View from the Pennines: Box Models of the Oceanic Conveyor Belt

Paul Glendinning, School of Mathematics, University of Manchester, Manchester M13 9PL

In 1964 the snow here was so deep that our next-door neighbours were snowed in for six weeks. I don't mean that their drive was too icy to negotiate, but that they had over six feet of snow all along the 50 yard track down to their farm. They are resourceful people and they coped. It was a notable experience and a subject of conversation for years to come, but only because of the length of time the snow lay on the ground; the fact of being snowed in was sufficiently common that this on its own was not worthy of comment. They have not been snowed in for more than a couple of days at a time in the last fifteen years (and then only in the sense that the drive was unpassable to cars).

Of course, this is only anecdotal evidence for climate change, but most scientists working on the problem agree that something fairly drastic is happening. The world's weather is notoriously hard to predict, and contains many complicated feedback loops so it may well be that one effect of global warming will be local cooling. For this reason a lot of attention has been paid to global mixing mechanisms and whether these might be disrupted.

One such mechanism, the oceanic conveyor belt (or thermohaline circulation) appears to be responsible for a large amount of heat transport from equatorial (or low latitude) to polar (or higher latitude) regions. The idea is that there is a loop of currents taking warm water from the equator to the poles near the surface. At the poles the surface water cools and sinks (as cooler water is denser), and is then transported back to the tropics as deep-water currents. To complete the cycle the salt concentration needs to be considered as well: evaporation at equatorial regions increases the salinity and hence density of the surface water, allowing an exchange of heat to deeper parts of the ocean and thus allowing the waters to rise [4], hence the name thermohaline circulation. This is the theoretical mechanism, observations show that just such a conveyor belt operates in the Atlantic Ocean, and actually connects to the warmer waters of the Indian and Pacific Oceans.

Modelling the physical and dynamic processes at work here is a real challenge, and there are two different extremes in approach. One is to create a massive model with almost all the possible effects included as accurately as possible, with each component validated as carefully as possible. The resulting grand model can then be run on a sufficiently large computer or set of computers, and the extraordinarily large data set generated can then be sifted for relevant information. This is hard, expensive, requires large teams, a lot of money, and at the end it is not completely clear how much of the information gained is really accurate, although it does provide the best guess of what might happen, and parts can be validated by comparison with data. At the other extreme one can build simplified models which are cheap, inaccurate, but contain enough of a germ of the truth to at least enhance our intuition about what mechanisms might be important. The end results are, of course, open to all sorts of objections but the process allows scientists to test their expectations.

In practice, both approaches are needed, as the intuition gained from the simpler models informs the questions asked of the complicated models. These extremes are the two poles of a spectrum of modelling; the level of resolution or physical complexity contained in a model is up to the modeller and determined by the type of question the modeller is attempting to answer.

There is evidence that in the distant past the Atlantic conveyor belt has been disrupted, and that associated with such disruption (or even reversal) there is a significant change in the distribution of temperature over the earth. It is therefore worth looking at the mechanism for these reversals, together with the question as to whether climate change could induce such a transition. This is an active area of research, and at the level of the simplest models possible, has led back to the 'box' models of circulation developed by Stommel fifty years ago (see [5], which has over five hundred citations on Google Scholar). In these models the polar region is described by one well-mixed box with temperature T_p and salinity S_p, whilst the equatorial region is represented by another well-mixed box with with temperature T_e and salinity S_e as shown in Figure 1. These boxes are connected near the surface and at depth by tubes which allow a flux q to flow between the boxes which depends upon the temperature and salinity differences between the boxes. The property which makes these models so compelling to theoreticians is that they predict that in some regimes there can be bistability between solutions representing the current conveyor belt and a weaker or reversed circulation. At more extreme parameter values the existing solution disappears and there is an abrupt transition to a weaker circulation (see Figure 3). It is this bistability which will be investigated here following Dijkstra [3].

Stommel [5] made the ansatz that the flux is dominated by the density difference between the regions,

$$q = \gamma\left(\frac{\rho_p - \rho_e}{\rho_0}\right)$$

where ρ_p and ρ_e are the polar and equatorial densities of the water, given in terms of reference temperatures, salinity and density T_0, S_0 and ρ_0 by

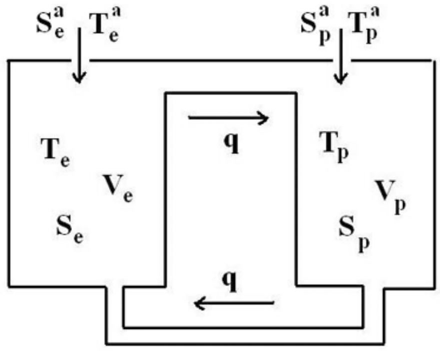

Figure 1: Schematic diagram of the box model. Variables with subscripts e are in the low latitude (equatorial) region and variables with subscripts p are in the higher latitude (polar) region (after Dijkstra [3]).

$$\rho = \rho_0(1 - \alpha_T(T - T_0) + \alpha_S(S - S_0))$$

Thus

$$q = \gamma(\alpha_T(T_e - T_p) - \alpha_S(S_e - S_p)) \tag{1}$$

Following Stommel as described by Dijkstra [3], assume that heat is added to the polar (resp. equatorial) box at a rate T_p^a (resp. T_e^a) from the atmosphere, with $T_e^a - T_p^a > 0$, and that salinity is increased at the equator (evaporation) and decreased at the pole (precipitation) at rates S_e^a and S_p^a with $S_e^a - S_p^a > 0$. Then the model is defined by four differential equations:

$$\begin{aligned}
V_e \frac{d}{dt} T_e &= C_e^T (T_e^a - T_e) + |q|(T_p - T_e) \\
V_p \frac{d}{dt} T_e &= C_p^T (T_p^a - T_p) + |q|(T_e - T_p) \\
V_e \frac{d}{dt} S_e &= C_e^S (S_e^a - S_e) + |q|(S_p - S_e) \\
V_p \frac{d}{dt} S_e &= C_p^S (S_p^a - S_p) + |q|(S_e - S_p)
\end{aligned} \tag{2}$$

It is standard to assume that the relaxation rates for the temperature are equal and constant, so $C_e^T / V_e = C_p^T / V_p = R_T$ and similarly for salinity $C_e^S / V_e = C_p^S / V_p = R_S$. With this simplifying assumption the equations can be combined to obtain two differential equations for the temperature difference $\Delta T = T_e - T_p$, and the salinity difference $\Delta S = S_e - S_p$ as

$$\begin{aligned}
\frac{d}{dt} \Delta T &= R_T([T_e^a - T_a^p] - \Delta T) - 2|Q|\Delta T \\
\frac{d}{dt} \Delta S &= R_S([S_e^a - S_a^p] - \Delta S) - 2|Q|\Delta S
\end{aligned} \tag{3}$$

where

$$Q = \gamma(\alpha_T \Delta T - \alpha_S \Delta S) \tag{4}$$

Now define $\Delta T^a = T_e^a - T_p^a$ and $\Delta S^a = S_e^a - S_p^a$, and rescale the equations by setting

$$x = \frac{\Delta T}{\Delta T^a}, \; y = \frac{\alpha_S \Delta S}{\alpha_T \Delta T^a}, \; \tau = R_S t \tag{5}$$

(this is a slightly different scaling than that used in [3], and will make it possible to use singular perturbation theory a little later; the same idea is used by Berglund and Gentz [1] on a slightly different model) giving

$$\begin{aligned}
\frac{d}{d\tau} x &= \frac{R_T}{R_S}(1 - x) - A|x - y|x \\
\frac{d}{d\tau} y &= \mu - (1 + A|x - y|) y
\end{aligned} \tag{6}$$

where

$$\mu = \frac{\alpha_S \Delta S^a}{\alpha_T \Delta T^a}, \; A = \frac{2\gamma\alpha_T \Delta T^a}{R_S} \tag{7}$$

To understand the behaviour of this system make one further (reasonable) assumption:

$$R_S << R_T \tag{8}$$

so that $\varepsilon = R_S / R_T$ is a small parameter. Then (6) can be rewritten as

$$\begin{aligned}
\varepsilon \dot{x} &= (1 - x) - \varepsilon A|x - y|x \\
\dot{y} &= \mu - (1 + A|x - y|) y
\end{aligned} \tag{9}$$

where the dot denotes differentiation with respect to τ. This is a slightly non-standard equation in the sense that the right hand side of the equations are piecewise smooth, but there is a discontinuity in their derivatives at $x = y$ (these systems arise naturally in many control systems [2]). Apart from this, (9) is in the standard form to apply singular perturbation theory, and assuming that the standard theorems of Tikhonov and Fenichel apply here we may assume that solutions are attracted close to a surface

$$x = 1 + O(\varepsilon) \tag{10}$$

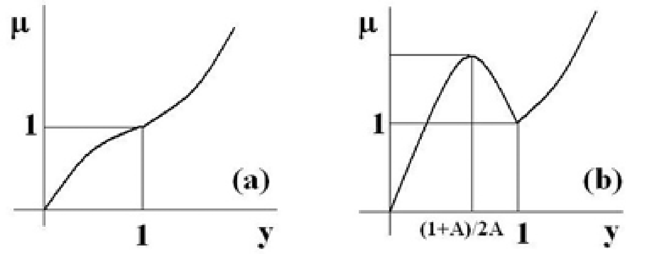

Figure 2: Sketch of the graphs of the right hand side of (12) for (a) $A < 1$; and (b) $A > 1$.

I have not checked that the standard theory does apply globally, although the piecewise smooth nature of the equations means that it applies locally provided $x - y$ is non-zero. Assuming (10) then the leading order equation for y is

$$\dot{y} = \mu - (1 + A|1 - y|) y \tag{11}$$

which we now consider with A constant and μ varying — this corresponds to changing the external salinity inputs. The stationary solutions, obtained by solving the quadratic equations when the left hand side of (9) is set to zero, satisfy

$$\mu = \begin{cases} (1 + A) y - Ay^2 & \text{if } y < 1 \\ (1 - A) y + Ay^2 & \text{if } y > 1 \end{cases} \tag{12}$$

and as shown in Figure 2 the nature of solutions depends on whether A is greater than or less than one.

If $A < 1$ then the right hand side of (12) is a monotonic increasing function with a discontinuity in the derivative at $y = 1$, so for each value of μ there is a corresponding fixed point which is stable (it is stable if the slope of the graph in Figure 2 is positive at the fixed point).

The case $A > 1$ is more interesting. In this case the turning point of the quadratic defined in $y < 1$ is $y = (1 + A)/2A$, where the function takes its maximum value $\mu = (1 + A)^2 / 4A$. Since this turning point now lies in $y < 1$, the graph of (12) has a turning point in $y < 1$ and the decreasing branch attaches to the increasing branch of the parabola in $y > 1$ at $y = 1$, at which the right hand side of (12) takes the value of unity. Thus there are three cases (imagine moving a horizontal line of constant μ up or down in Figure 2):

- if $\mu < 1$ the system has one stable solution in $y < 1$;
- if $1 < \mu < (1 + A)^2 / 4A$ there are three solutions: two of these are stable, one in $y < 1$ and the other in $y > 1$;
- if $\mu > 1$ then there is a single solution with $y > 1$.

The bifurcation at $\mu = 1$ is a non-smooth version of the saddlenode bifurcation, creating a pair of stable and unstable solutions; the bifurcation at $\mu = (1+A)^2/4A$ is a standard saddlenode bifurcation. The stability diagram is usually described as in Figure 3, which is essentially Figure 2b turned on its side, and shows the evolution of the fixed points as a function of μ.

Current estimations show that the ocean parameters are such that we lie on the upper branch of the stable solutions in Figure 3. This suggests two sources for concern: if the parameters are in the (middle) bistable region, then a perturbation of initial conditions, i.e. some extraneous effect not described by the box model of Figure 1, could push us onto the weaker lower stable solution; or the parameter μ may be changing in such a way that it approaches $\mu = (1+A)^2/4A$ at which point the system moves dramatically to the stable lower solution.

Whether such an event might really occur is a matter of speculation. There is geological evidence that it has happened in the past, and effort is being put into more sophisticated models to see whether the behaviour predicted by the box model is also observed in more complete models. The basic box model idea can be complicated by including further boxes, or making different assumptions about salinity and temperature changes, and the issue of noise has also been investigated.

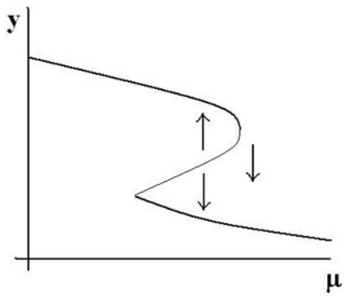

Figure 3: Sketch of the bifurcation diagram when $A > 1$. Stable stationary solutions are indicated by heavy lines and unstable solutions by thin lines. Note the region of bistability.

From a mathematical point of view, it is nice to find systems which climate scientists find interesting but which also pose interesting mathematical questions. More generally, it is clearly important that mathematicians continue to play a role in the discussions, modelling and simulations surrounding climate change.❑

Acknowledgements

The idea for this article came from an excellent lecture by Henk Dijkstra at the SIAM Conference on Dynamical Systems and its Applications, Snowbird, May 2009. I am grateful for support from the Royal Society which allowed me to attend that conference.

E-mail: P.A.Glendinning@manchester.ac.uk

REFERENCES

1 N. Berglund and B. Gentz (2002) Metastability in simple climate models: pathwise analysis of slowly driven Langevin equations, *Stochastics and Dynamics*, **2** 327—356.
2 M. di Bernardo, C.J. Budd, A.R. Champneys and P. Kowalczyk (2008) *Piecewise-smooth Dynamical Systems: Theory and Applications*, Springer, London.
3 H.A. Dijkstra (2005) *Nonlinear Physical Oceanography*, 2nd Edition, Springer, Dordrecht.
4 S. Rahmstorf (2006) Thermohaline Ocean Circulation, in *Encyclopedia of Quaternary Sciences*, ed. S.A. Elias, Elsevier, Amsterdam.
5 H.M. Stommel (1961) Thermohaline convection with two stable regimes of flow, *Tellus*, **13** 79—89.