature changes. The probability of a valid BJC is very high under reasonable climate feedback parameters, in both box model and observations.

This work complements previous theoretical studies of the BJC in a one-hemisphere model (Liu et al. 2016; Yang et al. 2016). This paper is organized as follows. In Sect. 2, the 9-box model is introduced and the BJC for both hemispheres is derived. In Sect. 3, perturbation experiments are performed to validate the BJC in the two hemispheres. In Sect. 4, the probability of a valid BJC is derived theoretically. In Sect. 5, the BJC is evaluated based on observational data. Summary and discussion are given in Sect. 6.

2 Two-hemisphere box model and the BJC

2.1 Basic equations

The 9-box coupled model consists of six ocean boxes and three atmosphere boxes (Fig. 1). The atmosphere model spans globally. The ocean model includes the THC explicitly, covering the region from 70°S to 75°N. The atmosphere and ocean boxes are separated into three zones by latitudes 45°N and 30°S, respectively. Here, the NH preference of the intertropical convergence zone (ITCZ) is considered, since the tropical box (30°S–45°N) is not symmetric about the equator, with its central latitude near10°N. The net radiative forcing at the TOA is negative in the extratropical boxes and positive in the tropical box. The atmosphere is assumed to be always in equilibrium with the surface ocean. The ocean model was originally designed by Stommel (1961), and then applied in many studies in different forms (e.g., Marotzke 1990; Huang et al. 1992; Nakamura et al. 1994; Tziperman et al. 1994; Marotzke and Stone 1995; Yang et al. 2016; Zhao et al. 2016). The ocean model used in this work is a two-hemisphere model (Rooth 1982), which was widely used in later studies on the interhemispheric THC's multi-equilibrium and stability (e.g., Rahmstorf 1996; Scott et al. 1999; Longworth et al. 2005). More details can be found in the publications mentioned above. In this work, the ocean is further divided into the upper ocean and lower ocean; so the ocean model has six boxes, which can explicitly describe the overturning THC. The final forms of the equations for the 6-box ocean system can be written as follows:

$$m_1\dot{T}_1 = \frac{1}{\epsilon c \rho_0 D_1}\left[(A_1 - B_1 T_1) + \chi(T_2 - T_1)\right] + q(T_2 - T_1)$$ (1a)

$$m_2\dot{T}_2 = \frac{1}{\epsilon c \rho_0 D_1}\left[(A_2 - B_2 T_2) - \chi(T_2 - T_1) + \chi(T_3 - T_2)\right] + q(T_3 - T_2)$$ (1b)

$$m_3\dot{T}_3 = \frac{1}{\epsilon c \rho_0 D_1}\left[(A_3 - B_3 T_3) - \chi(T_3 - T_2)\right] + q(T_6 - T_3)$$ (1c)

$$m_4\dot{T}_4 = q(T_1 - T_4)$$ (1d)

$$m_5\dot{T}_5 = q(T_4 - T_5)$$ (1e)

$$m_6\dot{T}_6 = q(T_5 - T_6)$$ (1f)

$$m_1\dot{S}_1 = \frac{S_0\gamma}{\epsilon_w D_1}\left[(T_1 - T_2)\right] + q(S_2 - S_1)$$ (2a)

$$m_2\dot{S}_2 = \frac{S_0\gamma}{\epsilon_w D_1}\left[(T_2 - T_1) - (T_3 - T_2)\right] + q(S_3 - S_2)$$ (2b)

$$m_3\dot{S}_3 = \frac{S_0\gamma}{\epsilon_w D_1}\left[(T_3 - T_2)\right] + q(S_6 - S_3)$$ (2c)

$$m_4\dot{S}_4 = q(S_1 - S_4)$$ (2d)

$$m_5\dot{S}_5 = q(S_4 - S_5)$$ (2e)

$$m_6\dot{S}_6 = q(S_5 - S_6)$$ (2f)

where m_i is the ratio of each ocean-box volume with respect to box 1, $m_1 = 1$, $m_2 = \frac{L_2}{L_1}$, $m_3 = \frac{L_3}{L_1}$, $m_4 = \frac{D_2}{D_1}$, $m_5 = \frac{L_2 D_2}{L_1 D_1}$, and $m_6 = \frac{L_3 D_2}{L_1 D_1}$. A_i and B_i ($i = 1, 2, 3$) are area-weighed net incoming radiation (Wm^{-2}) and climate feedback ($Wm^{-2}K^{-1}$), respectively. χ and γ are the bulk coefficients of atmosphere heat and moisture transports, respectively, which are related to the mean atmospheric circulations and eddy activities in the mid-to-high latitudes. c is the seawater specific heat capacity; ρ_0 is seawater density; S_0 is constant reference salinity (35 psu); q is the volume transport by the THC. Relative ocean coverages in all three areas are approximated to be the same, indicated by $\epsilon = G_1/G_{01}$. Here, G_{01} is the entire area of atmosphere box 1, and G_1 are the area of corresponding ocean box 1. ϵ_w indicates the ratio of ocean and catchment area, $\epsilon_w = G_1'/G_{01}$, where G_1' is the ocean and catchment area of the ocean basin. Table 1 lists all the parameters used in this study, mainly based on previous box-model studies (Nakamura et al. 1994; Marotzke and Stone 1995; Yang et al. 2016) and coupled model simulations (Yang et al. 2017), as well as observations (Rayner et al. 2003; Carton and Giese 2008).

The volume transport q by the THC is assumed to be linearly proportional to the density difference between the two extratropical boxes, as in Rooth (1982) and Scott et al. (1999):

Table 1 Parameters used in this study

Symbol	Physical meaning	Value
A_1, A_2, A_3	Net incoming radiative fluxes at boxes 1, 2, 3	$-55, 80, -30 \text{ Wm}^{-2}$
B_1, B_2, B_3	Local climate feedback parameters in boxes 1, 2, 3	$-0.6, 1.7, -0.5 \text{ Wm}^{-2}\text{K}^{-2}$
L_1, L_2, L_3	Meridional scale of north, tropical and south boxes	$30°, 75°, 40°$
D_1, D_2	Depths of upper and lower boxes	400, 4000 m
$c\rho_0$	Heat capacity of a unit water volume	$4 \cdot 10^6 \text{ Jm}^{-3} \text{ K}^{-1}$
S_0	Reference salinity	35.0 psu
α	Thermal expansion coefficient	$2.5 \cdot 10^{-4} \text{ K}^{-1}$
β	Haline contraction coefficient	$7.5 \cdot 10^{-4} \text{ psu}^{-1}$
G_{01}	Entire surface area north of 45°N	$1.25 \cdot 10^{14} \text{ m}^2$
ϵ	Ratio of ocean area of box G1 to box G01 (i.e., G1/G01),$\epsilon \leq 1$	0.2
ϵ_w	Ratio of ocean and catchment area to G01, $\epsilon \leq \epsilon_w \leq 1$	0.3
κ	Hydraulic constant	$3 \times 10^{-6}\text{s}^{-1}$
γ	Atmosphere moisture transport efficiency	$1.6 \cdot 10^{-10} \text{ ms}^{-1} \text{ K}^{-1}$
χ	Atmosphere heat transport efficiency	$1.7 \text{ Wm}^{-2}\text{K}^{-1}$

$$q = \kappa\left[\alpha\left(T_3 - T_1\right) - \beta\left(S_3 - S_1\right)\right] \tag{3}$$

where $\kappa(s^{-1})$ is a hydraulic constant; α and β are the thermal and haline expansion coefficients of seawater, respectively. The THC is simplified as a meridionally enclosed circulation, sinking in the NH extratropics and rising in the SH extratropics. Accompanying the mass transport, meridional heat transport is northward in the upper ocean and southward in the lower layer. Thus, the meridional OHT for the whole ocean depth is calculated as the difference of the heat transport in the upper and lower oceans. In the NH and SH, they are parameterized, respectively, as follows:

$$O_{tn} = \epsilon c\rho_0 D_1 G_{01} q(T_2 - T_4), \quad O_{ts} = \epsilon c\rho_0 D_1 G_{01} q(T_3 - T_5) \tag{4}$$

Based on the widely used Budyko-type model (Budydo, 1969), the meridional AHTs in the NH and SH mid-latitudes can be simply written as follows:

$$F_{an} = \chi G_{01} T_{sn}, \quad F_{as} = \chi G_{01} T_{ss} \tag{5}$$

Here, $T_{sn} = T_2 - T_1$ and $T_{ss} = T_2 - T_3$, denoting the upper ocean meridional temperature gradient in the NH and SH, respectively. This parameterization is appropriate for the atmosphere in mid-to-high latitudes (Stone and Yao 1990).

The net radiation flux at the TOA has a simple linear relation to surface temperature, which is widely employed in EBMs:

$$H_{01} = A_1 - B_1 T_1, \quad H_{02} = A_2 - B_2 T_2, \quad H_{03} = A_3 - B_3 T_3. \tag{6}$$

The surface heat fluxes at the air–sea interface are:

$$H_1 = \frac{1}{\epsilon c\rho_0 m_1 D_1}\left[\left(A_1 - B_1 T_1\right) + \chi T_{sn}\right], \tag{7a}$$

$$H_2 = \frac{1}{\epsilon c\rho_0 m_2 D_1}\left[\left(A_2 - B_2 T_2\right) - \chi(T_{sn} + T_{ss})\right], \tag{7b}$$

$$H_3 = \frac{1}{\epsilon c\rho_0 m_3 D_1}\left[\left(A_3 - B_2 T_3\right) + \chi T_{ss}\right]. \tag{7c}$$

Note that in (6) and (7), a positive (negative) value of B_i represents a negative (positive) climate feedback.

The Budyko-type model can also be applied to the meridional moisture transport in the atmosphere. In other words, the moisture transport is also assumed to be linearly proportional to the meridional temperature gradient and parameterized as follows,

$$F_{wn} = \gamma G_{01} T_{sn}, \quad F_{ws} = \gamma G_{01} T_{ss}, \tag{8}$$

which should be balanced by the net freshwater loss (gain) at low (high) latitudes.

2.2 Equilibrium state

The ocean heat budget as a whole is only determined by the net radiative forcing at the TOA:

$$\sum_{i=1}^{6} m_i \dot{T}_i = \sum_{i=1}^{3} m_i H_i = \frac{1}{\epsilon c\rho_0 D_1} \sum_{i=1}^{3} H_{0i}. \tag{9}$$

For the steady state, the total energy in the whole coupled box model is conserved, that is,

$$\sum_{i=1}^{3} m_i H_i = \sum_{i=1}^{3} H_{0i} = 0, \tag{10}$$

which depicts that the ocean uptake in the tropical latitudes is equal to the ocean heat release in the extratropical latitudes.

Without external freshwater source, the total salt content of the ocean in the box model is conserved:

$$\sum_{i=1}^{6} m_i \dot{S}_i = 0. \tag{11}$$

The equilibrium states of temperature and salinity, as well as AHT and OHT, can be obtained by letting the temporal tendency be 0 ($\dot{T}_i = \dot{S}_i = 0$):

$$B_1 T_1 + B_2 T_2 + B_3 T_3 = A_1 + A_2 + A_3 \tag{12}$$

$$T_1 = T_4 = T_5 = T_6, \quad S_1 = S_4 = S_5 = S_6 \tag{13}$$

$$O_{tn} = -G_{01}(H_{01} + \chi T_{sn}), \quad O_{ts} = -G_{01}(H_{03} + \chi T_{ss}) \tag{14}$$

$$F_{an} = \chi G_{01} T_{sn}, \quad F_{as} = \chi G_{01} T_{ss} \tag{15}$$

$$\begin{aligned} F_{tn} &= O_{tn} + F_{an} = -G_{01} H_{01}, \\ F_{ts} &= O_{ts} + F_{as} = -G_{01} H_{03} \end{aligned} \tag{16}$$

where F_{tn} and F_{ts} are total MHT in the NH and SH mid-latitudes, respectively. Note that O_{tn} (O_{ts}) is actually obtained by subtracting F_{an} (F_{as}) from the total heat transport F_{tn} (F_{ts}), instead of being directly calculated using Eq. (4). We also see that under the equilibrium condition, the ocean boxes 4–6 can be represented by ocean box 1. The 6-box ocean is reduced to the 3-box ocean, just as Rooth's 3-box model.

Using the parameters in Table 1, the equilibrium climate is obtained and listed in Table 2, which are consistent with values in previous box model studies (Nakamura et al. 1994;

Marotzke and Stone 1995; Scott et al. 1999; Longworth et al. 2005; Yang et al. 2016), coupled climate model results (Yang et al. 2017) and observations (Rayner et al. 2003; Carton and Giese 2008). For example, the poleward surface temperature gradient is about 23 °C in the NH, which generates a northward AHT of about 3.7 PW. The ocean mass transport by the THC is 14.0 Sv, which causes a northward OHT of about 1.3 PW in the NH. These amounts of AHT and OHT are also in line with the observations (Trenberth and Caron 2001). The gradients of meridional temperature and salinity are independent of the initial values, and only depend on parameters of the system. This mean climate is obtained with a strong negative feedback in the tropics ($-B_2 = -1.7$) and weak positive feedback in the extratropics ($-B_1 = 0.6$ and $-B_3 = 0.5$). For a stable Earth climate, the regional positive climate feedback is very likely, as long as the overall Earth climate is negative. Previous studies have suggested a weak positive feedback in the extratropics (e.g., Soden et al. 2008; Jonko et al. 2010; Vial et al. 2013; Yang et al. 2017), In Secs. 5, we also identify a weak positive feedback in the extratropics based on the HadISST data. Note that the mean climate is sensitive to model parameters. Under the global uniform climate feedback ($-B_1 = -B_2 = -B_3 = -1.7$), the mean THC is about 12.9 Sv, corresponding to a northward OHT of about 1.24 PW. We want to stress that slight difference in the mean state with different climate feedback does not affect the conclusions of this work.

2.3 Analytic solution of the BJC

Assuming there is a perturbation in the system, the changes in the heat transport components are:

$$\begin{aligned} \Delta O_{tn} &= -G_{01}(-B_1 \Delta T_1 + \chi \Delta T_{sn}), \\ \Delta O_{ts} &= -G_{01}(-B_3 \Delta T_3 + \chi \Delta T_{ss}). \end{aligned} \tag{17}$$

Table 2 Properties of the reference mean climate based on the parameters in Table 1

Symbol	Physical meaning	Value
$T_1, T_2, T_3, T_4, T_5, T_6,$	Temperature for boxes 1–6	2.6, 25.6, 3.4, 2.6, 2.6, 2.6 °C
$S_1, S_2, S_3, S_4, S_5, S_6,$	Salinity for boxes 1–6	35.0, 35.6, 34.4, 35.0, 35.0, 35.0 psu
T_{sn}, T_{ss}	Meridional temperature contrast $T_2 - T_1, T_2 - T_3$	23.0, 22.2 °C
S_{sn}, S_{ss}	Meridional salinity contrast $S_2 - S_1, S_2 - S_3$	0.6, 1.2 psu
q	Meridional ocean mass transport by THC	14.0 Sv
F_{an}, F_{as}	Atmosphere heat transport in NH and SH	3.7, − 3.6 PW
O_{tn}, O_{ts}	Ocean heat transport by THC	1.3, 0.05 PW
F_{tn}, F_{ts}	Total meridional heat transport	5.0, − 3.6 PW
F_{wn}, F_{ws}	Atmosphere moisture transport	0.34, − 0.33 Sv

For mass and heat transports, the positive (negative) value represents northward (southward). The subscripts n and s represent the NH and SH, respectively. 1 Sv $= 10^6$ m³/s, and 1 PW $= 10^{15}$ W

$$\Delta F_{an} = \chi G_{01} \Delta T_{sn}, \quad \Delta F_{as} = \chi G_{01} \Delta T_{ss}. \tag{18}$$

The constraint of the global total energy conservation in the Earth's system (Eq. (10)) requires that the temperature changes among the three boxes follow the relationship:

$$B_1 \Delta T_1 + B_2 \Delta T_2 + B_3 \Delta T_3 = 0. \tag{19}$$

The local climate feedback B_i must be in a reasonable range to maintain the stability of the coupled box model. This requires an overall negative feedback at least satisfying:

$$-(B_1 + B_2 + B_3) < 0. \tag{20}$$

The BJC rate C_R, defined as the ratio of AHT change and OHT change, can be obtained using:

$$C_{Rn} \equiv \frac{\Delta F_{an}}{\Delta O_{tn}} = -\frac{(\Delta T_2 - \Delta T_1)}{\Delta T_2 - \left(1 + \frac{B_1}{\chi}\right)\Delta T_1}, \quad C_{Rs} \equiv \frac{\Delta F_{as}}{\Delta O_{ts}} = -\frac{(\Delta T_2 - \Delta T_3)}{\Delta T_2 - (1 + \frac{B_3}{\chi})\Delta T_3}. \tag{21}$$

Here, C_{Rn} and C_{Rs} are compensation rates for the NH and SH, respectively. The lower ocean plays no role in the equilibrium BJC. It is apparent that in the two-hemisphere box model, the BJC rate depends largely on relative temperature changes in the surface ocean. This is different from the BJC rate derived in the one-hemisphere box model (Yang et al. 2016), in which the BJC rate is independent of temperature changes of the system. In the one-hemisphere box model, $\frac{\Delta T_2}{\Delta T_1} = -\frac{B_1}{B_2}$ is a constant due to the constraint of $B_1 \Delta T_1 + B_2 \Delta T_2 = 0$, and $C_R = -\frac{1}{1+B_1 B_2/\chi(B_1+B_2)}$ depends only on the internal climate parameters B and χ. In the two-hemisphere box model, however, temperature changes among the three boxes become more complex as formulized in Eq. (19), and $\frac{\Delta T_2}{\Delta T_1}$ is no longer a constant.

In fact, the two-hemisphere box model can be readily reduced to the one-hemisphere box model if the SH box (or the NH box) and the tropical box are combined. In this case, Eq. (21) can be re-written as follows,

$$C_{Rn} = -\frac{1}{1 + B_1 B_s/\chi(B_1 + B_s)}, \quad C_{Rs} = -\frac{1}{1 + B_3 B_n/\chi(B_3 + B_n)}, \tag{22}$$

where

$$B_s \equiv B_2 + \frac{\Delta T_3}{\Delta T_2} B_3, \quad B_n \equiv B_2 + \frac{\Delta T_1}{\Delta T_2} B_1. \tag{23}$$

are the combined climate feedbacks for the tropical-NH boxes and tropical-SH boxes, respectively. Equation (19) is then reduced to $B_1 \Delta T_1 + B_s \Delta T_2 = 0$ or $B_3 \Delta T_3 + B_n \Delta T_2 = 0$. Mathematically, C_{Rn} and C_{Rs} have symmetric forms about the equator, which are similar to the formula derived for the one-hemisphere box model in Yang et al. (2016), that is,

$C_R = -\frac{1}{1+B_1 B_2/\chi(B_1+B_2)}$. The physics behind this is simple: if the two hemispheres are perfectly symmetric about the equator, the BJC situations in the NH and SH should also be identical to that in the one-hemisphere box model. We want to emphasize again that the simpler form of the BJC in the one-hemisphere box model is independent of the climate change and heat transport change themselves, and depends only on internal climate parameters B_i and χ, which can then be thought as an eigen mode of the coupled climate system, as pointed out in Liu et al. (2016) and Yang et al. (2016).

The BJC establishes a link between the change in AHT and that in OHT. This link is valid at the decadal and longer timescales (Zhao et al. 2016), because the conservation of global total energy is required as a necessary condition that

can easily fail for short timescale changes. The BJC implies how and to what extent the AHT responds to the OHT change, or more generally, how the atmospheric meridional motion responds to the ocean change. The analytical solution of BJC, Eqs. (21) or (22), provides a simple and practical approach to scale the overall response in the atmosphere, given the relative temperature changes of ocean and local climate feedback in different regions.

Equations (21) or (22) suggests that the AHT change can perfectly compensate the OHT change in the mid-latitudes if $B_1 \Delta T_1 = 0$ ($B_3 \Delta T_3 = 0$) for the NH (SH), that is,

$$C_{Rn} = -1, \quad C_{Rs} = -\frac{1}{1 + B_2 B_3/\chi(B_2 + B_3)}, \quad \text{if} \quad B_1 \Delta T_1 = 0. \tag{24}$$

$$C_{Rn} = -\frac{1}{1 + B_1 B_2/\chi(B_1 + B_2)}, \quad C_{Rs} = -1 \quad \text{if} \quad B_3 \Delta T_3 = 0. \tag{25}$$

The physical mechanism behind this situation is simple: for example, $B_1 \Delta T_1 = 0$ means that the TOA net heat flux in the NH extratropics (Box 1 in Fig. 1) does not change in response to certain perturbation, so that any ocean heat gain in the extratropics through the horizontal OHT has to be exported completely to the tropics via horizontal AHT, in order to maintain an equilibrium climate in the extratropics, that is, a perfect compensation occurs. In other words, due to the lack of energy gain (or loss) in the vertical direction at the TOA ($B_1 \Delta T_1 = 0$), the horizontal energy inflow and outflow have to be exactly balanced in the atmosphere and ocean. This was discussed in details in Yang et al. (2016). Note that both C_{Rn} and C_{Rs} can be -1, i.e., perfect compensation can occur simultaneously in both the NH and SH if $B_1 = B_3 = 0$. This situation requires $-B_2 < 0$, i.e., the whole

system needs an overall negative feedback for its stability, so that $\Delta T_2 = 0$ based on Eq. (19). Therefore, there is no net heat flux change anywhere at the TOA. Significant changes in both the ocean and atmosphere circulations can still occur ($\Delta T_1 \neq 0$, $\Delta T_3 \neq 0$).

Equations (24) and (25) represent the best scenario of the BJC. Schematic diagram showing this best scenario is given in Fig. 2. Figure 2a shows an anomalous equatorward OHT is perfectly compensated by an anomalous poleward AHT, which can result in a symmetric change with warming (or zero warming) in the tropics and cooling in the extratropics. This pattern can be found in observations (Fig. 5b). Figure 2b shows an anomalous southward OHT is perfectly compensated by an anomalous northward AHT in both hemispheres, which can result in a dipole change with cooling in the NH and warming in the tropics and SH. This pattern can also be found in observations (Fig. 5c). Note that the changes in AHT (OHT) in the NH and SH do not have to be the same in magnitude, so the temperature changes in difference regions can be very different. Also note that in this scenario, the asymmetric THC does not cause any difference in the BJC in the NH and SH, because the zero TOA flux change requires the opposite but the-same-magnitude changes in OHT and AHT, regardless of the direction and magnitude of the background circulation.

3 Validating the BJC in perturbation experiments

In this section, we validate the theoretical formulae of Eq. (21) using box model perturbation experiments. Perturbing freshwater in the system does not affect the global total energy budget of the coupled box model, as discussed in Yang et al. (2016). Therefore, the precondition for a valid BJC is satisfied. Two freshwater perturbation experiments are performed, in which a constant freshwater flux is hosed into the NH extratropics (box 1), that is, a negative salinity tendency ($h = -5 \times 10^{-10}$psu/s, corresponding to a 0.5-Sv freshwater flux in box 1) is added in the salinity tendency equation of S_1 (Eq. 2a). In Exp. 1, we consider a global uniform climate feedback ($-B_1 = -B_2 = -B_3 = -1.7$) for simplicity. Exp. 2 is the same as Exp. 1, except for a non-uniform climate feedback ($-B_1 = 0.6, -B_2 = -1.7, -B_3 = 0.5$) as listed in Table 1. Each experiment has its own control run that has no external forcing. Other parameters are the same in these experiments. By comparing Exp. 1 and Exp. 2, we will see explicitly how the climate feedback and regional temperature changes affect the BJC.

The equilibrium changes in the two experiments, with respect to their own control runs, are shown in Fig. 3. The BJC values calculated from $\frac{\Delta F_a}{\Delta O_t}$ directly and from Eq. (21) are shown in Fig. 3a. The direct calculation of BJC ($\frac{\Delta F_a}{\Delta O_t}$) is nearly equal to the theoretical value as expected. Figure 3b-c show detail changes in temperature and heat transport. Under the globally uniform negative climate feedback, the changes in Exp. 1 are simple and clear (Fig. 3b): in response to the 0.5-Sv freshwater hosing in the NH extratropical box, the THC there is weakened by about 12% (1.6 Sv), leading to a 11% (0.132 PW) reduction of the OHT, a cooling in the NH extratropics ($\Delta T_1 = -0.35\,°C$), a weak warming in the tropics ($\Delta T_2 = +0.13\,°C$), and an even weaker warming in the SH ($\Delta T_3 = +0.02\,°C$), i.e., a dipole change in the upper-ocean temperature. The northward AHT is enhanced by about 2% (0.077 PW) due to the increased poleward
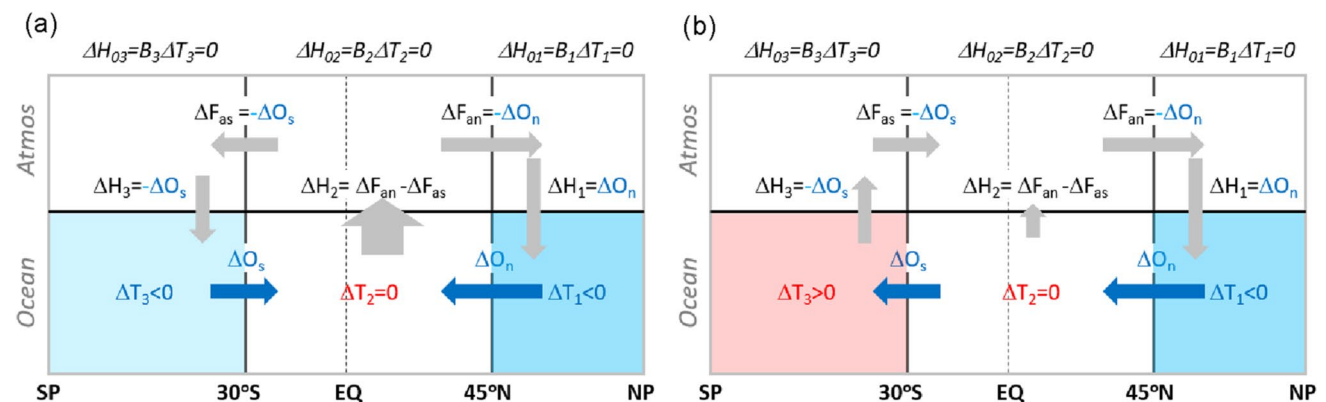

Fig. 2 Schematic diagram showing the perfect compensation, in which $B_1 = B_3 = 0$ so that Bjerknes compensation (BJC) rates in both hemispheres are equal to -1. Note that the changes in AHT and OHT in the Northern Hemisphere (NH) and Southern Hemisphere (SH) do not have to be of the same magnitude. **a** An anomalous equatorward OHT is perfectly compensated by an anomalous poleward AHT, which can result in a neutral change in the tropics and cooling in the extratropics of both hemispheres. **b** An anomalous southward OHT is compensated by an anomalous northward AHT in both hemispheres, which can result in a dipole change of cooling in the NH and warming in the SH
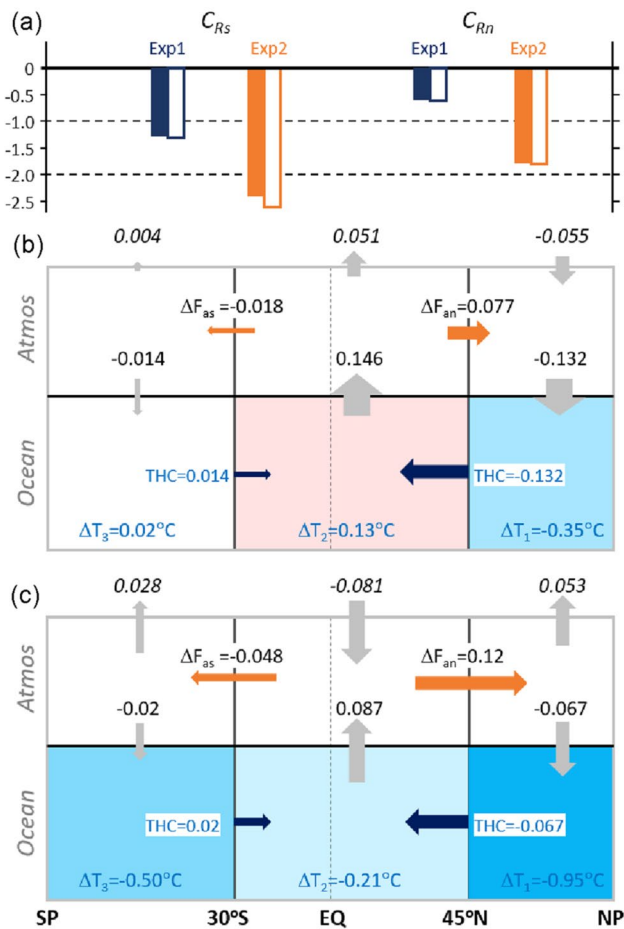
Fig. 3 Equilibrium response of climate changes to freshwater hosing in box 1. **(a)** Shows the BJC values in two experiments. Filled bars represent BJC from $\frac{\Delta F_a}{\Delta O}$, and unfilled bars, from Eq. (21). **b, c** Are for Exp.1 and Exp.2, respectively. In Exp.1, uniformly negative feedback ($-B_1 = -B_2 = -B_3 = -1.7$) is considered. In Exp. 2, weak positive feedbacks are given in the extratropical boxes ($-B_1 = 0.6, -B_3 = 0.5$) and a strong negative feedback is given in the tropics ($-B_2 = -1.7$). Orange arrow shows the AHT change, and dark blue arrows show the OHT changes due to THC. Positive (negative) value represents northward (southward) transport. Grey arrow shows the vertical heat transport change at the surface and top of the atmosphere, and positive (negative) value represents upward (downward) transport

temperature gradient $(\Delta(T_2 - T_1) = +0.48\,°C)$, which undercompensates the weakened OHT, because the negative climate feedback in the extratropics causes an additional heat gain (0.055 PW) from the TOA. The BJC rate in the NH is roughly -0.6 (Fig. 3a). The warming in the SH extratropics $(\Delta T_3 = +0.02\,°C)$ is much weaker than that in the tropics $(\Delta(T_2 - T_3) = +0.11\,°C)$. As a result, the AHT is enhanced southward $(-0.018$ PW). The OHT is increased northward (0.014 PW) because it is mainly determined by the temperature contrast between the two extratropical boxes $(\Delta(T_3 - T_1) = +0.37\,°C)$. The BJC in the SH is about -1.3 (Fig. 3a), indicating an overcompensation there.

Under non-uniform climate feedback, particularly considering the positive feedbacks in the two extratropical boxes, the temperature changes in Exp. 2 are much stronger than those in Exp. 1. The OHT change is weaker and the AHT change is stronger than those in Exp. 1 (Fig. 3c vs. Figure 3b). The THC in the NH extratropical box is weakened by about 9% (1.2 Sv), leading to a 5% (0.067 PW) reduction of the OHT, a strong cooling in the NH extratropics $(\Delta T_1 = -0.95\,°C;$ due to positive feedback in the extratropics) and a weak cooling in the tropics $(\Delta T_2 = -0.21\,°C)$. The northward AHT is enhanced by about 3% (0.12 PW) due to the increased poleward temperature gradient $(\Delta(T_2 - T_1) = +0.74\,°C)$, which overcompensates the weakened OHT, because the positive climate feedback in the extratropics causes an additional heat loss (0.053 PW) out of the TOA. The BJC rate in the NH is -1.79 (Fig. 3a). The cooling in the SH extratropics $(\Delta T_3 = -0.50\,°C)$ is also stronger than that in the tropics due to positive feedback $(\Delta(T_2 - T_3) = +0.29\,°C)$. As a result, the AHT is enhanced southward (-0.048 PW). The OHT in the SH is increased northward (0.02 PW) because of the enhanced temperature contrast between the two extratropical boxes $(\Delta(T_3 - T_1) = +0.45\,°C)$. The BJC in the SH is about -2.4 (Fig. 3a), indicating a strong overcompensation there.

Although mathematically C_{Rn} and C_{Rs} have the identical form as shown in Eq. (22) and they can also be identical in the situation of Fig. 2, they are very different as revealed in Exps. 1 and 2 under non-zero climate feedbacks. This can only be attributed to the asymmetric THC in the two hemispheres. Physically, the role of THC in C_{Rn} and C_{Rs} can be understood as follows. The asymmetric THC leads to a much stronger mean OHT in the NH than in the SH (1.28 PW vs. 0.05 PW; Table 2), because the vertical temperature difference in the NH is much bigger than that in the SH (23 °C vs. 0.8 °C; Table 2) (Eq. (4)). This background state implies that under certain small perturbation, there would be a much stronger OHT change in the NH than in the SH (Fig. 3b–c). For the atmosphere, the mean AHT is much stronger than the mean OHT. And the AHT changes in the NH and SH are usually comparable, because they are determined by the comparable *poleward* surface temperature gradient in the two hemispheres (Eq. (5)), which are also comparable to the OHT change in the NH. Therefore, the relative changes in the AHT and OHT in the NH are comparable, that is, a reasonable BJC in the NH can be expected. In the SH, the AHT change tends to be stronger than the OHT change, which would always result in an overcompensation regardless of the sign of climate feedback (Fig. 3b). The overcompensation can be exacerbated if there is a positive climate feedback in the SH (Fig. 3c), because the positive feedback affects the atmosphere more seriously than affecting the ocean.

The BJC in the NH (C_{Rn}) is more predictable, consistent with the suggestion by the theoretical formula $C_R = -\frac{1}{1+B_1B_2/\chi(B_1+B_2)}$ derived in the one-hemisphere box
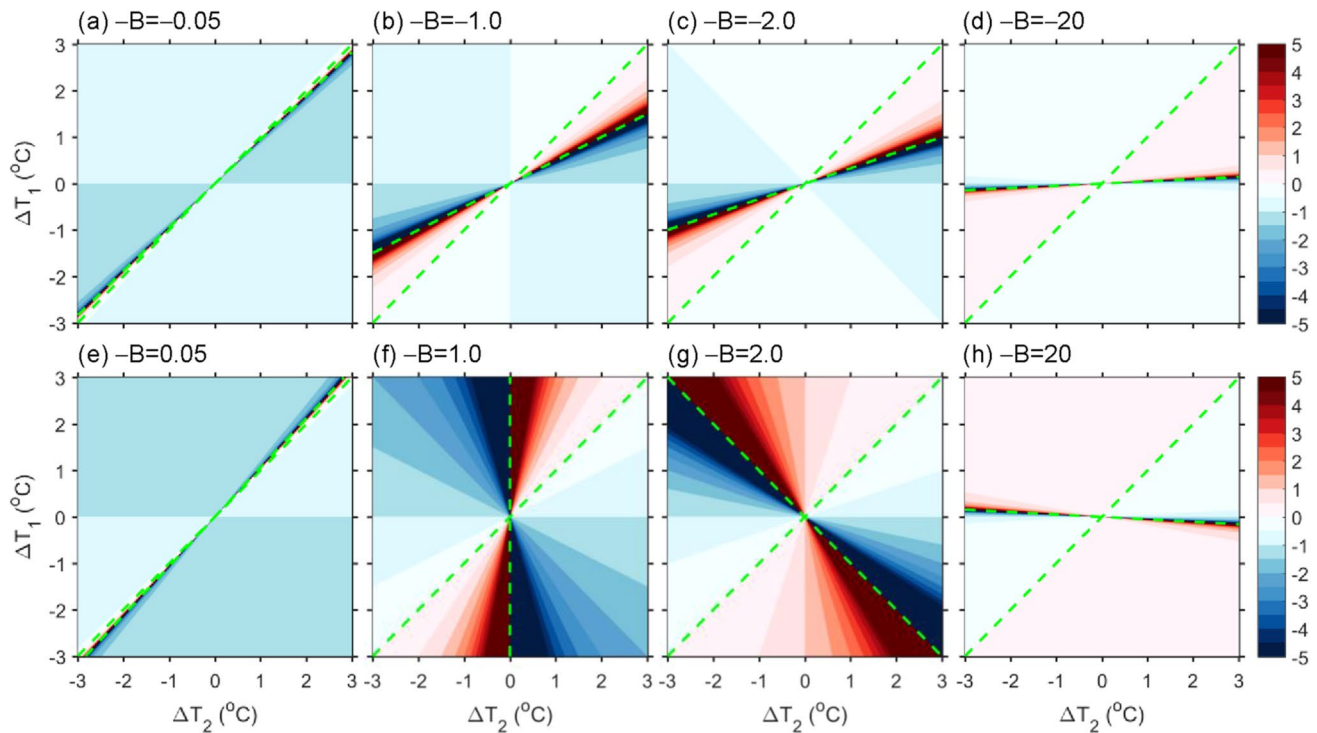
Fig. 4 Pattern of BJC rate in the phase space of temperature change based on Eq. (21). The x-axis represents the temperature change in the tropical box ΔT_2, and the y-axis represents the temperature change in the extratropical box (ΔT_1 or ΔT_3). The shaded contours show C_{Rn} or C_{Rs}, which are actually determined by the ratio of temperature changes in different boxes, i.e., $\Delta T_1/\Delta T_2$ or $\Delta T_3/\Delta T_2$; therefore, the contours consist of a cluster of straight lines that all avoid the singular point of (0, 0). The smaller area between two dashed green lines is the domain with $C_R > 0$, i.e., the domain with no com-

pensation. The upper (lower) panels are for C_R under the negative (positive) climate feedback. **a–d** Are for –B = −0.05, −1.0, −2.0, and −20, respectively; **e–h**, for –B = 0.05, 1.0, 2.0, and 20, respectively. Here, B = B_1/χ or B_3/χ is a non-dimensional parameter showing the relative strength of climate feedback with respect to the atmospheric heat transport coefficient. Note that B = ±0.05 represent the situations with nearly no climate feedback, and B = ±20 represent the situations with infinitely strong negative and positive feedbacks

model (Yang et al. 2016); that is, there should be undercompensation ($|C_{Rn}| < 1$) for global negative feedback, or overcompensation ($|C_{Rn}| > 1$) if there is a positive climate feedback somewhere, as validated in Exp.1 and Exp. 2 (Fig. 3a). However, the BJC in the SH (C_{Rs}) cannot be estimated merely based on the climate feedback. The changes of temperature patterns are critical to detailed BJC values in a global coupled system. However, if we only concern whether or not the BJC would occur in a global coupled system, or the probability of a valid BJC, the temperature change patterns do not matter anymore.

4 Probability of a valid BJC

To know how much the BJC in general depends on temperature changes of different regions, as well as the climate feedback, we plot C_{Rn} and C_{Rs} using Eq. (21) in Fig. 4. Note that C_{Rn} and C_{Rs} are actually determined by the ratio of temperature changes ($\frac{\Delta T_2}{\Delta T_1}$ and $\frac{\Delta T_2}{\Delta T_3}$). The contours of C_{Rn} and C_{Rs} in

Fig. 4 consist of a cluster of straight lines that avoid the singular point of (0, 0), in which $\frac{\Delta T_2}{\Delta T_1}$ (or $\frac{\Delta T_2}{\Delta T_3}$) is constant along each line. Since mathematically C_{Rn} and C_{Rs} have the identical form, only one of them needs to be plotted. The dashed green lines in Fig. 4 represent the situations of $\frac{\Delta T_2}{\Delta T_1} = 1$ and $\frac{\Delta T_2}{\Delta T_1} = (1 + \frac{B_1}{\chi})$, and the area enclosed by these two lines shows the regime where the BJC fails, i.e., $C_{Rn} > 0$ (denoted by warm colors in Fig. 4). In fact, $C_{Rn} = 0$ when $\frac{\Delta T_2}{\Delta T_1} = 1$, and $C_{Rn} = \pm\infty$ when $\frac{\Delta T_2}{\Delta T_1} = (1 + \frac{B_1}{\chi})$. Therefore, in the phase space of temperature change, theoretically the probability that the BJC fails can be defined as the ratio of the area within the green lines to the total area of the square, that is,

$$p = \begin{cases} (1-k)/4, & -B \le 0 \\ (1-\frac{1}{k})/4, & 0 \le -B \le 2 \\ (3+k)/4, & -B \ge 2 \end{cases}, \quad (26)$$

where $k = 1/(1 + B)$ is the slope. Here, we define a non-dimensional parameter, $B = \frac{B_1}{\chi}$, showing the relative strength of
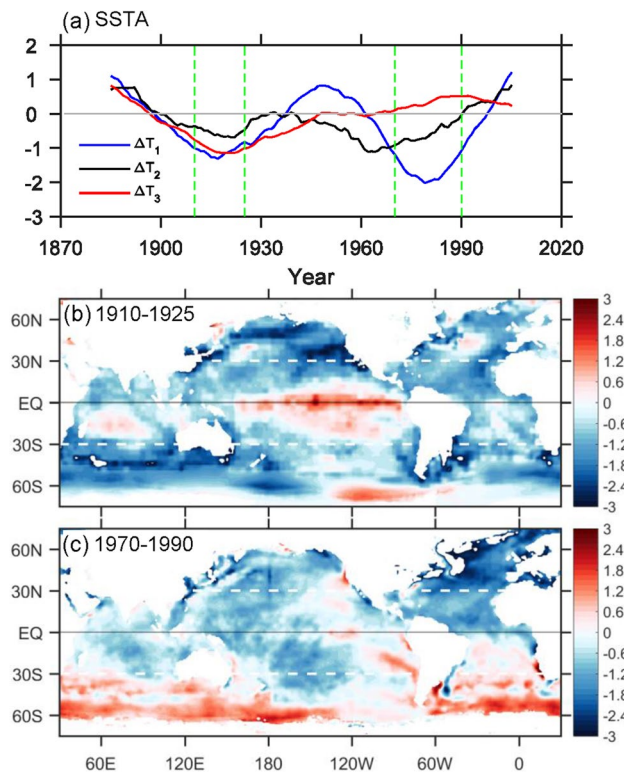
Fig. 5 Time series and patterns of sea surface temperature anomaly (SSTA; units: 0.1 °C) using the HadISST data. **a** SSTA time series averaged over 30°N–90°N (blue curve), 30°S–30°N (black curve) and 90°S–30°S (red curve). The dashed vertical green lines indicate two periods of 1910–1925 and 1970–1990. **b, c** are SSTA patterns averaged over the periods of 1910–1925 and 1970–1990, respectively, indicated by the two pairs of dashed green lines in (**a**). Annual mean data are used here. The long-term trend over the period 1870–2020 is first removed, followed by a 30-year running mean

of the local climate feedback (B_1) with respect to the meridional atmospheric transport coefficient (χ) in the same latitude band. Therefore, the probability of a valid BJC is $(1 - p)*100\%$, which is then determined only by the internal climate parameters and is *independent* of temperature change.

Figure 4 shows that generally, under a reasonable climate feedback, the probability of a valid BJC is more than 80% (blue region). Under an extremely negative climate feedback ($-B \to -\infty$), the probability of BJC is no less than 75% (Fig. 4d), no matter how the temperature changes, while under an extremely positive climate feedback ($-B \to \infty$), the probability of a valid BJC can be as low as 25% (Fig. 4h). Figure 4d, h represent two extreme situations, which are unrealistic but provide us useful information for understanding the BJC limit. A strong local negative feedback can efficiently dissipate local heat gain (or loss) in the vertical through the TOA, so that a more freedom of AHT change is allowed in response to OHT change, which could reduce the

probability of BJC by as much as 25% (Fig. 4d). A strong local positive feedback, instead, can exacerbate seriously the heat imbalance through the vertical process, so that the AHT and OHT have to change cooperatively (Liu et al. 2018), in order to maintain the local energy balance, resulting in probability of BJC failure being as high as 75% (Fig. 4h). Under a neutral climate feedback ($-B \to 0$), the BJC would be valid for nearly all temperature changes (Fig. 4a, e), with the probability of a perfect compensation (i.e., $C_R = -1$) nearly 100%. In a more realistic situation with the local negative climate feedback close to the atmosphere heat transfer coefficient ($-B = -1 \Rightarrow k = \frac{1}{2} \Rightarrow p = \frac{1}{8}$), the probability of a valid BJC is 87.5% (Fig. 4b). In contrast, under a realistic local positive climate feedback ($-B = 1 \Rightarrow k \to \infty \Rightarrow p = \frac{1}{4}$), the probability of a valid BJC is 75% (Fig. 4f). With enhanced local climate feedback, the probability of a valid BJC decreases, which is about 83% for $-B = -2$ (Fig. 5c) and 50% for $-B = 2$ (Fig. 4g). Figure 4 also shows that the probability of a valid BJC is much higher under negative feedback than under positive feedback.

In general, in a coupled box model system under the constraint of global energy conservation, the compensation changes in AHT and OHT are rather robust, regardless of the relative temperature changes in different latitude zones. The critical factor in determining the BJC rate is still the local climate feedback, as emphasized in our previous studies (Liu et al. 2016; Yang et al. 2016). How robust the BJC would be in a real world depends on what the climate feedback is in the reality. We will discuss this briefly next.

5 Evaluating the BJC in the real world

To evaluate the robustness of the BJC in the real world, we use the sea surface temperature (SST) data set of HadISST (Rayner et al. 2003) from the UK Met Office Hadley Centre to calculate the climate feedback and BJC. It is monthly data with a horizontal resolution of 1°·1°, and spans the period of 1870–2020. We use the annual mean data in our calculation. The long-term trend over 1870–2020 is removed, and then a low-pass filter of 30-year running mean is applied. BJC and global energy conservation are usually discussed on decadal and longer timescales, thus the de-trended low-frequency data is suitable.

Figure 5 shows the evolution of SST anomaly (SSTA) averaged over different latitude bands during 1870–2020, and the SSTA patterns averaged over two periods of 1910–1925 and 1970–1990. The 60–80 years' multi-decadal variation is clearly seen in the SSTA at all latitudes (Fig. 5a). The SSTA patterns over the two negative phases of the multi-decadal variation are quite different. For the period of 1910–1925 (Fig. 5b), the Atlantic SSTA had a roughly uniform cooling, while the Pacific SSTA showed a tri-polar

structure, with a warming in the central-eastern tropical Pacific and cooling in the extratropics. For the period of 1970–1990 (Fig. 5c), the SSTA in most region of the Pacific showed a cooling, while the Atlantic SSTA showed a dipole structure, with cooling in the NH and warming in the SH and Southern Ocean. Note that in the early period of the twentieth century, the ocean data coverage was very sparse (Deser et al. 2010). The SSTA pattern shown in Fig. 5b may not be accurate. Notwithstanding, the SSTA patterns in Fig. 5b and c are qualitatively consistent with the schematic diagrams in Fig. 2a and b, respectively. The SSTA evolution, patterns and their mechanisms have been studied comprehensively (Deser et al. 2010), and are not the focus of this paper. The purpose of Fig. 5 is to provide a general picture on the observational data used in this study.

To estimate the BJC in the reality, the climate feedback needs to be determined first. Based on the global energy conservation of Eq. (19), we can determine the climate feedback in different latitude bands. More specifically, we can obtain the relative magnitude of climate feedback in different regions. In fact, Eq. (19) can be generalized as follows,

$$\sum_{i=1}^{n} B_i \Delta T_i = 0. \tag{27}$$

Equation (27) forms homogeneous linear equations for B_i, provided that ΔT_i is given based on observations. Unfortunately, there is no non-zero solution to B_i in these homogeneous linear equations. However, if we happen to know one B_j in a specific latitude band, the B_i in the other regions can be easily obtained. Equation (27) can be re-written as follows,

$$\sum_{i=1,i\neq j}^{n} \frac{B_i}{B_j} \Delta T_i = -\Delta T_j. \tag{28}$$

Equation (28) forms non-homogeneous linear equations, and there are non-zero solutions to B_i. If B_j is known, B_i can be determined exactly. If B_j is unknown, the ratio $\frac{B_i}{B_j}$ can be determined at least. Equation (28) is simple, clean and clear in physical mechanism, but is powerful in determining the climate feedback parameters in every latitude band, as long as adequate surface temperature data are available.

A multivariable linear regression model is used to calculate coefficients $\frac{B_i}{B_j}$. Based on the time series shown in Fig. 5a, we obtain $\frac{B_1}{B_2} = -0.35$ and $\frac{B_3}{B_2} = -0.27$. Here, all of the 150-year data are used for the calculation. $\frac{B_1}{B_2}$ and $\frac{B_3}{B_2}$ could vary slightly by about ±10% if different lengths of data are used; and these values suggest that the sign of the tropical climate feedback (B_2) tends to be opposite to those of the extratropics (B_1, B_3). Note that in Fig. 5, ΔT_1, ΔT_2 and ΔT_3 are area-averaged quantities based on the regions slightly different from those defined in the box model (Fig. 1), so the

area-weighted quantities B_1, B_2 and B_3 here are also region-dependent. We want to emphasize that the fundamental principles, i.e., all formulae in this work, are independent of how the regions are defined.

In this work, we choose $-B_2 = -1.7$, so that $-B_1 = 0.6$ and $-B_3 = 0.5$. Previous studies suggested a strong negative feedback in the tropics and weak positive feedback in the extratropics (e.g., Soden et al. 2008; Jonko et al. 2010; Vial et al. 2013; Yang et al. 2017). The overall global climate feedback is negative, ensuring the stability of the current climate. The negative tropical feedback is mainly due to the strong negative feedback between the outgoing longwave radiation (OLR) and surface temperature (in association with high clouds), which dominates the positive feedback between shortwave radiation and surface temperature (Jonko et al. 2010; Vial et al. 2013; Yang et al. 2017). The positive feedback in the extratropics is mainly due to the strong positive feedback between the shortwave radiation and surface temperature, which overcomes the strong negative feedback between the OLR and surface temperature (Soden et al. 2008; Vial et al. 2013; Yang et al. 2017). The feedback parameter B_2 we choose here and B_1 and B_3 estimated based on the HadISST using Eq. (28) are consistent with those estimated using the so-called radiative kernel technique (Soden et al. 2008; Jonko et al. 2010).

The BJC situation based on the HadISST is shown in Fig. 6. Using Eq. (26), the calculated theoretical probabilities of a valid BJC under the weak positive feedback $-B_1 = 0.6$ and $-B_3 = 0.5$ are 91 and 93%, respectively (Fig. 6a, b). The temperature anomaly pairs of (ΔT_1, ΔT_2) and (ΔT_2, ΔT_3) are also scattered in Fig. 6a and b, respectively. We see that most of the circles are in the regime of a valid BJC. Figure 6c shows distributions of C_{Rn} calculated using ΔT_1, ΔT_2 and B_1 (blue bars) and C_{Rs} calculated using ΔT_3, ΔT_2 and B_3 (cyan bars). The occurrence of a valid BJC based on HadISST is greater than 90% for both hemispheres. The mean C_{Rn} (C_{Rs}) for the NH (SH) is -1.50 (-1.30). The "good" compensation ($C_R \in (-0.5, -1.5)$) occurs more than 50% for both hemispheres. Using the theoretical formulae Eq. (21), the estimated BJC based on the HadISST data suggests the robustness of compensation changes in AHT and OHT in the reality.

6 Summary and discussion

Using a coupled two-hemisphere model, we investigate the BJC in the presence of an interhemispheric thermohaline circulation. First, we obtain an analytical solution to the BJC, which is determined by both local climate feedback and temperature change. This is different from the BJC in a one-hemisphere model, which only considers local climate
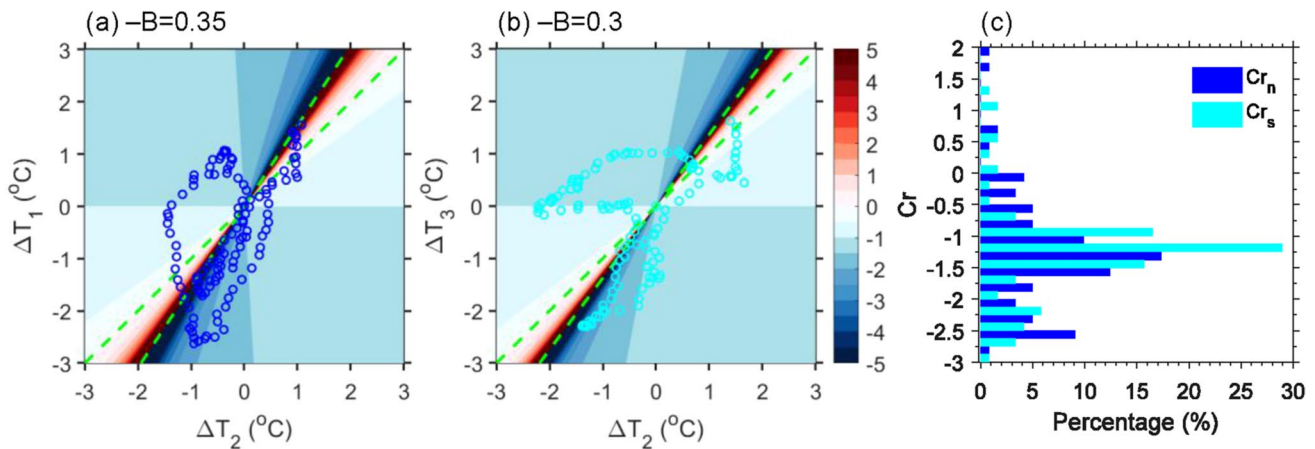
Fig. 6 Pattern of BJC rate in the phase space of temperature change and BJC distribution based on annual mean lowpass-filtered HadISST data. **a**, **b** Are the same as those in Fig. 4, except that the climate feedback values B_1 and B_3 are obtained from the HadISST data. The units for temperature anomaly are 0.1 °C. In (**a**), $-B = B_1/\chi = 0.35$, where $-B_1 = 0.6$. Blue open circles represent ΔT_1 vs. ΔT_2. In (**b**), $-B = B_3/\chi = 0.3$, where $-B_3 = 0.5$. Cyan open circles represent ΔT_3 vs. ΔT_2. Here $\chi = 1.7$. **c** Is the BJC distributions for C_{Rn} (blue bars) and C_{Rs} (cyan bars). The x-axis denotes the probability in percentage (%), and the y-axis shows the value of C_{Rn} or C_{Rs}

feedback. Second, we derive a formula for the probability of a valid BJC, i.e., the possibility for a negative BJC ($C_{R<0}$). We illustrate that the probability of a valid BJC depends only on the local climate feedback and is independent of temperature change. The probability of a valid BJC is usually higher than 80% under reasonable choice of climate feedback parameters. Third, the BJC and the probability of a valid BJC are evaluated using observational data of the HadISST; and both are found to be robust. The BJC in box model suggests a better BJC in the NH than in the SH (Fig. 3a) because of the asymmetric interhemispheric THC, while the observational data suggests that the asymmetric THC appears to have no effect on the BJC (Fig. 6c). This is an interesting question to be explored in our next work.

The progresses of this work with respect to previous studies (e.g., Marotzke 1990; Nakamura et al. 1994; Tziperman et al. 1994; Marotzke and Stone 1995; Liu et al. 2016, 2018; Yang et al. 2016) are the two-hemisphere box model used and the formula for the probability of a valid BJC. This model is one step closer to the reality. For example, in the one-hemisphere box model, the BJC is independent of the heat transports themselves, which fail to resemble the full range of behaviors suggested by complex general circulation models, as pointed out by Rose and Ferreira (2013). The two-hemisphere box model proposed by Rooth (1982) is dynamically superior to the one-hemisphere model (Longworth et al. 2005), by considering an interhemispheric THC. In the two-hemisphere box model, the BJC does depend on the pattern of temperature change. However, under reasonable choice of climate feedback parameters, the relative temperature changes in different regions do not affect the BJC too much, as suggested by the probability of a valid BJC. The fundamental mechanism revealed in the one-hemisphere box model remains valid in a global system to some extent.

The BJC is one of the fundamental mechanisms that constrain the global climate change. This mechanism may be crucial to the overall Earth's climate stability, and may shed light on a potential self-restoring mechanism in a complex climate system. The BJC also suggests a remote climate change in response to a local forcing, such like the teleconnection between the SH ocean-atmosphere system and the polar amplification in the NH (Liu et al. 2018). The present-day's Earth climate is experiencing a rising global mean temperature and a diminishing cryosphere in the high latitudes and over mountains. Knowing the BJC in the real world has a realistic significance for us to assess future climate change and the possible changing teleconnections between the two hemispheres in the future.

The coupled box model has many limitations. The linear relationship between the AHT and poleward surface temperature gradient is appropriate for the atmosphere in the mid-to-high latitudes (Stone and Yao 1990), but not accurate in the low latitudes. The ocean model is constructed based on Rooth's (1982) box model and considers an interhemispheric THC, but it does not consider the effects of wind forcing and vertical mixing in the Southern Ocean on the THC (e.g., Toggweiler and Samuels 1995, 1998). Moreover, the wind-driven circulation (WDC) is not included in the ocean model. In reality, the WDC has roughly an symmetric structure about the equator, which transports heat poleward in collaboration with the AHT. The southward OHT by the WDC is important in the SH, which dominates over the northward OHT by the THC and leads to roughly symmetric poleward OHTs by global oceans (Trenberth and Caron 2001). The role of WDC OHT in the global total meridional heat transport has been examined through many different numerical models (e.g., Vallis and Farneti 2009).

However, how the WDC would affect the BJC theoretically in a box model remains to be explored in-depth. Will the WDC lead to a symmetric BJC in the two hemispheres, or cause a failure of the BJC in the SH? Actually, Fig. 6 suggests that the WDC might not be critical to the BJC of the global coupled system, since the HadISST has included the effect of the WDC. Still, a theoretical study on WDC's role in the BJC is needed.

References

Bjerknes J (1964) Atlantic air-sea interaction. Adv Geophys 10:1–82

Budyko MI (1969) The effect of solar radiation variations on the climate of the earth. Tellus 21:611–619. https://doi.org/10.3402/tellusa.v21i5.10109

Carton JA, Giese BS (2008) A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). Monthly Weather Rev 136:2999–3017. https://doi.org/10.1175/2007MWR1978.1

Czaja A, Marshall J (2006) The partitioning of poleward heat transport between the atmosphere and ocean. J Atmos Sci 63:1498–1511. https://doi.org/10.1175/JAS3695.1

Deser C, Alexander MA, Xie S-P, Phillips AS (2010) Sea surface temperature variability: patterns and mechanisms. Annu Rev Mar Sci 2:115–143. https://doi.org/10.1146/annurev-marine-120408-151453

Donohoe A, Marshall J, Ferreira D, Mcgee D (2013) The relationship between ITCZ location and cross-equatorial atmospheric heat transport: from the seasonal cycle to the Last Glacial Maximum. J Clim 26:3597–3618. https://doi.org/10.1175/JCLI-D-12-00467.1

Enderton D, Marshall J (2009) Explorations of atmosphere-ocean-ice climates on an aquaplanet and their meridional energy transports. J Atmos Sci 66:1593–1611. https://doi.org/10.1175/2008JAS2680.1

Farneti R, Vallis G (2013) Meridional energy transport in the coupled atmosphere-ocean system: compensation and partitioning. J Clim 26:7151–7166. https://doi.org/10.1175/JCLI-D-12-00133.1

Frierson DM, Huang Y-T (2012) Extratropical Influence on ITCZ shifts in slab ocean simulations of global warming. J Clim 25:720–733. https://doi.org/10.1175/JCLI-D-11-00116.1

Held IM (2001) The partitioning of the poleward energy transport between the tropical ocean and atmosphere. J Atmos Sci 58:943–948

Huang RX, Luyten JR, Stommel HM (1992) Multiple equilibrium states in combined thermal and saline circulation. J Phys Oceanogr 22:231–246

Jonko AK, Hense A, Feddema JJ (2010) Effects of land cover change on the tropical circulation in a GCM. Clim Dyn 35:635–649. https://doi.org/10.1007/s00382-009-0684-7

Kang SM, Held IM, Frierson DMW, Zhao M (2008) The response of the ITCZ to extratropical thermal forcing: Idealized slab-ocean experiments with a GCM. J Clim 21:3521–3532. https://doi.org/10.1175/2007JCLI2146.1

Kang SM, Frierson DMW, Held IM (2009) The tropical response to extratropical thermal forcing in an idealized GCM: The importance of radiative feedbacks and convective parameterization. J Atmos Sci 66:2812–2827. https://doi.org/10.1175/2009JAS2924.1

Langen PL, Alexeev VA (2007) Polar amplification as a preferred response in an idealized aquaplanet GCM. Clim Dyn 29:305–317. https://doi.org/10.1007/s00382-006-0221-x

Lindzen RS, Farrell B (1977) Some realistic modifications of simple climate models. J Atmos Sci 34:1487–1501

Liu Z, Yang H, He C, Zhao Y (2016) A theory for Bjerknes compensation: the role of climate feedback. J Clim 29:191–208. https://doi.org/10.1175/JCLI-D-15-0227.1

Liu Z, He C, Lu F (2018) Local and remote responses of atmospheric and oceanic heat transports to climate forcing: compensation versus collaboration. J Clim 31:6445–6460. https://doi.org/10.1175/JCLI-D-17-0675.1

Longworth H, Marotzke J, Stocker TF (2005) Ocean gyres and abrupt change in the thermohaline circulation: a conceptual analysis. J Clim 18:2403–2416

Marotzke J (1990) Instabilities and multiple equilibria of the thermohaline circulation. Christ-Albrechts-Univers. https://doi.org/10.1016/S0032-3861(98)00713-7

Marotzke J, Stone P (1995) Atmospheric transports, the thermohaline circulation, and flux adjustments in a simple coupled model. J Phys Oceanogr 25:1350–1364

Nakamura M, Stone PH, Marotzke J (1994) Destabilization of the thermohaline circulation by atmospheric eddy transports. J Clim 7:1870–1882

North GR (1984) The small ice cap instability in diffusive climate models. J Atmos Sci 41:3390–3395

Rahmstorf S (1996) On the freshwater forcing and transport of the Atlantic thermohaline circulation. Clim Dyn 12:799–811

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J Geophys. https://doi.org/10.1029/2002JD002670

Rooth C (1982) Hydrology and ocean circulation. Progr Oceanogr 11:131–149

Rose BE, Ferreira D (2013) Ocean heat transport and water vapor greenhouse in a warm equable climate: a new look at the low gradient paradox. J Clim 26:2117–2136. https://doi.org/10.1175/JCLI-D-11-00547.1

Shaffrey L, Sutton R (2006) Bjerknes compensation and the decadal variability of the energy transports in a coupled climate model. J Clim 19:1167–1181. https://doi.org/10.1175/JCLI3652.1

Scott JR, Marotzke J, Stone P (1999) Interhemispheric thermohaline circulation in a coupled box model. J Phys Oceanogr 29:351–365

Soden BJ, Held IM, Colman R, Shell KM, Kiehl JT, Shields CA (2008) Quantifying climate feedbacks using radiative kernels. J Clim 21:3504–3520. https://doi.org/10.1175/2007JCLI2110.1

Stommel H (1961) Thermohaline convection with two stable regimes of flow. Tellus 13:224–230. https://doi.org/10.1111/j.2153-3490.1961.tb00079.x

Stone PH (1978) Constraints on dynamical transports of energy on a spherical planet. Dyn Atmos Oceans 2:123–139

Stone PH, Yao M-S (1990) Development of a two-dimensional zonally averaged statistical-dynamical model. Part III: the parameterization of the eddy fluxes of heat and moisture. J Clim 3:726–740

Swaluw EVD, Drijfhout SS, Hazeleger W (2007) Bjerknes compensation at high northern latitudes: the ocean forcing the atmosphere. J Clim 20:6023–6032. https://doi.org/10.1175/2007JCLI1562.1

Toggweiler JR, Samuels B (1995) Effect of Drake Passage on the global thermohaline circulation. Deep-Sea Res Part I: Oceanogr Res Pap 42:477–500. https://doi.org/10.1016/0967-0637(95)00012-U

Toggweiler JR, Samuels B (1998) On the ocean's large-scale circulation near the limit of no vertical mixing. J Phys Oceanogr 28:1832–1852

Trenberth KE, Caron JM (2001) Estimates of meridional atmosphere and ocean heat transports. J Clim 14:3433–3443

Tziperman E, Toggweiler JR, Feliks Y, Bryan K (1994) Instability of the thermohaline circulation with respect to mixed boundary conditions: Is it really a problem for realistic models. J Phys Oceanogr 24:217–232

Vallis GK, Farneti R (2009) Meridional energy transport in the coupled atmosphere–ocean system: Scaling and numerical experiments. Quart J Roy Meteor Soc 135:1643–1660. https://doi.org/10.1175/JCLI-D-12-00133.1

Vellinga M, Wu P (2008) Relations between northward ocean and atmosphere energy transports in a coupled climate model. J Clim 21:561–575. https://doi.org/10.1175/2007JCLI1754.1

Vial J, Dufresne J-L, Bony S (2013) On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. Clim Dyn 41:3339–3362. https://doi.org/10.1007/s00382-013-1725-9

Yang H, Zhao Y, Liu Z (2016) Understanding Bjerknes compensation in atmosphere and ocean heat transports using a coupled box model. J Clim 29:2145–2160. https://doi.org/10.1175/JCLI-D-15-0281.1

Yang H, Wen Q, Yao J, Wang Y (2017) Bjerknes compensation in meridional heat transport under freshwater forcing and the role of climate feedback. J Clim 30:5167–5185. https://doi.org/10.1175/JCLI-D-16-0824.1

Zhang R, Delworth T (2005) Simulated tropical response to a substantial weakening of the Atlantic thermohaline circulation. J Clim 18:1853–1860. https://doi.org/10.1175/JCLI3460.1

Zhao Y, Yang H, Liu Z (2016) Assessing Bjerknes compensation for climate variability and its time-scale dependence. J Clim 29:5501–5512. https://doi.org/10.1175/JCLI-D-15-0883.1