Government Office for Science

Foresight

Prediction of hydro-meteorological, meteorological and climatological hazards

Commissioned Review

Foresight, Government Office for Science

Contents

Prediction of hydro-meteorological, meteorological and climatological hazards

Tim Palmer

27 November 2012

Report produced for the Government Office of Science, Foresight project 'Reducing Risks of Future Disasters: Priorities for Decision Makers'

What is the status of hydro-meteorological, meteorological and climatological hazard event forecasting in terms of time, space and severity?

i Hydro-meteorological events: from hours to decades

In this paper, the prediction of severe hydro-meteorological events, with timescales of hours to decades is reviewed. These events include pluvial flooding from intense thunderstorms and fluvial flooding associated with larger scale rain-bearing weather systems, potentially remote in space and time from the flooded region. Such storms and systems can bring additional hazards with them such as tornadoes, lightning and damaging winds. Additionally we consider inundation from the sea arising from storm surges on the one hand, to climate-change induced sea-level rise on the other. We also discuss the opposite problem: drought and related heat waves. Related to these are the key questions: how well placed are we in predicting such hydro-meteorological events, and what are the prospects for significantly improving our predictive capability in the coming two decades?

Meteorological forecasting is often delineated by terms that denote different forecast timescales: nowcasting, short- and medium-range forecasting, intraseasonal and seasonal forecasting, decadal forecasting and so on. These terms are often obscure to the non-specialist. Hence, in discussing forecast timescales, let us focus instead on the specific hydro-meteorological phenomena that underpin the rationale for these various terms.

On the shortest timescale of interest consider the flash flood associated with an intense convective weather system. In meteorology the word “convection” is used to denote a type of vertical motion arising when the vertical profile of temperature is unstable to small perturbations. Such convective instabilities occur in the oceans and atmosphere, but in the atmosphere the strength of these up-down motions can be amplified substantially by the release of latent heat when water vapour, evaporated from the surface, condenses as it is drawn aloft. The thunderstorm is a familiar example of such a convective system. In situations of exceptionally strong convective instability, individual convective cells with scales of a few kilometres are organised dynamically into what are called mesoscale convective weather systems with scales of tens of kilometres. These are associated with exceptional precipitation amounts (often rain, but sometimes hail or even snow), and can be associated with destructive wind. The timescale for the evolution of an individual severe convective weather system is typically measured in hours, and this also determines the timescale over which these individual convective weather systems can be predicted.

Because of the fundamental nonlinearity of the underlying equations of motion representing the laws of physics and thermodynamics, weather systems do not always fall neatly into scale-dependent categories. However, on scales of hundreds to thousands of kilometres there is a large class of “synoptic scale” weather system. The dynamics of these weather systems is strongly constrained by the daily rate of rotation of the earth about its axis. The familiar extratropical low-pressure weather systems certainly lie in this class, but so too do the tropical cyclones (also known as hurricanes in the Atlantic basin and typhoons in the Indian and western Pacific Ocean regions). For the more intense extratropical weather systems, and for all tropical cyclones, the role of latent heat release in the embedded cloud systems is critical for maintaining the often-destructive character of these systems, both in terms of wind strength and precipitation intensity. It is well known, for example, that tropical cyclones lose some of their intensity as they track over land and are shut off from their source of oceanic moisture. At the same time, tropical cyclones that move northward undergo what is called mid-latitude transition where the modified lower latitude phenomenon can inflict severe weather on Europe. As mentioned, the timescales for development and evolution of these synoptic-scale weather systems can be measured in timescales of days rather than hours.

Of course, when considering the hydrological effects of severe weather systems on coastal regions, an important aspect is inundation associated with wind-driven storm surges. One only has to think of the impact of hurricane Katrina on New Orleans, or the loss of life in Myanmar following Tropical Cyclone Nargis to realise the seriousness of this for society.

Up to now we have been discussing individual weather systems. However, additionally there are atmospheric disturbances with scales of tens of thousands of kilometres (meteorologists call these “planetary scale”) that can modulate the strength of individual weather systems. For the purposes of this paper, the most important of these is the so-called Madden-Julian Oscillation (MJO). The MJO can be described as a planetary-scale wave packet which propagates from west to east around the tropics with a periodicity between 30 and 60 days. Associated with the active phase of this wavelike structure, mesoscale convective systems, bringing intense rainfall, are more likely to develop. That is to say, the MJO acts as a sort of “envelope” disturbance that either encourages or suppresses individual weather systems. This east-west MJO is not the only type of envelope disturbance. In the Asian monsoon region, there are active north-south propagating envelope disturbances defining what are known as “active” and “break” phases of the monsoon with the former inflicting periods of intense rainfall and flooding and the latter extended periods of drought.

Although the MJO is often described as a 30-60 day oscillation, it is rarely coherent over such timescales. Rather it is a more intermittent phenomenon, typically developing over the Indian Ocean and weakening over the Eastern Pacific. Its characteristic predictability timescale can be considered to be several weeks.

Although the phenomena discussed above are primarily atmospheric, their dynamics depend critically on the oceans through the atmosphere responding to the differences in sea-surface temperature from one region to another and by the addition of water vapour to the atmosphere by evaporation. As we approach the monthly timescale, two-way interactions between the ocean and atmosphere become more and more important in describing the dynamics of these phenomena. For example, the propagation of the MJO cannot be simulated accurately without such two-way interactions.

However, two-way ocean-atmosphere interactions are a *sine qua non* for our next geophysical phenomenon, the El Niño event, associated with a weakening of the westward trade winds across the Pacific Ocean which reduces the upwelling of subsurface cold water and provides an essential warming across the entire Pacific Ocean. Weather patterns across the Pacific basin and surrounding regions are profoundly changed during an El Niño event (or its opposite “La Niña” phase where the trade winds are enhanced rather than weakened and induce stronger than normal upwelling of colder subsurface water in the eastern tropical Pacific Ocean). For example, as the waters of the eastern and central Pacific warm during an El Niño event, so too do the thermally-driven convective weather systems. The associated changes in spatial patterns of latent heat release then have profound impacts on planetary scale circulations in the atmosphere. Hence, for example, the Indian monsoon is often especially active during a La Niña event, as is hurricane activity in the Caribbean. During an El Niño event, on the other hand, hurricanes occur preferentially in the eastern Atlantic Ocean and the Indian summer monsoon rains are usually below average, with an enhanced possibility of drought.

El Niño events typically develop on timescales of a season or two, and the possibility of predicting these events and their global consequences has spawned considerable developments in the field of “seasonal prediction”, ie prediction on timescales of three to six months.

The oceans exhibit natural variations on timescales longer than seasons and these also have consequences for rainfall in surrounding land regions. For example, it is now understood that the droughts of the African Sahel in the 1970s and 1980s were linked to multi-decadal fluctuations in sea surface temperatures, primarily in the tropical Atlantic basin. Variations in these sea surface temperatures are in turn associated with what are known as “overturning” circulations in the oceans. The most well-known of these is known as the planetary-scale “thermohaline” circulation – its strength being controlled by density fluctuations associated with both thermal and salinity effects. The thermohaline circulation encompasses most of the Atlantic basin, and in turn modulates the shallower overturning circulations in the tropical Atlantic that affect Sahel rainfall. Just as the existence of El Niño has driven the development of seasonal forecast systems, so the existence of the thermohaline circulation has driven the development of decadal-timescale prediction studies.

However, there is another independent reason for developing a decadal timescale forecasting system. Year on year, the atmospheric concentration of carbon dioxide is increasing due to burning of fossil fuels. It is well known that this will lead to an increase in global mean temperature, but increasing greenhouse gas concentrations may also be expected to lead to changes in regional weather patterns, and indeed to changes in oceanic patterns of variability, for example associated with El Niño or the thermohaline circulation. The IPCC Special Report on Climate Extremes (SREX, 2012) has addressed in detail the impacts of climate change on weather extremes. For the purposes of providing scientific input into global policy discussions to reduce carbon emissions, the focus has been on estimating the changes to climate in one hundred years. However, if these attempts fail or are only partially successful, humanity must learn how to adapt to these regional changes in climate: indeed in the coming decades humanity must learn to adapt to the effects of changes already set in stone by historical emissions. For this, predictions of changes in climate on timescales shorter than the century timescale are needed. Hence the decadal and multi-decadal prediction problem has taken on a new sense of importance – as providing information on near term changes in climate that are partly naturally occurring and partly man made. These predictions must include hydrological information: are intense tropical cyclones likely to be more or less likely, are rivers more or less likely to burst their banks, will long-term drought be more or less commonplace?

Just as inundation from the sea associated with storm-surge activity can lead to devastation, and in some regions this might increase with climate change, so climate change will also lead to global sea level rise as sea temperatures warm and ice sheets melt. Countries like Bangladesh could be hit by a combination of these two effects.

A review of weather and climate prediction across this range of timescales has been given by Hoskins (2012).

ii How are predictions made?

Man has been trying to forecast weather since time immemorial. Until the advent of the digital computer, attempts at forecasting were based on what can be called “empirical models” – using past data and the sequences of past events to produce analogues, these models encode some empirical functional relationship between quantities one wants to predict and quantities one can observe ahead of time. Although for weather prediction such attempts can now considered to be utterly archaic and of historical interest only, on seasonal and longer timescales such empirical predictions are still used in some quarters. However, for the purposes of this paper and for all space and time scales of interest, we only consider predictive models based on the primitive equations of physics: Newton’s laws of motion, the laws of thermodynamics and, where appropriate, the quantum theory of atomic spectra (which determine how photons from the sun are absorbed, scattered and re-emitted by the atmosphere).

In terms of mathematics, these various laws are described by nonlinear partial differential equations (PDEs). They describe a continuum of scales from the planetary scales down to microscopic viscous scales. There is no way - in the conceivable future and certainly not on the timescale covered by this Foresight Report – that supercomputers will be powerful enough to solve these equations across this vast range of scales. In practice this means that the underlying PDEs have to be approximated by projecting them onto some finite grid. Processes occurring on scales less than the grid scale (also known as the truncation scale) have to be approximated in some way - a procedure sometimes called “parametrisation”. Hence a typical seasonal forecast model may have a truncation scale of around 100km, whilst a global weather forecast model may have a truncation scale of around 10km. For both of such models, cloud systems have to be parametrised. For very short-range forecasting (less than one day), a finer scale model can be nested inside the global weather prediction model. For this finer scale model, convective cloud systems are simulated explicitly, whilst sub-cloud turbulence and other small scale processes are parametrised.

Some theoretical points need to be made at this stage. Firstly, there is no good theoretical reason for truncating these global models at 100km or 10km. Rather the choice of truncation scale is determined by available supercomputing capability. Indeed, from a theoretical point of view, it would seem desirable for a global weather or climate model to be able to simulate explicitly the phenomenon of atmospheric convective instability (not only for predicting severe weather, but more generally for simulating the energy-momentum budget of the climate system accurately). However, this would mean that global models should have truncation scales of 1km or less, and at present this is not practicable as discussed below. This means that a key source of forecast uncertainty arises from the way we represent key processes by parametrisation. In truth, parametrisations are in part quasi-empirical formulae. This is a key point for providing reliable forecasts to users, to which we shall return later.

Mathematically speaking, weather forecasting is an initial-value problem – ie given a set of initial conditions for the atmosphere, ocean and land surface, and a set of prognostic equations, estimate the future state of the system. Initial values are obtained from the array of observations that are made each day of the atmosphere, ocean and land surface through in situ measurement or remote measurement from satellite or radar and lidar. However, converting this multitude of observations (made at different locations and at different times) into a set of initial conditions on the model grid, is a highly complex mathematical procedure known as Data Assimilation. For example, a satellite radiometer might observe pixel-scale atmospheric radiances in the wings of a spectral infra-red absorption band of carbon dioxide. Data Assimilation “inverts” these radiance measurements to give an estimate of atmospheric temperature on the 3D atmospheric grid. Such inversions are as computationally expensive as the weather forecast itself, and therefore are major cost drivers for supercomputing.

Of course, for predicting river discharge, or storm surge, then output of the weather forecast model has to be coupled with a hydrological model for the relevant catchment basin, or with a ocean storm surge model than can predict the sea level rise associated with wind strength.

A discussion of the “seamless” or “unified” nature of weather and climate prediction can be found in Hurrell *et al*, (2009) and Shapiro *et al*, (2010).

iii Probability forecasting

Traditionally, weather forecasts have been deterministic in character. However, such forecasts are unreliable. Our understanding of why weather forecasts can go wrong, and what to do about it, has built on the back of chaos theory. A chaotic system is, by definition, one whose evolution is generically sensitive to uncertainties in initial conditions. The initial conditions for weather forecasts are inevitably uncertain, not only due to imprecision in the basic observations, but also due their limited coverage (the infra-red satellite sounder of the last paragraph cannot see through cloud) but also due truncation errors in the model in which the observations are assimilated.

Superficially, it might be thought that weather forecasting is a hopeless task in the light of chaos. However, a couple of key points need to be made. Firstly the growth of initial errors is scale dependent. Uncertainties in the initial values for a thunderstorm may double in an hour, whilst uncertainties in the initial values for a synoptic scale weather system may double in a couple of days. For planetary scales (e.g. associated with the MJO) the doubling time is longer still. Secondly, the growth of errors is state dependent (technically, the local Lyapunov exponents of a chaotic system vary according to position on the system's attractor). This means that on some occasions it may be possible to make useful predictions on timescales much longer than one would on average expect to be possible, whilst on other occasions, the range of useful skill would be extremely curtailed due to rapid error growth.

A single-shot deterministic forecast system (sometimes right and sometimes wrong) is unreliable. How would a user know whether he or she could depend on today's forecast or a farmer on how much rainfall may occur in the next week or two? This was more or less the state of weather forecasting before the advent of ensemble prediction, and, not surprisingly, users were not confident in being able to use weather forecast data with any degree of reliability.

All this has been changed by ensemble prediction. The key idea is that at each initial time, we can now run the forecast system many times (say 50 or 100 times) varying the key uncertain aspects of the initial conditions and the model equations. On the occasions where the atmosphere system is in a relatively predictable state, the individual members of the forecast system track one another closely. On the occasions where the atmosphere system is in a relatively unpredictable state, the individual members of the forecast system will diverge rapidly.

The way to synthesise this information is through the notion of probability. For example, suppose that 20 out of 100 ensemble members predict rainfall over Dhakar, in excess of 100mm/day at day 5 of the forecast initialised on July 1st. Then, by elementary frequentism, one would estimate the probability that rainfall will exceed 100mm/day over Dhakar on July 6th, to be 20%.

Reliable probabilistic predictions will necessarily lead to better decision making than could be possible with unreliable single-shot, yes-no, best-guess deterministic forecasts. To be more specific, suppose it is possible for a decision maker to take some sort of protective action, at cost C, to mitigate the damage caused by some severe weather event. If the severe weather event occurs, assume the benefit in taking protective action in terms of reduced weather-inflicted damage, is equal to B. If B exceeds C, a key question for a decision maker is when to take protective action? With a single-shot deterministic prediction (the severe weather event either will or will not occur), then the only rational decision is to take protective action when the weather event is predicted. However, if the deterministic forecast is wrong then either the protective action will have been a waste of money (“false positive”), or unmitigated damage will have occurred (“false negative”). However, with a reliable probabilistic prediction, the rational decision is significantly different: it is to take protective action when the forecast probability of the event exceeds C/B. If C is close to B, protective action should be taken only if it is almost certain that the event will occur. Conversely, if C is much smaller than B, it makes sense to take protective action even if the forecast probability of the event is relatively small. It is easy to show that the economic benefits are greater of taking decisions using the probabilistic rather than deterministic strategy.

Of course, this analysis is highly idealised. In practice, the values C and B are not well known (especially when human lives are at risk – e.g. when deciding whether to evacuate a town or city in the path of a possible tropical cyclone). Moreover decisions are not always of a simple binary type (take action or don’t take action); there may be a continuum of types of protective action that could be taken (for example evacuate the upper N percentile of those at risk).

Nevertheless, the simple analysis demonstrates the value of probabilistic forecasts, providing the forecast probabilities are themselves reliable.

What does it mean to say that a probabilistic forecast system is reliable? This is a crucial concept and we will discuss it below in some detail. Suppose an ensemble forecast system predicts some weather event (e.g. occurrence of rainfall exceeding 100mm/day over Dhakar) with a probability of 100%. Then if the system is reliable, we would expect the weather event to definitely occur in reality. Conversely, if an ensemble system predicts the same event to have a probability of 0%, we would expect the weather to definitely not occur. However, what if the probability of rainfall exceeding 100mm/day over Dhakar was 60% - how do we know if that is a reliable prediction or not? After all, both the occurrence and the non-occurrence of the rainfall event are consistent with a forecast probability of 60%!

The first point to make is that we can’t judge the reliability of a probability forecast from just one forecast event. However, we can if we have a large enough statistical sample of forecast events. For example, if we have a large statistical sample of probability forecasts of rainfall for Dhakar and we select from this a subsample where the probability of rainfall exceeding 100mm/day was 60%. Then, if the forecast probabilities are reliable, we would expect it to actually rain at least 100mm/day on 60% of occasions in this subsample. More generally, from the occasions a reliable ensemble forecast system predicts some weather event with a probability p, then the event should actually occur on a faction p of occasions.

However, a key question to ask is whether current operational ensemble forecast systems do produce perfectly reliable probabilities? For predicting synoptic-scale weather systems on timescales of days, probability forecasts are known to be very reliable. For predicting intense convective storms on very short range timescales of a few hours, it is currently not known the extent to which prediction systems are reliable or not – high resolution limited-area model ensemble prediction systems are only now being constructed for these short timescales. For longer range predictions eg on seasonal timescales, it is known that probabilistic ensemble predictions of regional rainfall are not fully reliable. On these timescales, climate models develop significant biases against observations – these biases can be as large as the climatic signal one is wishing to predict. The existence of such biases is indicative of errors in the models themselves.

It therefore becomes crucial for reliable probability forecasting, that there is some representation in the ensemble, not only of observation uncertainty, but also model uncertainty. We do not yet have adequate representations of model uncertainty in our seasonal prediction systems. The traditional “multi-model ensemble” methodology appears to be not good enough. Using several quasi-independent climate models with different parametrisations, instead of just one, multi-model ensembles provide a pragmatic approach to the representation of model uncertainty. Although multi-model ensembles do provide more reliable probability forecasts on the seasonal timescale than single-model ensembles, they suffer from systemic failure if all component models have the same class of error. A more theoretically inspired approach is to formulate the parametrisation problem stochastically, treating the sub-grid processes in terms of probability distributions (constrained by the grid-scale variables). However, research in this area is only just beginning and is one that can be expected to blossom in the coming decades.

The existence of overt model deficiencies (either in the weather or climate prediction model or in the hydrological impact models) can to some extent be ameliorated by empirical calibration. For example, if a model has a persistent bias towards wetter than observed conditions, a simple empirical bias correction can be made which brings the forecasts closer to the observations on average. Similarly, forecast probabilities for some weather event E can be calibrated to bring the probabilities closer to the climatological probability of E.

The problem of how to correct for model bias is one of the key issues for increasing the impact of hydrometeorological predictions in the coming years and will be discussed further below.

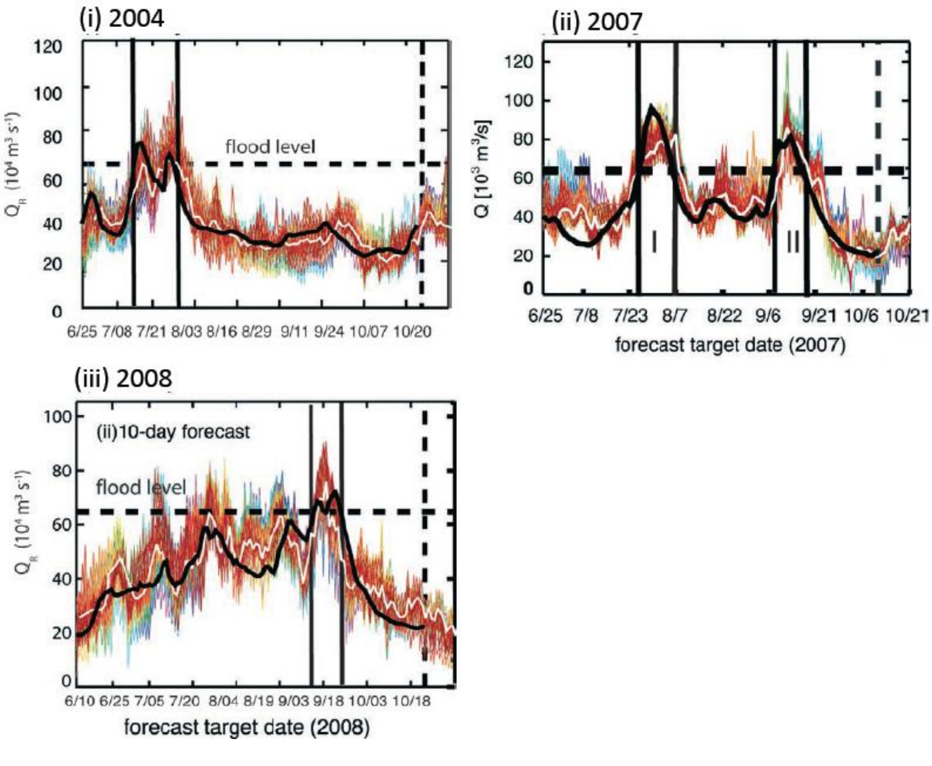
A recent review of issues related to predictability of weather and climate are given in Slingo and Palmer (2011) and Palmer (2012).

iv Some examples

There are a number of examples of existing probabilistic forecasting applied directly to hazards. Three examples are given: flood forecasting in Bangladesh, flooding in Pakistan and the North Indian Ocean. In addition, it is anticipated that major insights into the forecasting of drought may be gained in the relatively near future.

(a) Probability of flood occurrence

Figure 1: 10-day forecasts of the Brahmaputra River discharge into Bangladesh for (i) 2004, (ii) 2007 and (iii) 2008. The coloured swath of lines denote the 51 realizations made using the ECMWF allowing the determination of the probabilities of river flow 10-days in the future. The horizontal dashed lines show the flood level at the entrance point of the river into Bangladesh. The solid bold line shows the observed river level. The vertical lines indicate the duration of the four flood periods. The mean of the ensembles is shown in white. In each case, the forecasting system indicated extremely high probability of floods 10 days in advance. As well as the reduction of loss of life, preparatory actions allowed savings measured in multiples of annual incomes. From Hopson and Webster (2010) and Webster *et al*. (2011)



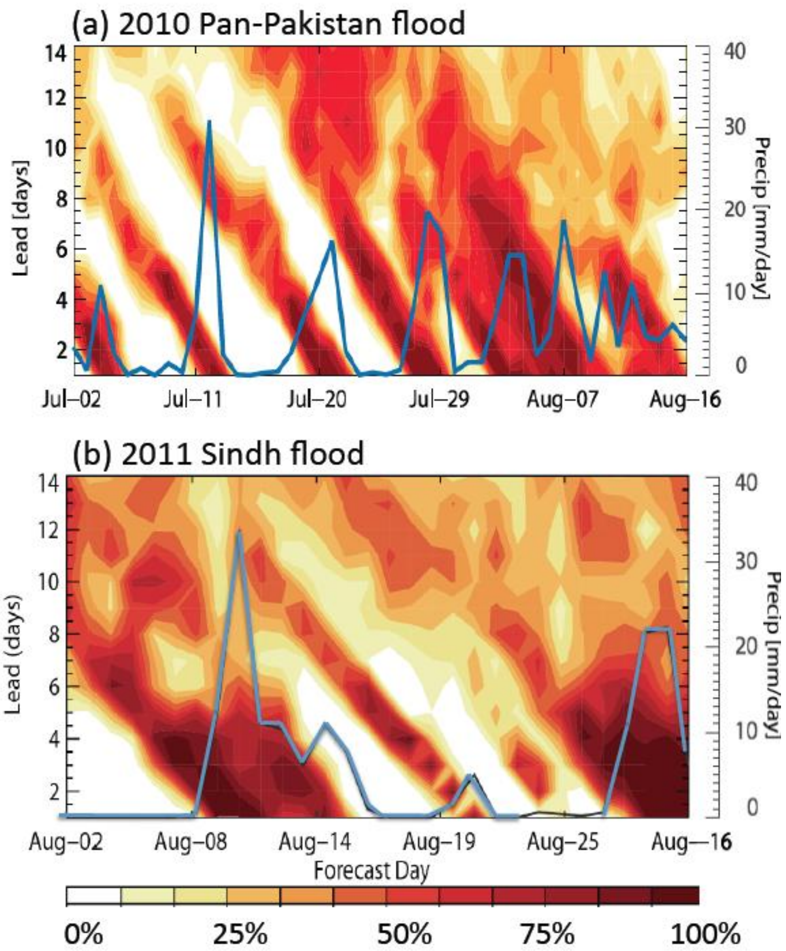
Major flooding occurred in Bangladesh during the summer of 1998. The flooding arrived unannounced and inundated 60% of the country for over three months. The impacts were devastating and the loss of life and property catastrophic. A major effort by international agencies was initiated to produce long lead-time forecasts of Bangladesh floods. The forecasts needed to be probabilistic so that the risk could be computed. Each year there are floods but only a few have the vigour and the longevity of the 1998 floods. Two hurdles had to be navigated: there was no upstream data available from India so that the meteorological forecasts had to be used to “synthesize” upstream river flow. Second, rainfall forecasts possess significant bias and it was necessary to remove this bias. A nonlinear “quantile-to-quantile” statistical technique based on almost three decades of satellite data rendered the forecasts statistically consistent with the observations. Figure 1, from Hopson and Webster (2010) and Webster *et al*. (2011) describe probabilistic flood forecasts made using ECMWF ensemble forecasts. The forecasts were made with an absence of upstream data with the ECMWF rainfall forecasts, in conjunction with a hydrological model, computing the upstream flow.

The Bangladesh system became operational in 2006 and major Brahmaputra floods were forecast 10 days in advance in 2007 and 2008. Ahead of the floods, communities in 6 unions (counties) were trained in what to do with crops, farm animals and personal property in addition to the storing of fresh water and food for the period of inundation following the floods. An economic analysis showed that the savings to farms and households were measurable in units of annual income. Warnings of the probability of tropical cyclone formation, track and landfall are also produced routinely with lead-times of 10 days or so (see also SREX, 2012).

(b) Probability of extreme rainfall events

The system described in Figure 1 is transferable to other regions of the developing world. Figure 2 provides two examples of probabilistic forecasts of extreme rainfall events that occurred in Pakistan in 2010 (panel a) and 2011 (panel b) (Webster *et al*. 2011). Both flooding events had catastrophic consequences. The figures show the probability of rainfall occurring greater than one standard deviation greater than climatology. Both diagrams show that very heavy rainfall was predicted 8-10 days in advance. If these forecasts had been available and coupled to an advanced hydrological model, the resulting floods would have been forecast over a week in advance and catastrophic losses of life and property could have been reduced.

Figure 2: (a) Forecast lead time diagram of the probability that the predicted ECMWF rainfall forecast (adjusted using the q-to-q statistical technique) for the red region exceeds the observed July-August climatology plus 1 standard deviation. The blue line represents the satellite observed rainfall [mm/day] averaged for the same region and the same time period. The July 12 and June 21 events were forecast at 50% probability level to exceed observed climatology plus 1 standard deviation, at least 10 days in advance. An additional 2 days of predictability is evident at the 50% level for the July 28 event. All events were forecast at >70–80% level of probability 6 days in advance. (b) Same as (a) but for precipitation forecasts leading the major flooding of 2011 that occurred in the Sindh region. From Webster *et al* (2011).

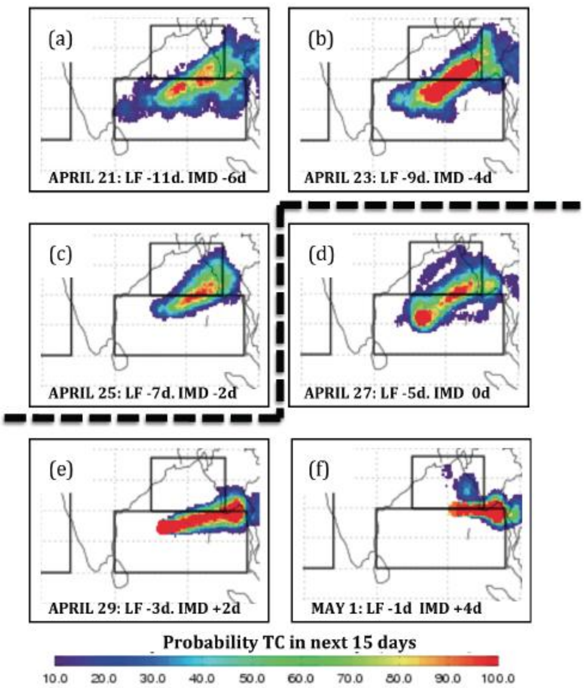


In 2010 and 2011 Pakistan did not have a system to produce flood forecasts nor access to the ECMWF system's forecast system that would translate the rendered ECMWF precipitation forecasts to flood forecasts. Nor did they have a disaster system that could have communicated the warning effectively. Yet, there is ample evidence the development of such a system (e.g., Fig. 2) is quite possible.

(c) Extended prediction of tropical cyclone genesis probability

The landfall of a tropical cyclone in any part of the world usually results in destruction and the loss of life. In the United States, Katrina in 2005 had devastating impacts. In South Asia, the impacts are often calamitous because of two factors: the extensive deltas and the large populations that reside within them. Recently, Tropical Cyclone Nargis killed over 130,000 Mynamar citizens as a storm surge inundated the Irrawaddy delta. Yet, a high probability of tropical cyclone development and subsequent movement over Myanmar was forecast over 10 days in advance (Belanger *et al*, 2012). Figure 3 shows the sequence of forecasts of Nargis. The loss of life resulted mostly from government inaction (Webster 2008).

Figure 3: Sequence of forecasts of Tropical Cyclone Nargis. The colour code swath indicates the probability of the location of a tropical cyclone with the region during the next 15 days. Panels (a)-(c) show tropical cyclone forecasts ahead of the observed development of the storm. Note that 11 days before landfall there was a high probability of the cyclone passing through Myanmar. Panels (d)-(f) show forecasts following genesis. Forecasts from Belanger *et al*. (2012) If the forecasts had been heeded, disaster managers would have had ample time to arrange evacuation (Webster 2008).



(d) Probability of droughts and water resource management

Drought impacts vast areas and large populations for extended periods of time (SREX, 2012). Advanced warning of drought allows water resource management to alleviate the drought, the planting drought resistant crops and planning for the migration of herd animals. There are three types of drought (meteorological, hydrological and agricultural) and the impact of each can be lessened by probabilistic seasonal forecasts.

Meteorological drought involves a reduction in regional rainfall over a specific period (day, month, season and year) defined formally as some percentage of the long-term climatological average. Such droughts are often associated with extremes of temperature that lead to an increase of evapotranspiration and consequent serious loss of soil moisture. A meteorological drought is especially devastating for regions that depend principally on precipitation for agriculture. Such rain-fed areas constitute 70% of the less developed world. A hydrological drought is defined by a reduction in water resources (for example, streamflow, reservoir levels, ground water, sub-surface aquifer levels) within a region. A hydrological drought may result from a meteorological drought elsewhere. For example, a hydrological drought may exist in the southern Indus Valley due to reduced winter snowfall in the Himalaya or poor monsoon rainfall during the summer. River flow, so reduced, may limit water for irrigation. This form of drought is the compound consequence of hydrological and meteorological droughts. About 30 percent of the total area of Pakistan is cultivated. Of this cultivated land, over 75 per cent is irrigated, the rest being rain-fed. The irrigated land may avoid a regional meteorological drought if there is no hydrological drought and water is supplied from regions receiving more abundant rainfall.

To date, most hydrological forecasting has been aimed at relatively short duration precipitation events resulting in flooding. But an ability to forecast regional rainfall presents the opportunity to forecast the three different types of droughts. Clearly there is the opportunity to optimize water resource management through the provision of probabilities of rainfall and streamflow.

What are the most interesting/important areas of scientific progress in the next 20 years?

The key elements needed to produce more reliable forecasts can be summarised as:

a) raw observations (atmosphere, ocean, land-surface, cyrosphere) to initialise and evaluate the forecasts
b) techniques to assimilate the observations into the forecast models
c) the numerical forecast models themselves
d) techniques to quantify the inherent uncertainty in the forecasts
e) empirical calibration models
f) linkage between the meteorological forecast variables (rain, temperature, wind strength) to user-relevant impact variables (malaria incidence, crop yield, river discharge, property damage)

It is beyond the scope of this review paper to discuss the development of the observational network, suffice to say that development of the network is pivotal to the improvement of forecast capability. On timescales of hours and days, many of the key observational platforms are space based, and the space agencies play a key role in the maintenance and development of the observational network. Ground based measurements, important in their own right, are also key to calibrate the space-based measurement systems.

On longer timescales, in situ oceanic observations are paramount. Because of the types of model deficiencies discussed above it is sometimes hard to show the value of the considerable investments that have been made in ocean observations. The key lies in improving the models – see below.

As discussed above, the assimilation of observations into the forecast models is a complex numerical process. Data assimilation is an active topic of research. Traditionally, data assimilation schemes have been designed to find the most likely state of the atmosphere or oceans, given a set of raw observations. Increasingly, data assimilation schemes are being designed to determine a full probability distribution of states, consistent with the raw observations and their uncertainty. Secondly, whilst data assimilation has been traditionally performed separately in the atmosphere and oceans; increasingly data assimilation schemes are being developed for the combined coupled atmosphere/ocean system.

The development of the forecast models continues in many weather and climate institutes around the world. Research can be broadly divided into efforts to improve the numerical algorithms to integrate the partial differential equations of climate, and work to develop the sub-grid scale parametrisation schemes (these are sometimes delineated by the words “numerics” and “physics”). As discussed above, research is beginning to incorporate the inherent uncertainty associated with parametrisation into the numerical equations, through stochastic (rather than deterministic) parametrisations. Since this work is still relatively nascent, it remains important to calibrate model output empirically before inputting into impact models, and new types of calibration schemes are continually being developed.

What would it take to provide a step change in forecasting capability and how big an impact could this change have in reducing impacts?

Leaving aside detailed theoretical developments, one could say that there are four key ingredients for improving forecasts: i) human resources needed to develop the forecasting systems, ii) observations needed to initialise and validate the forecasts, iii) supercomputers needed to integrate the equations forward in time, iv) strategies to communicate forecasts to the public in general and, especially, those in peril.

As discussed above, since weather and seasonal-to-decadal climate prediction is essentially an initial value problem, the role of observations is central. However, many weather and climate institutes struggle to find sufficient human resources to develop the models needed to integrate these initial conditions forward, especially since these models increasingly include representations of the chemistry and biology of the Earth System, as well as basic physics.

However, perhaps the single most important bottleneck to progress is lack of supercomputing capability. It is important to recognise that supercomputing is not only needed to provide the basic forecasts, it is needed to do the research that leads to the development of the forecast systems. Included in this research is the testing of candidate ensemble forecast systems on past events. In the case of seasonal forecasting, such testing must go back over the last 30 years or more (otherwise sufficient statistics will not be available to make assessments of reliability). This requires enormous amounts of supercomputing. A key question for the development of models is that of resolution – what is an ideal grid spacing for a weather or climate model?

There is no doubt that in numerical weather prediction, forecast skill has been improved through the use of higher resolution models. This improvement arose not only through more accurate integration schemes, but also because the raw observations could be assimilated more accurately into models with higher resolution.

A key aspiration for weather and climate prediction is to be able to resolve the key convective instabilities that on the one hand are key to hydro-meteorological prediction, but more generally are central to the overall energy-momentum budget of the climate system. There is considerable evidence, not only that large-scale aspects of climate simulations can be extremely sensitive to the details of convective parametrisation closure, but that all such closures may be deficient in certain ways (eg in failing to account for the backscatter of kinetic energy associated with the release of convective instability, onto resolved scales). It therefore becomes a matter of urgency that we can assess the accuracy of climate simulations with explicit convection, suggesting a requirement of a minimum of around 1 km grid. That is to say, such resolutions may not only be needed for ensemble-based short range weather forecasts of individual storms, but also for longer-range seasonal, decadal or indeed centennial climate forecasts.

This means that reliable weather and climate predictions may require supercomputers with performance in the range hundreds of petaflops to exaflops- ie 10^{17}- 10^{18} (floating point) operations per second. It is unlikely that exaflop computers will be constructed until late in the decade.

Of course, it is not a matter of “all or nothing” as far as resolution is concerned. Current numerical weather prediction and climate models have parametrised representation of convective cloud systems, and there is considerable scope for improving model resolution, whilst continuing with such parametrised cloud systems. For example, the Earth’s topography, including the land-sea boundary is clearly represented more accurately on a 10km grid compared with the 100km grids currently used for seasonal or decadal prediction. Also, some of the nonlinear effects of extratropical weather systems (the so-called enstrophy cascade, believed to be important in maintaining blocking anticyclones against radiative relaxation) can be better represented with model truncation scales around 10km or so.

One might also ask whether there is additional value in weather and climate institutes pooling human resources more than is currently done (Palmer, 2012). This is a more controversial area. Some would argue that we need a number of quasi-independent climate models to provide an estimate of model uncertainty (cf the multi-model ensemble concept discussed above). On the other hand, the development of stochastic parametrisation schemes, and the fact that multi-model ensembles may suffer from systemic failures, may vitiate this argument somewhat. Others may argue that having multiple quasi-independent model development teams fosters a sense of competition, which in turn will stimulate creativity. The counter argument is that in so doing, none of development teams has a critical mass of scientists needed to make optimal progress. At the least, the arguments need to be aired more openly.

What are the barriers to realising these scientific developments?

One key aspect which could make a step change to the quality of our hydro-meteorological forecasts – improving model resolution – has been discussed. We have made a *prime facie* case that reliable hydro-meteorological forecasting particularly for the seasonal and longer timescale, may be an exascale problem. Exascale computers are unlikely to be on the market until the end of the decade. However, in order to be prepared for their arrival, weather and climate forecast centres need to be planning now. Consistent with earlier studies by Shukla *et al* (2010) – see also Palmer (2011) - the Beddington report on the Met Office Hadley Centre provides a potential route forward for European Countries in the coming years: “It will be important to actively engage with European stakeholders to facilitate and pursue opportunities for the future provision of European supercomputing infrastructures.” (Beddington, 2010)

On the other hand, there are barriers to realising these scientific developments. For example, where would the funding for such European supercomputing infrastructure come from? Some might fear that funding would in part come from the budgets that support existing national weather and climate centres.

There may also be barriers to the notion of pooling human resources to create a smaller number of more advanced weather and climate models. Weather and climate institutes have traditionally been built up as part of national capability (in the US around Agency capability) and there accrues a certain amount of pride and prestige in having a world-leading capability in this area. World leading institutes (and hence those with the most influence) may view the notion of resource pooling as effectively relinquishing their position on the league table.

There is a final problem, perhaps less technical and more prosaic, but certainly major. It involves the communication of forecasts to the public in general and to those at risk in particular. In the more developed world, in close proximity to the major forecasting centres, communication can be rapid allowing the public to respond in a rapid fashion to hazard warnings. Yet, in the developing world, for example in Bangladesh, forecasts have a greater difficulty in reaching those in peril for two reasons: Communication is difficult logistically and many nations do not have access to extended forecasts generated in the more-developed world. Furthermore, the constraints on hazard forecasting for the developing world are more exacting that in the more developed world. See also SREX, 2012. For example, even though

Katrina was a major disaster, a large fraction of the population of southern Louisiana/New Orleans was able to evacuate using the highway infrastructure that exists in the United States. Thus, a 3-day warning of impact was sufficient to mitigate some of the potential damage caused by Kartina. But in the developing world, the time needed for evacuation is measured in terms of the slowest members of society, the walking pace of citizens and their cattle. A 3-day forecast is not sufficient and a forecast of 7-8 days is necessary.

Most developing world countries do not have the capacity to build forecasting systems that have been developed eg in Europe and the United States. A clear strategy would be to develop world-wide hazard prediction systems emanating out of the major centres providing extended forecasts to less-developed nations for national or regional dissemination. This strategy could have far-reaching implications for both the well-being of the developing nations and the global community as well. The developing nations encompass the largest populations on the planet and also the most rapidly growing. An unanticipated hazard can have extremely severe ramifications stretching the resources of a region and the humanitarian assistance from the more developed world. As we have indicated with the examples of flood forecasting in Bangladesh, in-country mitigation, resulting from probabilistic forecasts can allay the disastrous impacts of hazards. As we move forward through this century, the need for hazard anticipation in the developed world becomes even more critical. Hazards are anticipated to become more severe and possible more frequent, impacting an even greater proportion of the world’s population (SREX, 2012).

References

Beddington, J. (2010) *Review of climate science advice to Government and Met Office Hadley Centre role, governance and resourcing*. Government Office for Science.

Belanger, J.I, P J. Webster, J A. Curry, And M. T. Jelinek (2012) *Extended Prediction of North Indian Ocean Tropical Cyclones. Weather and Forecasting*, (in press)

Hopson, T. M., P. J. Webster (2010) *A 1-10 day ensemble forecasting scheme for the major river basins of Bangladesh: forecasting severe floods of 2003-2007*. J. Hydromet. Volume: **11** Issue: **3** Pages: 618-641 DOI: 10.1175/2009JHM1006.1

Hoskins, B.J. (2012) *Predictability beyond the deterministic limit*. Quart. J. Roy. Meteorol. Soc., In press.

Hurrell, J., G. Meehl, D.Bader, T.Delworth, B.Kirtman and B. Wielicki (2009) *A unified approach to climate system prediction*. Bull Amer Meteor Soc., 90, 1819-1832.

Jian, J., P.J. Webster and C.D. Hoyos (2009) *Large-scale controls on Ganges and Brahmaputra river discharge on intraseasonal and seasonal time-scales*. Q. J. R. Meteorol. Soc., 135: 353-370

Palmer, T.N. (2011) *A CERN for climate change*. Physics World, 24, 14-15.

Palmer, T.N. (2012) *Towards the probabilistic earth-system simulator: a vision for the future of weather and climate prediction*. Quart. J. Roy. Meteorol. Soc., DOI:10.1002/qj.1923

Slingo, J. and T.N.Palmer (2011) *Uncertainty in weather and climate prediction*. Phil. Trans. R. Soc. A, 369, 4751-4767.

Shukla, J., T.N.Palmer, R. Hagedorn, B. Hoskins, J. L. Kinter III, J. Marotzke, M.Miller and J. Slingo (2010) *Towards a new generation of world climate research and computing facilities*. Bull Amer Meteor Soc., 91, 1377-1388.

Shapiro M., and co-authors (2010) *An Earth-System Prediction Initiative for the 21st Century*. Bull. American Meteorological Society, 91, 1377-1388. DOI:10.1175/2010BAMS2944.1

SREX (2012) *Managing the risks of extreme events and disasters to advance climate change adaptation*. Special report of the Intergovernmental Panel on Climate Change.

Webster, P. J. (2008) *Myanmar's deadly daffodil*. Nature Geoscience doi: 10.1038/ngeo257

Webster P. J., V. E. Toma and H-. M Kim (2011) *Were the 2010 Pakistan Floods predictable?* Geophys. Res. Lettrs.,38, L04806, doi: 10.1029/2010GL046346

Webster P. J. and J. Jian (2011) *Probability, Uncertainty and prediction: A pathway towards the alleviation of poverty in the developing world*. Phil. Trans. Roy. Soc. vol.369 no. 1956

Webster, P.J., J. Jian, T. M. Hopson, C. D. Hoyos, P. Agudelo, H-R. Chang, J. A. Curry, R. L. Grossman, T. N. Palmer, A. R. Subbiah (2010) *Extended-range probabilistic forecasts of Ganges and Brahmaputra floods in Bangladesh*. Bull. Amer. Meteor. Soc. 91, 11, 1493-1514

Foresight
1 Victoria Street
London SW1H 0ET
www.foresight.gov.uk
URN 12/1302