Durham University

Insights

Institute *of* Advanced Study

Palaeoclimate Histories

John Haslett and Peter Challenor

Volume 3 **2010** **Number 1**

ISSN 1756-2074

About Insights

Insights captures the ideas and work-in-progress of the Fellows of the Institute of Advanced Study at Durham University. Up to twenty distinguished and 'fast-track' Fellows reside at the IAS in any academic year. They are world-class scholars who come to Durham to participate in a variety of events around a core inter-disciplinary theme, which changes from year to year. Each theme inspires a new series of Insights, and these are listed in the inside back cover of each issue. These short papers take the form of thought experiments, summaries of research findings, theoretical statements, original reviews, and occasionally more fully worked treatises. Every fellow who visits the IAS is asked to write for this series. The Directors of the IAS – Ash Amin, Michael O'Neill, and Colin Bain – also invite submissions from others involved in the themes, events and activities of the IAS.

About the Institute of Advanced Study

The Institute of Advanced Study, launched in October 2006 to commemorate Durham University's 175th Anniversary, is a flagship project reaffirming the value of ideas and the public role of universities. The Institute aims to cultivate new thinking on ideas that might change the world, through unconstrained dialogue between the disciplines as well as interaction between scholars, intellectuals and public figures of world standing from a variety of backgrounds and countries. The Durham IAS is one of only a handful of comparable institutions in the world that incorporates the Sciences, Social Sciences, the Arts and the Humanities.

The focal point of the IAS is a programme of work associated with, but not exclusive to, an annual research theme. At the core of this work lies a prestigious Fellowship programme. This programme gathers together scholars, intellectuals and public figures of world standing or world-promise to address topics of major academic or public interest. Their mission is to anticipate the new and re-interpret the old, communicating across and working between disciplinary boundaries.

Every year, the Institute invites as many as twenty highly creative individuals to spend up to three months in Durham. They are located in Cosin's Hall, a magnificent and spacious 18th century mansion which, together with Durham Cathedral and Durham Castle, forms part of Palace Green, dominating the World Heritage Site of Durham Peninsula. During their stay, Fellows engage with departments and colleges, deliver public lectures and seminars, and, above all, join an international community of researchers to address the theme selected for that year. Further details of the IAS and its Fellowship programme can be found at www.durham.ac.uk/ias/fellows

Copyright

PALAEOCLIMATE HISTORIES

Two sources of complementary information about the palaeoclimate are (a) dynamical models of the Earth climate system and (b) proxy data such as pollen, from which reconstructions may be attempted. Both of these are subject to considerable uncertainty, but of very different types. Recent advances in quite different parts of statistics allow us to anticipate a new level of inter-comparison focussed on palaeoclimate dynamics. This brief paper outlines a research agenda.

Introduction

This paper sets out some potential lines of research at the interface between palaeoclimate modelling and palaeoclimate reconstruction from proxies. It is our contention that advances in both – and in particular advances in methods for the study of uncertainty in such systems – are such that it is now time to consider an entirely different level of inter-comparison, to the mutual benefit of both. Our focus here is regional climate since the last glacial maximum (LGM), about 21,000 years before present (BP). Our concern in this paper is with the history over this period of the space-time system which is climate.

The importance of the underlying scientific questions rests on the fact that, in the context of predictions on global climate change over the next century, there is very little instrumental data on past climate prior to the nineteenth century. Indeed, only in the last century do we have good quality data even on global atmospheric temperature; and only in the last few decades do we have any useful information on aspects of the climate system such as the global ice-cover, which we now know to play a crucial role in the dynamics of climate change. There is thus great uncertainty – and some futile debate – associated with claims concerning past and future climates.

For example, the 'Hockey Stick' debate spawned by the paper of Mann, Bradley and Hughes (1998) (hereafter MBH98) on the climate of the past six centuries led to heated scientific and even political debate. Using various environmental proxies and bore-hole temperatures this seemed to show that the climate of the Northern Hemisphere has been *relatively* stable from about the tenth to the nineteenth century (the 'handle' of the hockey stick, roughly constant, with estimated variations of ±0.25°C), but that rapid warming began in the early twentieth century (showing rise of about 0.4°C since then). An update was published recently (Mann et al., 2008).

MBH98 was much cited by the 2001 Intergovernmental Panel on Climate Change (IPCC) report. Yet its conclusions led to a heated debate in the scientific literature, much of that concerning the quality of the statistical methods, and in particular the uncertainties associated with these methods. In the light of the importance of the paper, this debate led to inquiries by the United States National Research Council at the request of the US Congress and another for the (Congress) Committee on Energy and Commerce – the Wegman report (2007).[1] In the context of this paper, the following remark in the Wegman report (p. 4) is pertinent: 'It is important to note the isolation of the paleoclimate community; even though they rely heavily on statistical methods they do not seem to interact with the statistical community.'

There have been several scientific responses. Examples include an editorial 'On uncertainty and climate change' (Dessai et al., 2007) in a special issue of *Global Environmental Change*. This urges the wide climate community (including politicians and lawyers) to invest in the science of uncertainty. The UK Climate Impacts Programme[2] (UKCIP08) now offers a fully probabilistic framework for the assimilation of historic data and the presentation of future scenarios. Projects and networks such as PalaeoQUMP[3] (Quantifying Uncertainties in Model Prediction) have placed uncertainty modelling at the heart of the debate. An international network, SUPRanet, focussed on the Study of the Uncertainty in Palaeoclimate Reconstruction, has been established. Its first meeting (in Sheffield in June 2008) has itself contributed to several of the ideas in this paper.

Of particular interest is that some past climates are very different from the recent past. The methodologies of climate modelling have been created for, and validated by, climates very similar to the current climate. But although the climate of the recent past (the last two centuries) has been relatively stable (notwithstanding MBH98), atmospheric levels of carbon dioxide (CO_2) have increased to about 380 ppm (from a pre-industrial base of about 280 ppm, relatively stable over the past 10,000 years) during that period; indeed global ice-cover also has changed. For example, the National Oceanic and Atmospheric Administration (NOAA) reports[4] that the extent of perennial Arctic sea ice is now less than 50% of its value in 1957. These raise questions about the value of such modelling in a rapidly changing environment. Such questions provide much of the scientific interest in inter-comparisons. Thus, for example, the Palaeoclimate Modelling Intercomparison Project (PMIP)[5] – already more than 10 years old – has generated almost 50 papers since 2002; its focus is the LGM and the mid-Holocene. But, as outlined below, the modelling of the uncertainties in such inter-comparisons raises many challenges; and it is on these that we focus.

Uncertainty in the discussion of climate is inevitable. But ignoring it – or using over-simple methods – is not. Modelling uncertainty (and indeed communicating it) is neither simple nor cheap; however it cannot be set aside. In fact, confidence about the handling of uncertainty not only pre-empts the sometimes sterile debate encouraged by unqualified (or poorly qualified) statements about climate change, past and future; but it goes further. It provides a basis for pooling the many different sources of information and thus for reducing the uncertainty associated with at least some statements. In this paper we discuss some of the challenges and look forward to the fruits of recent developments. First, we must define some terms.

Climate modelling

In the following, we use the term 'climate model' to refer to models that study – using equations of energy transfer – the physical processes that transfer solar energy to the climate. Climate models can be divided loosely into three classes: energy balance models, general circulation models (GCMs) and Earth system models of intermediate complexity (EMICs) (McGuffie and Henderson-Sellers, 2000). Energy balance models are the simplest; they balance the radiation input from the Sun with a simple description of the Earth's climate system. They do not have a spatial component. GCMs on the other hand use atmospheric models (similar to, but simpler than, the models used for weather forecasting). Increasingly these are coupled with models of the ocean and sea ice of similar complexity; some also include further components dealing with both the terrestrial and oceanic biosphere. For example, the Hadley Centre models such as HadCM3 have components HadOCC and TRIFFID which deal with the oceanic and terrestrial carbon cycle components respectively. They have provided much of the basis for the predictions of the IPCC. They are expensive and can take months of time on the world's fastest computers; the best are available only to a few groups. The EMICs fall between these two classes in terms of complexity.

Modelling iteratively updates the climate at cells defined on a spatial grid over discrete time intervals. The largest, and thus most expensive, climate models work on spatial scales of the order of 100 × 100 km; energy balance models have one spatial cell for the entire globe. The large GCMs are integrated with similar models for the oceans, as we discuss; this adds to model complexity. The time step for the most detailed GCMs is of the order of an hour. Output is usually summarised as regionally averaged values over a period of a year. Although the computer programmes required to implement GCMs run on very fast computers, the run-time can nevertheless be very slow. For example, depending on the model configuration and the computer, between 6 and 12 hours are needed to run HadCM3 for one model year.

EMICs work with much cruder space and time scales, and ignore – or crudely model – aspects such as change in ice-cover and the biological environment. Examples are GENIE (also known as C-GOLDSTEIN) (Edwards and Marsh, 2005), the University of Victoria model (Weaver et al; 2001) and CLIMBER (Petoukhov et al., 2000). They represent a cut down version of the full physics present in the GCMs but in a form that is computationally efficient. Some modern models, for example GENIE-1, have a 3D ocean but often only have a simple energy balance atmosphere. Recently some, such as GENIE-2, have been produced with a dynamic atmosphere. The advantage of EMICs is that they make it possible to look at very long time periods covering, for example, glacial to interglacial transitions. Until advances in computing hardware make GCMs much more accessible, EMICs will form the basis of much research on climate uncertainty. The project we envisage, focussing on the climate since the LGM, will be utterly reliant on these in the short term. But they allow us to foresee the use of GCMs in the medium to long term.

Palaeoclimate reconstructions

By 'reconstruction from proxies,' we mean statistical inferences for aspects of past climate given data on various types of proxies; typically these are based on the analysis of chemical and/or environmental features in cores such as ice and sediment. These proxies are indirect methods of measuring climate; their basis depends on an observed correlation between these and aspects of measurable modern climate. Pollen in lake sediment is one example of a bio-environmental proxy; we use this as a running example. Ocean sediments and ice provide other proxies. Of course, much more proxy data is available to study the past few centuries than for the LGM, for example, and the data there are much more uncertain.

Different proxies can provide information on different aspects of climate and at quite different scales, in time, and most particularly, in space. Ice-cores, like tree-rings and corals can show annual variations clearly; they are said to be laminar and thus to permit dating with relatively little uncertainty. For example, in the Greenland GISP[6] ice core the stable oxygen isotope ^{18}O data provides indirect information on atmospheric temperature at temporal resolution of one year, at least for a few millennia. This allows the IPCC Working Group 1 to state, in 2007, 'During the last glacial period, abrupt regional warmings (probably up to 16°C within decades over Greenland) occurred repeatedly over the North Atlantic region' (IPCC, 2007). But in the Northern Hemisphere only Greenland has ice sheets thick enough to provide cores of sufficient depth to permit such statements; there is little spatial scale.

By contrast, there are thousands of sites with ancient pollen in lake sediment. The spatial resolution is excellent. But the temporal resolution is poor; for example, it may take more than a century for an oak forest to respond to climate change. In addition there can be much uncertainty in the dates attributed by radio-carbon dating to the samples. Ice, being laminar, allows much higher precision in dating, and can resolve climate changes annually. However, pollen can serve as a proxy for several aspects of climate, including moisture-availability, annual temperature minima and the length of the growing season, whereas ^{18}O provides information only on mean temperatures.

Multi-proxy reconstructions are thus attractive, but the modelling for these is still under-developed. For example, the MBH98 study used multiple proxies (primarily Northern Hemisphere tree-rings, ice-cores and coral-reefs) together with bore-hole temperatures (which may be thought of as direct (although imperfect) instrumental measurements of past surface temperature, reflected in modern measurements at different depths). Their procedure was – in essence – multivariate linear regression. The first step correlated various proxy space-time series with the 'known' instrumental space-time history of the temperature of the Northern Hemisphere for the past century; multivariate (i.e., spatial) temperature was regressed on the proxies. The (known) proxy series for the previous nine centuries were then used, with the fitted model, to predict the unknown Northern Hemisphere temperature. The procedure also employed various technical statistical procedures (all linear) for 'dimension reduction.'

We will argue below that the critical failing of such an approach is that it concatenates the several sources of random error into one; it does not model them. Indeed there is only an implicit use of the normal probability distribution and a presumption that this is satisfactory. Little attention was paid to the dating uncertainties or to auto-correlation. The modelling is based exclusively on correlation, and linear correlation at that. It is this, we argue, that undermines the confidence in the uncertainties associated with such calculations. In defence of the method, it represents the only method available with which to form regional reconstructions from multi-proxy data from classical statistical tools.

Comparisons

The challenges facing both palaeoclimate modelling and reconstructions are such that comparisons, while important, have been limited. For example, the PMIP project, has focussed on only two points in the past, BP and the mid-Holocene (~6000). Thus longer term climate dynamics, including the study of possible abrupt changes in the future climate, have been much understudied. As a case in point, the abrupt changes referred to by the IPCC were driven it seems by changes in the Thermohaline Circulation, the process that governs the flow of the major ocean currents. These events raise questions on its future stability in the context of changes in atmospheric composition. We propose here to look at the entire period from the LGM to the recent past – the palaeoclimate history – and as such we will shed light on the dynamics of change.

This short paper elaborates on some of these issues and challenges. It illustrates how recent developments in the methodologies in both fields may now permit a great deal more inter-comparison than it has previously been reasonable to aspire to. We review in the following sections the state of the art in climate modelling, in reconstruction and in inter-comparisons. We subsequently outline research that may be able to exploit much further the interface between these different approaches to the uncertain and multi-scale space-time history that is the palaeoclimate.

Uncertainty in Climate Models

Climate models, based on geophysical fluid dynamics, are deterministic in principle. That is, running the same model twice with the same inputs will give the same outputs.[7] Notwithstanding this, there are very many sources of uncertainty. Here we lay out some of the difficulties and some of the ways forward, building on recent progress in modelling uncertainty in complex computer models. We focus on model inputs and outputs.

We use the term 'model inputs' to describe all external inputs to the model. These include time-independent parameters, for example the physical constants needed to solve the Navier-Stokes equations, the initial state of the model and any forcing variables. The parameters include physical constants such as the acceleration due to gravity, for example, that are well-known. However it also includes empirically determined parameters. For example in the ocean component, the numerical value for the viscosity of sea water, as used in computation, is several orders of magnitude higher than the value as measured in the laboratory. This is because the model viscosity is being used to parameterise processes that occur on a scale smaller than the grid on which the model equations are solved on. The smaller the model grid the closer we expect these 'constants' to be to their measured physical values. For the biological components of the models there is often no measurable real world quantity. These empirical quantities are estimated through a combination of expert opinion and initial model runs (so-called informal tuning). With the exception of well-known physical constants we expect all the model parameters to be subject to a greater or lesser degree of uncertainty.

Minimally, climate models, given a set of defined inputs, sequentially compute and (optionally output) values $C_M(s,t)$ for values of time t over a specified period, where s labels grid-cells on a sphere. Here $C(s,t)$ denotes many aspects of multivariate climate at spatial locations s and times t and $C_M(s,t)$ is its model equivalent. These aspects include, for example, measures of atmospheric pressure, temperature and moisture, and ocean temperature and salinity; these will include values for these at different heights in the atmosphere and depths in the ocean. (For simplicity we are excluding anything other than surface data from C_M; it is simple to extend s to include climate variables at different levels.) If a detailed GCM is coupled with an ocean model, for example, many aspects of the changing ocean must also be computed. For GCMs the spatial resolution can be 1° to 3° corresponding to grid boxes (away from the poles) of the order of 100 to 300 km square. For simpler models the spatial resolution is much more coarse.

Jointly, for all (s,t) we can refer to these as $C=\{C_M(s,t); all\ s,t\}$; the model equivalents are $C_M(s,t)$ and C_M. The length q of C (and C_M) is the number of locations in space-time; we use the index $u \subset \{u_i; i=1,q\}$ to refer to these; thus $C(u) \equiv C(s,t)$ and equivalently $C_M(u) \equiv C_M(s,t)$. Potentially C_M is a very high-dimensional variable. However, interest usually lies in a few regional summaries. For notational simplicity in the following, C_M can refer to such summaries; context will provide guidance. The values of C_M depend on several inputs whose values must be specified. In this paper the challenge lies: in the analysis of the uncertainty surrounding C_M; in how the inputs impact uncertainty; in the comparison under uncertainty of C_M with its equivalent from proxy-based reconstructions C_P; and in the uncertainties for C in the light of either or both of C_M and C_P.

The generation of such summaries of climate history is thus both computationally challenging in the extreme and subject to considerable uncertainty. Even 30-year averages of, say, three climate dimensions for a region such as Northern Europe (comprising some tens of grid-cells) over the period since the LGM is a vector of length q=3x21000/30 = 2100 for each region; this vector defines a regional climate history. Any change in parameters or initial conditions will result in a different history; given the cost of computing even one C_M, studying its uncertainty is a challenge.

Thus, while in principle deterministic there is much uncertainty in the output from climate models. However, several recent initiatives have developed new methods for studying uncertainty in computer models in general and climate models in particular. For example, Research Councils UK (RCUK) has funded the Managing Uncertainty in Complex Models project as part of its basic technology programme and the National Environment Research Council (NERC) has funded the Probability, Uncertainty and Climate Modelling (PUCM) workshops and research playgrounds. (See http://www.noc.soton.ac.uk/pucm for the results of the first PUCM research playground.) In this section, we discuss methods for the statistical analysis of climate models; focussing on one aspect – the use of emulators. First, we expand on the sources of uncertainty.

Sources of uncertainty in climate models

As we have seen, these climate models are nominally deterministic; nevertheless uncertainty enters in two ways: structural and input. Models are not perfect; some processes are missing and others have been simplified by parameterising empirically. Thus we are even uncertain about how well the model captures the 'true' model structure. In particular, spatial resolution is always problematic. Important climate processes (e.g. turbulence) are not seen on certain scales. Assessing the uncertainty due to imperfect model structure remains a challenge in many areas, including climate; see Goldstein et al (2008) for one framework for discussion. This issue is not pursued here.

The second form of uncertainty concerns the inputs to the model. The simplest version of this concerns parametric uncertainty. In essence, if we specify our uncertainty on the parameters we can propagate it through the model to obtain uncertainty estimates on the outputs. For example, we may be interested in the probability that the global mean temperature in 2050 is greater than 4°C above the present; observe that this is the spatial average of one component of $C(_s,2050)$. To calculate this probability we might use Monte Carlo methods to sample from the probability distributions that represent our input uncertainty and to run the model for each sampled set of inputs, thus building up a distribution of outputs $C_M(_s,2050)$. From this distribution we could then estimate the required probability; see for example, Challenor et al (2006). The problem with such methods is that they require huge numbers of model runs (~10000), particularly if we want to estimate small probabilities. For detailed models this is therefore completely infeasible.[8]

Another challenge arises from 'forcing functions.' In principle if $C(s,t)$ is fully defined, then it determines *all* aspects of the changing climate, including, for example, the changing ice-sheets and the changing CO_2 concentration. We can presume that in the past mankind drove neither of these. The only exogenous aspect is now the past changing energy production of the Sun and the past variations in the Earth's tilt. The former is well understood (despite the circularity that such knowledge comes to us via past climates), but is of course uncertain; the latter is of course known. If, as is typical for most simple models, $C_M(s,t)$ does *not* contain a dynamic model of glaciation and ice sheets and of changing CO_2 concentration, then without intervention, the model will run under (implicitly) constant conditions. If the model has a dynamic atmosphere the climate will vary in a 'chaotic' way; if it has an energy balance atmosphere the climate will go to a steady state or near steady state without forcing.

For prediction on future climate, the anticipated changes in CO_2 are a critical part of the modelling; these changes are input via externally specified *forcing* functions, and interest lies in climate response to different specifications. These are known as 'scenarios', reflecting different choices that society might make in the hope of ameliorating the impact of climate change. For past climate, it is natural to input our knowledge of past CO_2, imperfect as it is, and indeed past values for the size of the ice sheets. Effectively these functions are parameters of the model, and their uncertainties also impact the uncertainties in C_M (and its relevance to C). However, the methodologies of using uncertain forcing functions are under-developed, at least for all but the least computationally demanding models; see for example Pastres and Ciavatta (2005).

The initial conditions – an input required by iterative models – represent a yet more difficult challenge. If interest lies in $C(s,t)$ for many time points between the LGM and the present, then the initial conditions are in principle $C_M(s,t_{LGM})$ for all spatial locations s. But these are completely unknown. One widely used procedure is to specify, rather arbitrarily, values for $C_M(s,t_0)$ for the climate at some time t_0 well prior to the LGM (that is, $t_0 << t_{LGM}$), to run the model forward to t_{LGM} and to use the resultant $C_M(s,t_{LGM})$ as a starting value. The idea is that the model climate will 'forget' the initial values during the spin up and come to the same value of $C_M(s,t_{LGM})$ regardless of the state at t_0.

In summary, therefore, although climate models are in principle deterministic functions built on well-established physics, in practice they are approximate. The output can be high dimensional; it is potentially sensitive to *many* pre-specified inputs whose values are uncertain. Further, the models can be slow. This inhibits greatly the very necessary experimentation needed to understand such sensitivities and uncertainties. Nevertheless, problems involving uncertainty and complex computer codes, such as climate models, have received much attention in the statistical literature in recent years. Emulators are new and important tools permitting fast experimentation in the context of uncertainty. Recent developments allow us to anticipate future developments for palaeoclimate modelling.

Emulators

An emulator is a statistical approximation to the climate model; see for example, Kennedy and O'Hagan (2002), O'Hagan (2006), Kennedy et al (2007). Alternatively, from a Bayesian point of view, it is a summary of our beliefs about C_M given specified input values, referred to below as θ. (This notation is natural for a few parameters; it is perhaps stretched when we consider forcing functions and initial conditions, but the uncertainty about these may itself be focussed on a few parameters.) Using an emulator we can produce an estimate $C_M(\theta_0)$ of what the model will produce at a specified value θ_0. The basis for this is a set of values $C_{M,train}=\{C_M(\theta_i); i=1,n\}$ obtained by running the climate model at a designed set of n locations in parameter space (i.e., values of θ) to produce a 'training set' of inputs and outputs. We then use these runs to build the emulator. We can now think of $C_M(\theta)$ as an interpolator between the points in the training set, an interpolator whose precision is known.

The emulators produced are univariate, we can reproduce what the climate model would produce at any point in space or time. We would like to emulate the entire multivariate field, C(s,t), simultaneously so we can look at the relationship between different positions and times. In this case where our outputs are indexed we can change the multivariate problem to a univariate one by adding the index as an additional input. We can now emulate the entire space-time field but at the cost of hugely increased dimensionality in the input space. As noted by Rougier (2008), implemented naively this requires $O((nq)^3)$ flops, where n is the number of original inputs and q is the size of the space time field. By making some assumptions of separability (e.g. that the spatial characteristics of the climate model do not change with time), Rougier proposes the exploitation of special computing structures that are particularly well-adapted to emulating space-time functions and make such calculations feasible.

The implications of this are that modern emulation methods can in fact be used to experiment with exactly the situation we have outlined above; that is, where C and C_M are space-time series. The prerequisite is still that we have available n runs of the climate model under study corresponding to different values of θ, where n is such that we can cover the input space associated with the dimensionality of θ, itself having many dimensions. Thus for the near future we must anticipate that the research we propose here is based on EMICs, relies on spin-up to render the starting conditions irrelevant and to hold the forcing functions to be sufficiently well understood to be treated as known.

In conclusion, within this context we can compute random draws of palaeoclimate history $C_M(\theta)$ from a model that can generate such histories for each of a specified set of n inputs θ_i. We now consider how we may do the equivalent for proxy-based reconstructions.

Uncertainties in Palaeoclimate Reconstruction

We describe in this section how to quantify information on the palaeoclimate available in proxy data, such as pollen composition at given depths in sediment samples. Technically the task is not 'reconstruction' per se; it is in fact the statement of our uncertainties about the palaeoclimate conditional on data and models; but we use the term here, as it is in wide currency. With existing models, these uncertainties are currently large; but versions under development allow pooling information from many sources and the uncertainties are anticipated to decrease. Perhaps paradoxically, we shall find that it is in many ways easier to *state* the uncertainties in palaeoclimate reconstruction than it is in physical climate modelling. The key to the new advances lies in modelling the complex system by which climate changes cause changes in the proxy data. This system is not as well understood as the physical climate. There are no 'constants of nature' to draw on; empirical methods are inevitable. But much modern statistical research is focussed on inference for complex systems and much can be anticipated for the future.

Methods based on proxies use frequency data from cores (often rendered as relative frequencies, e.g. pollen composition, ratios of oxygen or carbon isotopes). The details of the relationships between the proxy frequencies (sometimes highly multivariate) and the climate (typically low dimension) are established by statistical means using 'training data,' being modern and/or laboratory data; these relationships are referred to here as response surfaces; alternative phrases include calibration curves and/or transfer functions. With a presumption that the observed *modern* relationship is of relevance to the fitting of such surfaces (the 'space-for-time substitution'), sample compositions obtained from *ancient* data in core samples are transformed by fitted surfaces to estimates of the corresponding ancient climate. We consider below recent developments in this area. For illustration, we continue to rely on pollen-based reconstructions.

There are three broad approaches on which we focus: dimension reduction techniques; nearest neighbour algorithms (see for example, ter Braak, 1995); and model-based analyses (Haslett et al., 2006). The first are characterised by methods such as dimension reduction and multiple regression where the algorithms are based on means and covariances. One widely used such method draws on 'correspondence analysis.' In the second, nearest neighbours are preferred; here, a multivariate proxy measurement is allocated to a particular climate on the basis of some distance measure between it and those in the training data; the allocated climate is that of its nearest neighbour in pollen space, or variants thereon. One such variant is Huntley (1993); there interpolation in climate space yields surfaces and changes the meaning of 'neighbour.'

We describe all three approaches, including Haslett et al, together with MBH98, as 'purely correlational'; there is no attempt to model a data generating *process*. Observe however the key difference between these three and the MBH98 procedure: scientific theory about the data generating process tells us that *selected modern* training data are of relevance to inference on the details of the response of the proxies to climate change. MBH98 relied only on *past* measurements of temperature (and only temperature) corresponding to little over a century. As there is much modern data (many thousands in the European Pollen Database, though fewer for other proxies), the inference process is thus potentially more precise.

Linear methods, the simplest of the three, have inherent difficulties. A specific example is that climate proxies can disagree, sending conflicting messages, 'preferring' two different climates. A natural statement of the uncertainty would be a bimodal distribution. Linear methods will necessarily 'average' these, even perhaps suggesting a climate at odds with both sets of evidence. Nearest neighbour algorithms – including Huntley (1993) – reflect a desire to avoid the dangers of linear methods (although they often average over the nearest ten neighbours, for example). They have in the past been used in many clustering algorithms in many areas of science, and in climate reconstruction; see for example Huntley (1993) and Hill et al (2003).

The great difficulty with these two methods is again there is no principled modelling of the data generating process, and thus of the uncertainties. Therefore, for example, there is no clear way to combine information from different proxies. Indeed reconstructions at all space-time locations are conducted independently. Typically the location is time inferred (without uncertainty) from the depth. Such methods describe, but do not model, uncertainty; the only available measures of uncertainty are based on one-sample-at-a-time cross-validation. Nevertheless they are computationally fast and available for many different types of proxy.

Model-based approaches are increasingly able to overcome such challenges. Such methods can in principle study, and be optimised for, the multiple sources and different types of interconnected uncertainties. However, model-based methods are, relatively speaking, new and remain computationally challenging. The outputs of such methods are random 'palaeoclimate histories' that are consistent with the data and with the properties of the processes involved. We develop this below.

Model-based analyses involve building specific statistical models of the processes involved. The key is to model the (forward) data generating process from climate to data (e.g. pollen counts at given depths in multiple lakes), and then to *invert the inference* which flows from data to climate. Thus plant taxa (such as oak and hazel) respond to changes in climate; this is reflected in 'pollen rain,' some of which accumulates in the sediment of lakes across Europe; and some is sampled and examined in the laboratory. Note that at any one site, many taxa will be completely absent for long periods; the climate – or other aspects of the environment – are completely unsuitable; the process will produce many counts of precisely zero which carry less strength for inference. There are presence-absence variables and count variables. There is decay in the ^{14}C in the organic material collected in the sediment; this process yields information on the 'radio-carbon ages' of the samples at different depths; the stochastic decay and the random errors induced by the laboratory measuring procedures are well understood. Ultimately a sample is examined under a microscope, pollen of multiple types is counted and organic material is dated. Sampling and counting errors arise in this process.

Mostly, climate changes slowly in space and in time; but sometimes abruptly; oak responds more slowly than hazel; counts are zero-inflated; ^{14}C decays as a Poisson process; sedimentation is monotone. The entire system is a stochastic process; but the system interconnections can be modelled. The broad features of all these processes are well understood scientifically. The technical task, and the focus of much modern statistical research, is to invert the process to permit statistical inference, allowing information to flow from data to climate. Progress in statistical inference from such systems – in many other areas of application – is what is driving the advances. Much of this is built on Bayesian hierarchical models and is computationally intensive. This inversion involves computational shortcuts, and approximations are involved. But the underlying model is well-defined and it is this that provides the principled basis for modelling uncertainty.

Haslett et al (2006) (H06) offered a proof-of-concept for pollen and by extension for other proxies. But the computational challenges were large, and relied on Markov chain Monte Carlo (MCMC) for which there remain many difficult technical issues, such as 'convergence' especially in large models and even more so in cross-validation. Unpublished work in the PhD thesis of Salter-Townshend (2008) shows that it is now possible to use Bayesian modelling while completely MCMC, retaining rich models for proxy response (e.g. both presence/absence and count data; fast cross-validation; correlation induced by nested structures such as broad 'woody' and 'herb/grassy' taxon groupings with specific taxa (e.g. *Quercus*, *Corylus* within 'woody').

The research discussed in H06 has as its objective the use (jointly) on all samples in a core to yield a 'palaeoclimate history'; that is a (joint) climate reconstruction based on every sample in the core. (Developments since (Salter-Townshend, 2008) overcome many of the difficulties discussed there.) Indeed it envisages more than this, for a space-time history can be envisaged as the simultaneous goal for analyses of multiple cores, possibly with different proxies. Model-based methods can use elementary knowledge of climate dynamics (such as climate change in succeeding decades being generally small, but occasionally large) to generate histories that are realistically smooth; formally these are joint ('all-samples-together') reconstructions.

There are very many challenges to this. For example, all too often joint analyses involving *all* the samples in one or more cores rely on further ad-hoc procedures that ignore the uncertainty in, for example, the dating. Recent progress (Haslett and Parnell, 2008; Parnell et al., 2008) has however shown that there are fast and flexible statistical procedures which can recognise and work with dating uncertainties in the context of proxy-based palaeoclimate reconstruction from many cores simultaneously.

The technical approach of H06 is as follows. The essential idea is that, *given* the true space-time history of the palaeoclimate in regions as large as Northern Europe or as small as Durham, it becomes possible to define simple and natural joint statistical models for all observables. One specific and rich model is: the observed count N (for example, of *Corylus* at the location s and time t associated with its depth $d(t)$) is modelled as a realisation of a zero-inflated Poisson distribution; that is, with probability $1-q(s,t)$, $N=0$, and with probability $q(s,t)$, $N \sim Poisson(\lambda(s,t))$, these parameters depending on climate $C(s,t)$. Furthermore we can assert that the count for every other taxon is similarly modelled, and further that all such variation (across taxa, space and time) is (conditionally) statistically independent. Indeed we could go further and include similar remarks about all other observable bio-proxies. Thus, conditional on complete knowledge of $C(s,t)$ and of the functions $(q(s,t), \lambda(s,t)$, we can write down the joint likelihood of all the observed data.

If further we can assert that $(q(s,t), \lambda(s,t))$ be smooth non-linear functions of $C(s,t)$ and that $C(s,t)$ must itself be a smooth function of (s,t), we can make more progress. Modelling such smoothness by regarding each of the parts of $(p(s,t), \lambda(s,t), C(s,t))$ via random functions, then we can write down a posterior probability distribution for $C(s,t)$ given ancient proxy data and modern training data. To emphasise that the distribution refers to a model of the palaeoclimate derived from a specific model for proxy data, we refer to this as the distribution of $C_P(s,t)$. If the data are derived from lake sediment, then of course such histories reflect the 'catchment.'

If proxy data are in fact available at sites s_1,s_2 (and there are several thousand such sites in Europe alone) then we can go further than making separate (or even joint) inferences on $C(s_1,t)$, $C(s_2,t)$. For, given a model of the *spatial* process $C(s,t)$ for fixed t, it is now entirely possible to make inference – in a principled fashion – on *regional* climates from local proxy data. Indeed, even if the data represent multiple proxies this is not difficult. Further, there is no impediment in the fact that the ages of the various samples are themselves unknown (see Haslett and Parnell, 2008; Parnell et al., 2008).

Observe that our focus is the *joint* posterior for the entire random vector C_P comprising all values $C_P(s,t)$, that is for the entire space-time palaeoclimate history. We can thus sample from such distributions and thus produce (equally likely) random space-time histories for a specific point in space indexed by s. Such histories are consistent both with the data and with known smoothness properties for the processes involved.

One key aspect of the Bayesian approach is that there is no attempt to find the *best* reconstruction. The variation between alternate histories represents all the uncertainties, and simultaneously. Although this may be summarised for presentation – and depending on the focus (e.g. *average* temperature over a region at 6000BP; the *largest* one hundred year change in this temperature since the LGM) different summaries may be appropriate – the essential output is a space-time series of climate. It is this that suggests that the procedure is particularly suited for the comparison with regional histories generated otherwise, for example by emulation of a regional climate model.

In conclusion, we can generate random samples C_P that represent the true palaeoclimate history and our uncertainty about it. We can do this within a context of rich models of the many sources of uncertainty.

Interfacing Reconstructions and Models: a Research Agenda

There are two main conclusions from the previous sections. Firstly, climate modelling, especially concerning the palaeoclimate, is rife with uncertainties – notwithstanding the supposed deterministic nature of the modelling. These uncertainties are not easy to state. By contrast, modern methods of reconstruction – although noisy – do explicitly state the uncertainties. Secondly, climate modelling has poor resolution at fine spatial scale. Reconstructions on the other hand are explicitly spatial. For lake sediments the spatial scale is fine; ocean sediments will have poor spatial scale. Recent modelling developments as above suggest that it will soon be possible to use multi-site, multi-proxy data to form regional reconstructions in such a way that the uncertainties can be stated explicitly. Using both sources of information for joint inference is clearly desirable. But how shall we use these jointly?

In PMIP, comparisons have been made between model computations for the LGM and for the mid-Holocene; the comparisons are 'time slice' and the discussion of uncertainty is limited. The reasons for this are several: (i) detailed models for the dynamics of climate change are not yet available for all but the simplest energy balance type of model; (ii) no detailed studies of the many uncertainties have been attempted for climate models; (iii) no detailed studies of spatial aggregation have yet been attempted for model-based reconstructions from proxies; and (iv) no formal methods have been proposed in any detail for the comparison of the uncertain inferences.

We suggest that the recent methodological developments outlined above now allow us to foresee a future where emulated regional time series, covering extensive periods of time, can be constructively contrasted with palaeoclimate histories sampled from distributions informed by multiple proxies at multiple locations in space and time; further, the contrasts can be made within a formal framework of inference. There are several challenges here: but a research agenda is clear, if ambitious. It presumes the existence of proxy data banks (such as the European Pollen Data bank), both modern (for training) and ancient, accompanied by well-stated information on data quality (i.e., meta-data).

The agenda focusses on palaeoclimate histories and includes the following:
(a) To develop further forward models of the processes by which climate changes generate changes in proxy data.

This can be pursued in parallel and at different levels by groups focussed on different proxies; it includes the study of processes other than climate which may impact on the proxy, and it includes system dynamics. Note that as the focus is on *data* these processes also include counting, laboratory and dating errors. The research outputs are: stochastic models with parameters; and statements about the values of these parameters together with the inferential basis for these values.

(b) To develop a modelling framework for bringing these together as a unified process which, given parameters and given changing climate (and possibly other confounding covariates, e.g. anthropogenic), can generate stochastic multi-proxy data.

This framework will impose some common structure to the models above.

(c) To develop a variety of physical climate models for the palaeoclimate.

This will continue to be pursued in parallel by different research groups with different priorities. Some models will be quite simple, but will facilitate fast and flexible study of the climate since the LGM and its uncertainties. Others will be much richer, and will continue to focus on rather specific periods. All models will involve parameters; some will involve forcing functions and the choice of initial conditions. The assembly of agreed meta-data on these models will be the on-going challenge.

(d) To develop inverse inference procedures for models as in (b) given data as in (b).

Inference is focussed on (aspects of) past climate, but necessarily involves many parameters. The most flexible methods are likely to be Bayesian, but a variety of inference procedures will be used; some will need to be fast and thus approximate. For comparison purposes, some of these must be 'purely correlational.'

(e) A crucial part of (d) will be the way that these inferences are informed by the climate models at (d), given their stated uncertainties. This will involve much systematic research and many approaches will be used. These will include:

Qualitative descriptions of climate smoothness in space and time: are the spatial aspects of climate change adequately modelled by spatially isotropic processes? Are the spatio-temporal aspects of this separable? Does climate smoothness change in time?

Different ways of using forcing functions such as ice-cover and past CO_2: these include using them within the climate models and using them empirically as covariates within the inference process in (e).

For those climate models for which it is possible to use emulators to generate random samples of the space-time histories, how can these be used constructively with space-time samples from (e)? Options include 'history matching' where the proxy data, through the reconstructions, can be used in a relatively informal sense to 'screen-out' regions of (climate model) parameter space.

The ultimate objective of this agenda is of course to combine modelling and proxy reconstruction to make joint inference on the history of the palaeoclimate since the LGM. Comparisons are but a first step. As Bayesian perspective provides a framework for this also, the next steps can also be foreseen.

Conclusions

We have argued above that climate models and proxy reconstructions are complementary sources of information about the palaeoclimate. However there are difficulties in using them jointly. The reasons for this are two-fold. Firstly they are currently used at different scales: climate models are used to make much aggregated statements, concerning for example the Northern Atlantic at the LGM: palaeoclimate reconstructions from proxies are often specific to one location at many points in time. But developments in the methodologies are such that (some) climate models can now be used to make statements that are less aggregated. In parallel, developments with proxy reconstructions mean that we can foresee good quality regional multi-proxy reconstructions.

Secondly, there are quite different uncertainties associated with the methods. Climate models are often thought of as deterministic. But this reflects the very great challenges in articulating and studying the many sources of uncertainty, and there has been much progress in recent years, with methodologies such as EMICs and emulation. Proxy reconstructions are thought of as noisy. But the noise is well-understood and regional multi-proxy reconstructions are now becoming available.

It is thus possible to envisage a new type of research, working jointly with both sources of information. Such research will be challenging in many ways. We have sketched some of the possibilities above.

Notes

[1]http://www.uoguelph.ca/~rmckitri/research/WegmanReport.pdf

[2]http://www.ukcip.org.uk under ‘Scenarios Gateway’

[3]http://www.researchpages.net/PQUMP

[4]http://www.arctic.noaa.gov/reportcard/seaice.html

[5]http://pmip2.lsce.ipsl.fr/

[6]http://www.ncdc.noaa.gov/paleo/icecore/greenland/gisp/gisp.html

[7]There is a technical issue manifested as ‘instabilities’ in some models. These can be confused with stochastic variation and are regarded by some as evidence of ‘chaos.’ They can also arise from using starting values that are far from the model equilibrium.

[8]We remark that it is the *joint* uncertainty in the inputs that is important. The marginal uncertainty surrounding some of the inputs is all we are likely to have.

Reference List

Challenor, P. G., Hankin, R. K. S. and Marsh, R. (2006) Towards the Probability of Rapid Climate Change. In Schellnhuber, H. J., Cramer, W., Nakicenovic, N., Wigley, T. and Yohe, G. (eds.) *Avoiding Dangerous Climate Change*. Cambridge University Press, pp. 55–63.

Dessai, S., O'Brien, K. and Hulme, M. (2007) On uncertainty and climate change. *Global Environmental Change* 17(1): 1–3.

Edwards, N. and Marsh, R. (2005) Uncertainties due to transport-parameter sensitivity in an efficient 3-D ocean-climate model. *Climate Dynamics* 24 (4): 415–43, doi: 10.1007/s00382-004-0508-8

Goldstein, M., House, L. and Rougier, J. (2008) Assessing model discrepancy using a multi-model ensemble. Technical Report, University of Bristol. http://www.maths.bris.ac.uk/~mazjcr/#pubWIP

Haslett, J. and Parnell, A. (2008) A simple monotone process with application to radiocarbon-dated depth chronologies. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 57(4): 399–418.

Haslett, J., Whiley, M., Bhattacharya, S., Salter-Townshend, M., Wilson, S. P., Allen, B., Huntley, J. R. M. and Mitchell, F. J. G. (2006) Bayesian palaeoclimate reconstruction. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 169(3): 395–438.

Hill, J. K., Thomas, C. D. and Huntley, B. (2003) Modelling Present and Potential Future Ranges of European Butterflies using Climate Response Surfaces. In Boggs, C. L., Watt, W. B. and Ehrlich, P. R. (eds.) *Butterflies: Ecology and Evolution Taking Flight*. University of Chicago Press: Chicago, pp. 149–67.

Huntley, B. (1993) The use of climate response surfaces to reconstruct palaeoclimate from Quaternary pollen and plant macrofossil data. *Philosophical Transactions of the Royal Society of London*, Series B - Biological Sciences 341: 215–23.

IPCC (2007) *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press.

Kennedy, M. C. and O'Hagan, A. (2002) Bayesian calibration of computer models. *Journal of the Royal Statistical Society*: *Series B (Statistical Methodology)* 63(3): 425–64.

Kennedy, M., Anderson, C., O'Hagan, A., Lomas, M., Woodward, I., Gosling, J. P. and Heinemeyer, A. (2007) Quantifying uncertainty in the biospheric carbon flux for England and Wales. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171(1): 109–35.

McGuffie, K. and Henderson-Sellers, A. (2000) *A Climate Modelling Primer*. Chichester: John Wiley & Sons.

Mann, M. E., Bradley, R. S. and Hughes, M. K. (1998) Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* 392: 779–87 doi:10.1038/33859

Mann, M. E., Zhang, Z., Hughes, M. K., Bradley, R. S., Miller, S. K., Rutherford, S. and Ni, F. (2008). Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *PNAS* 105, 132520–57, doi:10.1073/pnas.0805721105

National Research Council (2006) *Surface Temperature Reconstructions for the Last 2,000 Years.* Washington, DC: The National Academies Press.

O'Hagan A. (2006) Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering and System Safety* 91 (10-11): 1290–300.

Parnell, A., Haslett, J., Allen, J. R. M., Buck, C. E. and Huntley, B. (2008) A flexible approach to assessing synchroneity of past events using Bayesian reconstructions of sedimentation history. *Quaternary Science Reviews*, in press.

Pastres, R. and Ciavatta S. (2005) A comparison between the uncertainties in model parameters and in forcing functions: its application to a 3D water-quality model. *Environmental Modelling and Software* 20 (8): 981–9.

Petoukhov, V., Ganopolski, A., Brovkin, V., Claussen, M., Eliseev, A., Kubatzki, C. and Rahmstorf, S. (2000) CLIMBER-2: a climate system model of intermediate complexity. Part I: model description and performance for present climate. *Climate Dynamics* 16(1): 1–17.

Rougier, J. C. (forthcoming, 2008) Efficient emulators for multivariate deterministic functions. *Journal of Computational and Graphical Statistics*. (Available at http://www.maths.bris.ac.uk/~mazjcr/OPemulator1.pdf)

Salter-Townshend, M. (2008) Fast approximate Bayesian inference in non-parametric multivariate regression. PhD submitted, Trinity College Dublin.

ter Braak, C. J. F. (1995) Non-linear methods for multivariate statistical calibration and their use in palaeoecology: a comparison of inverse (*k*-nearest neighbours, partial least squares and weighted averaging partial least squares) and classical approaches. *Chemometrics and Intelligent Laboratory Systems* 28 (1): 165–80.

Weaver, A. J., Eby, M., Wiebe, E. C., Bitz, C. M., Duffy, P. B., Ewen, T. L., Fanning, A. F., Holland, M. M., MacFadyen, A., Matthews, H. D., Meissner, K. J., Saenko, O., Schmittner, A., Wang, H. X. and Yoshimori, M. (2001) The UVic Earth system climate model: model description, climatology, and applications to past, present and future climates. *Atmosphere-Ocean* 39 (4): 361–428.

Wegman, E. J., Scott, D. W. and Said, Y. H (2007) Ad hoc committee report on the 'hockey stick' global climate reconstruction. http://republicans.energycommerce.house.gov/108/home/07142006_Wegman_Report.pdf

Backlist of Papers Published in Insights

2008 Volume 1

No.	Author	Title	Series
1	Boris Wiseman	Lévi-Strauss, Caduveo Body Painting and the Readymade: Thinking Borderlines	General
2	John Hedley Brooke	Can Scientific Discovery be a Religious Experience?	Darwin's Legacy
3	Bryan R. Cullen	Rapid and Ongoing Darwinian Selection of the Human Genome	Darwin's Legacy
4	Penelope Deutscher	Women, Animality, Immunity – and the Slave of the Slave	Darwin's Legacy
5	Martin Harwit	The Growth of Astrophysical Understanding	Modelling
6	Donald MacKenzie	Making Things the Same: Gases, Emission Rights and the Politics of Carbon Markets	Modelling
7	Lorraine Code	Thinking Ecologically about Biology	Darwin's Legacy
8	Eric Winsberg	A Function for Fictions: Expanding the Scope of Science	Modelling
9	Willard Bohn	Visual Poetry in France after Apollinaire	Modelling
10	Robert A. Skipper Jr	R. A. Fisher and the Origins of Random Drift	Darwin's Legacy
11	Nancy Cartwright	Models: Parables v Fables	Modelling
12	Atholl Anderson	Problems of the 'Traditionalist' Model of Long-Distance Polynesian Voyaging	Modelling

2009 Volume 2

No.	Author	Title	Series
1	Robert A. Walker	Where Species Begin: Structure, Organization and Stability in Biological Membranes and Model Membrane Systems	Darwin's Legacy
2	Michael Pryke	'What is Going On?' Seeking Visual Cues Amongst the Flows of Global Finance	Modelling
3	Ronaldo I. Borja	Landslides and Debris Flow Induced by Rainfall	Modelling
4	Roland Fletcher	Low-Density, Agrarian-Based Urbanism: A Comparitive View	Modelling
5	Paul Ormerod	21st Century Economics	Modelling

No.	Author	Title	Series
6	Peter C. Matthews	Guiding the Engineering Process: Path of Least Resistance versus Creative Fiction	Modelling
7	Bernd Goebel	Anselm's Theory of Universals Reconsidered	Modelling
8	Roger Smith	Locating History in the Human Sciences	Being Human
9	Sonia Kruks	Why Do We Humans Seek Revenge and Should We?	Being Human
10	Mark Turner	Thinking With Feeling	Being Human
11	Christa Davis Acampora	Agonistic Politics and the War on Terror	Being Human
12	Arun Saldanha	So What *Is* Race?	Being Human
13	Daniel Beunza and David Stark	Devices For Doubt: Models and Reflexivity in Merger Arbitage	Modelling
14	Robert Hariman	Democratic Stupidity	Being Human

Insights

Insights is edited by Michael O'Neill, IAS Director and Professor of English.
Correspondence should be directed to Audrey Bowron (a.e.bowron@durham.ac.uk).