Joseph D. Daron and David A. Stainforth
Assessing pricing assumptions for weather index insurance in a changing climate

Article (Published version)
(Refereed)

Original citation:
DOI: 10.1016/j.crm.2014.01.001

© 2014 The Authors. Published by Elsevier B.V.

This is an open access article under the CC-BY 3.0 license

This version available at: http://eprints.lse.ac.uk/59154/
Available in LSE Research Online: August 2014

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (http://eprints.lse.ac.uk) of the LSE Research Online website.
Assessing pricing assumptions for weather index insurance in a changing climate

J.D. Daron\textsuperscript{a,\textcopyright}, D.A. Stainforth\textsuperscript{b,c,d}

\textsuperscript{a}Climate System Analysis Group, University of Cape Town, Rondebosch, Cape Town, South Africa
\textsuperscript{b}The Grantham Research Institute on Climate Change and the Environment, Centre for the Analysis of Time Series, London School of Economics, Houghton Street, London, UK
\textsuperscript{c}Department of Physics, University of Warwick, Coventry, UK
\textsuperscript{d}Environmental Change Institute, University of Oxford, Oxford, UK

\textbf{Article info}

\textbf{Article history:}
Available online 11 February 2014

\textbf{Keywords:}
Climate modeling
Uncertainty
Bayesian Networks
Adaptation
India

\textbf{Abstract}

Weather index insurance is being offered to low-income farmers in developing countries as an alternative to traditional multi-peril crop insurance. There is widespread support for index insurance as a means of climate change adaptation but whether or not these products are themselves resilient to climate change has not been well studied. Given climate variability and climate change, an over-reliance on historical climate observations to guide the design of such products can result in premiums which mislead policyholders and insurers alike, about the magnitude of underlying risks. Here, a method to incorporate different sources of climate data into the product design phase is presented. Bayesian Networks are constructed to demonstrate how insurers can assess the product viability from a climate perspective, using past observations and simulations of future climate. Sensitivity analyses illustrate the dependence of pricing decisions on both the choice of information, and the method for incorporating such data. The methods and their sensitivities are illustrated using a case study analysing the provision of index-based crop insurance in Kolhapur, India. We expose the benefits and limitations of the Bayesian Network approach, weather index insurance as an adaptation measure and climate simulations as a source of quantitative predictive information. Current climate model output is shown to be of limited value and difficult to use by index insurance practitioners. The method presented, however, is shown to be an effective tool for testing pricing assumptions and could feasibly be employed in the future to incorporate multiple sources of climate data.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/3.0/).

1. Introduction

1.1. Index insurance

Many private and public insurance schemes covering risks in the developed world are considered infeasible in developing countries (Skees et al., 1999). Traditional forms of \textit{claims based} crop insurance cover multiple perils, including weather related perils such as hail and drought, as well as non-weather perils such as pest and disease outbreaks. However, the
associated premiums are usually unaffordable for low-income small-holder farmers and the speed at which payouts are made is often too slow to alleviate the adverse impacts of loss events on their livelihoods. Alternative forms of crop insurance are being sought and over the past decade, a number of index insurance products have been developed, offering lower premiums and speeding up the payout process.

Index insurance products for agriculture can be divided into two types: area yield index insurance and weather index insurance (WII). For area yield index insurance, “the indemnity is based on the realised average yield of an area such as a county or district, not the actual yield of the insured party”, while for WII, “the indemnity is based on realisations of a specific weather parameter measured over a pre-specified period of time at a particular weather station” (World Bank, 2011). Thus WII loss estimates are based on a proxy for loss rather than upon the individual loss of each policyholder (Skees et al., 2007) and once the proxy is triggered (e.g. rainfall accumulation falls below a certain threshold) all policyholders in the area covered receive a payout. The administration costs per policy of index insurance are lower than traditional claims-based forms of insurance as the traditional claims-handling process is bypassed. As a consequence premiums can be lower, so index insurance provides an attractive alternative for small-holder farmers and insurers in developing countries. It is unsurprising therefore that over the past decade there has been an increase in the provision of crop index insurance for weather related perils such as drought, excess rainfall and wildfire. Yet from an insurer’s perspective, setting the thresholds at which payouts for a given peril are triggered is a difficult task, not least because a relationship between meteorological events and the incurred losses must be established and quantified. As a result WII is associated with “basis risk” which represents the risk that the payout a policyholder receives does not match the actual loss experienced (USAID, 2006; Barnett and Mahul, 2007). Skees (2008) states that because of such risks, index-based schemes are not replacements for traditional insurance but rather serve as a foundation for developing financial risk transfer mechanisms in new markets.

In contrast to traditional claims-based insurance, however, the uncertainties associated with the response of the physical climate system are disaggregated from socio-economic considerations in WII products. It is therefore much simpler to assess the sensitivity of the pricing structure to climate variability and change.

The United Nations advocates the use of WII (Schwank et al., 2010; United Nations, 2012) as a “soft” climate change adaptation measure (Hallegatte, 2009) within developing countries but whether or not index insurance policies are themselves resilient to climate change has received little attention in the climate change adaptation discourse. Climate science has an important role to play in addressing this problem (Hellmuth et al., 2009). Alderman and Haque (2007) note that in order for index insurance to be effective, probability distributions of the relevant variables need to be reliably estimated. In the context of climate change, meteorological variable time series are unlikely to be stationary (McGregor, 2006) so WII products designed using observed data alone will be at risk of under- or over-estimating the probabilities of triggering payments. This risk is present for policies designed for use both today and in the future because under climate change past observations do not represent today’s climate (Daron and Stainforth, 2013; Stainforth et al., 2013).

Climate models are obvious sources of information on the nonstationarity of the relevant time series (Peicai et al., 2003; McGuffie and Henderson-Sellers, 2005). However, given the much debated uncertainties in their interpretation (Stainforth et al., 2007a), it is questionable whether they can be reliably used to inform WII products. Some research has been conducted to examine the benefits of incorporating seasonal climate model forecasts into the provision of WII (Osgood et al., 2008; Carrquiry and Osgood, 2012). Osgood et al. (2008) present an approach to combine the provision of loans with index insurance using seasonal forecasts. The authors conclude, “a scheme that uses skillful seasonal forecasts to adjust the bundled loan-insurance contract according to expected rains can substantially benefit participating farmers”. Considering much longer time scales, Bell et al. (2013) analyse the value of paleoclimate reconstructions (specifically tree-ring data) for index insurance in providing additional information to improve our understanding of the underlying climate distributions and past variability. In this paper the focus is on decadal and multi-decadal time scales and the assessment of WII pricing assumptions both today and in the future, in the context of climate change. The key questions relate to how we evaluate the available probability distributions and how we integrate multiple sources of climate information. While our study is focused on WII, the implications of our results have broader relevance to climate-sensitive decisions across the insurance sector.

Hochrainer et al. (2008) state that the viability of microinsurance schemes can be viewed from two perspectives: from the view of insurers (the supply-side perspective) and from the view of the clients (the demand-side perspective). This study investigates the extent to which climate model information can inform the supply-side. For this purpose a Bayesian Network (BN) structure is presented as a means of combining different sources of quantitative climate information.

1.2. Bayesian networks for climate change applications

A BN is a graphical/conceptual model with an underlying probabilistic framework which characterizes and quantifies an outcome of interest, along with the relevant variables and the causative interactions (Donald et al., 2009). BNs propagate probabilities throughout the network according to Bayes Theorem (Box and Tiao, 1973). They can be used to explicitly quantify risk and uncertainty drawing on evidence from diverse sources, including both subjective beliefs and objective data (Fenton and Neil, 2004). Cain (2001) distinguishes between two potential uses of a BN:

---

1. The process of verifying that a loss has occurred and administering a payout appropriate for the level of damage.

2. Insurance characterized by low premium and low caps or low coverage limits.
1. To provide a mathematically optimal decision on the basis of the information provided to the BN.
2. To promote an improved understanding of the system, leaving the decision makers to reach their own conclusions on the basis of that knowledge.

Given the size and nature of the uncertainties associated with climate predictions on the scales relevant for WII, it is only conceivably appropriate to use BNs to promote improved understanding rather than attempting to provide optimal decisions (Stainforth et al., 2007a). For climate change applications, a BN can be used as a tailorable decision tool, providing a means for quantifying and communicating the relationship between output from complicated computer models and other factors and pressures facing a decision maker. Furthermore, the network does not have to be complete in order for the BN to act as an “instrument for consultation” (Ouerdane and Tsoukias, 2009).

This approach has been used extensively in a wide range of research fields and industrial sectors as a tool for structuring and informing the decision-making process (Heckerman et al., 1995; Fenton, 2007; Fenton and Neil, 2007). The tool has also been applied to a small number of climate-related studies which each address specific decision problems (Kuikka and Varis, 1997; Wooldridge et al., 2005; Musango and Peter, 2007). These studies have relied on the output of a single model or on individual time series of observations; they have not been used to test the sensitivity of decisions to multiple sources of climate information. While acknowledging that the BN approach is only one method to combine data, in this paper we show how a BN might be used to investigate the sensitivity of WII pricing assumptions to different sources of climate information; observations, reanalysis data and climate model forecasts. The overall aim of the paper is to demonstrate a method for assessing the validity of WII pricing assumptions in the context of climate change, and in doing so the analysis aims to expose the benefits and limitations of the BN approach, WII as an adaptation measure and climate simulation data as a source of quantitative predictive information on the relevant scales.

The next section outlines the adopted methodology, including a description of the case study focus region. An analysis of the model and observational data used in the BNs is presented in Section 3. Section 4 shows the results of the case study, presenting a number of BNs incorporating different assumptions and combinations of climate data. The implications of the results for insurers and climate information providers are discussed in Section 5, and Section 6 concludes the paper, linking the results to the specific aims of the study.

2. Methodology

2.1. Case study description

The focus region for the case study is Kolhapur district, in western India (Fig. 1). Kolhapur district has a total land area of 768,500 hectares, of which 104,000 hectares (13.5%) are devoted to rice production, making it the main crop for local farmers and the region's primary commercial crop (Collector Office, Kolhapur, 2009). The rice crop is a Kharif crop; i.e., the main sowing period is during the summer monsoon season (June to September). This region was chosen for a number of reasons. Firstly, a large volume of high resolution Regional Climate Model (RCM) data for India has recently been made available from the HighNoon project – a project exploring the future water resources of the Ganges river basin (Wiltshire et al., 2010) – and gridded observed daily rainfall data is also available from the APHRODITE project (Yatagai et al., 2008, 2009). Secondly, the rain-fed agricultural sector in India is largely dependent on the timing and magnitude of the summer monsoon rains (Gadgil and Rupa Kumar, 2006) and Kolhapur is located at the edge of the core monsoon region (Fig. 1). Thus, any shifts in the location, timing and/or magnitude of the monsoon under climate change would have a significant impact on the region's crop yields. Finally, the microinsurance company MicroEnsure have a pilot WII project for farmers in the area (Hazel et al., 2010) suggesting that there is indeed a market for appropriately designed microinsurance.

When purchasing WII cover, policyholders can usually specify the level of coverage (a function of the expected yield) as well as the dates on which the policy begins and terminates. Typically these dates will correspond to the growing season or harvesting period, where a lack or surplus of rainfall can damage crops and adversely affect yields. For simplicity, the analysis presented in this case study focuses on cumulative rainfall in July, the month historically associated with the highest accumulations of rainfall in the monsoon season in Kolhapur, and a key month for the success of the rice crop (Bhuiyan, 1992). The BNs developed here examine the rainfall triggers at which payouts might occur for rice crop losses associated with both deficit and excess rainfall.

2.2. Climate model and observational data

The European Commission project “HighNoon” was established to assess the impact of Himalayan glacier retreat and possible changes in the Indian summer monsoon on the water availability in northern India (HighNoon, 2009). The project utilised data from the UK Hadley Centre regional model, HadRM3 (Jones et al., 2004). The HadRM3 configuration used in the model simulations has a horizontal resolution of 25 km, with 19 vertical levels in the atmosphere and includes MOSES II (Met Office Surface Exchange Scheme), a tiled land surface scheme (Essery et al., 2001). As part of the project, RCM runs were performed for a domain covering India and a broad area of southern Asia (Fig. 2). The HadRM3 model was run using boundary conditions from two General Circulation Models (GCMs); ECHAM5 (Roeckner et al., 2003; Hagemann et al., 2006) and
HadCM3 (Gordon et al., 2000; Pope et al., 2000; Collins et al., 2001). The GCM simulations were run according to the A1B SRES emissions scenario (Nakićenović et al., 2000) for the period 1960–2099. In addition, the HadRM3 model was run using boundary conditions from ERA-Interim reanalysis data (Simmons et al., 2010) for the period 1990–2008. Monthly rainfall totals for the model grid cell encapsulating Kolhapur city (16°70′N, 74°23′E) were extracted from the HighNoon data archive. Observed data was obtained for the period 1961–2004 from the APHRODITE project which has developed a high resolution (0.25° x 0.25°) gridded daily rainfall dataset for Asia utilising rain-gauge observations (Yatagai et al., 2008, 2009).

2.3. Constructing the Bayesian networks

In this paper, the BN tool is used to examine how WII pricing decisions might be affected when incorporating different sources of climate information. To determine the most appropriate quantitative output of the BNs to inform insurer’s WII pricing strategies, discussions were held with insurance experts from Lloyd’s of London and Willis Reinsurance. From these discussions, it was clear that developing a BN which shows the sensitivity of WII premiums to different choices of input data (climatic and non-climatic) would potentially be a valuable decision aid.

The methodology developed by Ames et al. (2005) was used to guide the construction of the BNs. This methodology details the steps required to construct a BN, taking into account the selection of variables, constants and causal relationships between the network nodes. A number of BNs were constructed with different levels of complexity. The results are presented in Section 4 and the output is discussed in the context of uncertainty and model error in Section 5. In the following section, the raw climate model data and historical observed data used in the analysis are examined in more detail.

3. Climate data analysis

The raw model data extracted from the HadRM3 runs driven by HadCM3, ECHAM5 and ERA-Interim is first compared to the observed data. Fig. 3 shows cumulative July rainfall totals from each source of data for the period 1960–2050. Table 1 contains summary statistics for the period 1961–2004 and the 30 year period centered on the 2030s (2020–2049) which is considered the extent of the time horizon for strategic microinsurance decisions (Daron, 2012).

The data reveals large systematic biases between the raw model output and the observational data (see Table 1). For the period 1961–2004, the HadRM3/HadCM3 combination produces a mean rainfall which is just over half that observed (234.01 mm compared to 405.06 mm), while the mean rainfall from the HadRM3/ECHAM5 and HadRM3/ERA-Interim combinations are even further from the observations (111.57 mm and 120.90 mm, respectively); note that the HadRM3/ERA-Interim data is for 1990–2008. The annual variability is also greater in the observations than in the RCM output; by a factor of three to four. The HadRM3/ERA-Interim reanalysis data produces particularly low variability (year to year standard deviation of 37.93 mm) relative to the observed rainfall values (year to year standard deviation of 162.36 mm). This may be partly due to the shorter length of this time series (19 years cf. 44 years) but is also likely to be a consequence of the reanalysis/model combination not capturing the annual rainfall patterns accurately. Both the HadRM3/HadCM3 and HadRM3/ECHAM5 output show an increase in rainfall variability under future climate conditions.

Fig. 1. Indian states in the core monsoon region (highlighted in green): the black circle denotes the location of the Kolhapur district (IITM, 2011). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Over the period analysed, the HadRM3/ECHAM5 model results are in close agreement with the HadRM3/ERA-Interim results. In the absence of observational data, one might therefore conclude that results should be weighted more strongly in favour of the HadRM3/ECHAM5 model and against HadRM3/HadCM3. However, the HadRM3/HadCM3 model results, although still much lower, are closer to the observational data than the HadRM3/ECHAM5 results. On this evidence one might conclude that the HadRM3/HadCM3 data should be weighted more favourably. Yet the large systematic errors in all RCM output suggests that weighting of models is inappropriate (Stainforth et al., 2007b); a more extensive discussion of model weighting is found in Section 5.

**Fig. 2.** "HighNoon" geographical model domain used for the HadRM3 RCM runs. The domain extends from 60°E to 100.25°E and 4°N to 40.25°N. Each model grid box has a resolution of 0.25° meaning there are 23,345 grid cells covering the horizontal model domain and 443,555 grid cells in total accounting for the 19 vertical levels in the atmosphere.

**Fig. 3.** Model and observational cumulative rainfall data for Kolhapur, India in July. Observational data (black) is shown for the period 1961–2004; HadRM3/ERA-Interim data (purple) is shown for the period 1990–2008; and the HadRM3/HadCM3 data (blue, short-dashed line) and HadRM3/ECHAM5 data (red, long-dashed line) is shown for the entire period 1960–2050. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
Table 1
Summary statistics for cumulative rainfall in Kolhapur for July from observations and the HadRM3 output driven by ERA-Interim, HadCM3 and ECHAM5 for given time intervals.

<table>
<thead>
<tr>
<th>Time period</th>
<th>Data source</th>
<th>$\mu$ (mm)</th>
<th>$\sigma$ (mm)</th>
<th>Linear trend (mm/yr)</th>
<th>Min (mm)</th>
<th>Max (mm)</th>
<th>Range (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1961–2004</td>
<td>Observations</td>
<td>405.06</td>
<td>162.36</td>
<td>−6.42</td>
<td>155.99</td>
<td>1043.76</td>
<td>887.77</td>
</tr>
<tr>
<td></td>
<td>HadRM3/HadCM3</td>
<td>234.01</td>
<td>54.72</td>
<td>0.27</td>
<td>121.82</td>
<td>361.96</td>
<td>240.14</td>
</tr>
<tr>
<td></td>
<td>HadRM3/ECHAM5</td>
<td>111.57</td>
<td>40.57</td>
<td>0.12</td>
<td>27.07</td>
<td>229.69</td>
<td>202.63</td>
</tr>
<tr>
<td>1990–2008</td>
<td>HadRM3/ERA-Interim</td>
<td>120.90</td>
<td>37.93</td>
<td>−3.00</td>
<td>31.31</td>
<td>185.80</td>
<td>154.50</td>
</tr>
<tr>
<td>2020–2049</td>
<td>HadRM3/HadCM3</td>
<td>269.00</td>
<td>66.89</td>
<td>0.45</td>
<td>146.39</td>
<td>409.41</td>
<td>262.83</td>
</tr>
<tr>
<td></td>
<td>HadRM3/ECHAM5</td>
<td>119.27</td>
<td>46.44</td>
<td>−0.36</td>
<td>31.64</td>
<td>229.68</td>
<td>198.05</td>
</tr>
</tbody>
</table>

The linear trend (column 5 of Table 1) shows a significant decrease in observed total July rainfall (−6.42 mm/yr) over the observed period; significant at the 99.9% confidence level. While the magnitude of the trend is influenced by the anomalously high value in the first year, there is still a decrease of −4.75 mm/yr even if this value is omitted; the negative trend remains significant at the 99.5% confidence level. The HadRM3/ERA-Interim shows a weaker decreasing trend (albeit for a shorter time period), significant at the 90% confidence level, but the HadRM3/HadCM3 and HadRM3/ECHAM5 output do not show any significant trends over the hindcast period. The inability of the models to capture the observed trends over the hindcast period suggests that any conclusions based on the trends evident in the forecast period should be, at best, tentative. The BNs presented in Section 4 should therefore only be interpreted as illustrative examples of how one might use additional climate data to inform WII design.

The discrepancies in the mean rainfall statistics may be attributed either to an overestimate of rainfall accumulation in the observations or an underestimate in the model results. Systematic errors in GCM rainfall fields are not uncommon (Schmidli et al., 2006) and it is these GCM fields which provide the boundary conditions for the RCM simulations. It is unsurprising therefore to see large differences in the rainfall totals between the HadRM3/HadCM3 output and the HadRM3/ECHAM5 output, despite their use of the same RCM.

The Intergovernmental Panel on Climate Change (IPCC) fourth assessment report (AR4) shows a multi-model mean increase in summer rainfall over western India of between 5% and 10% for the 2080–2099 period; using 1980–1999 as a base period (Christensen et al., 2007). However, only 13 of the 21 AR4 GCMs analysed show an increase while the remainder show a decrease. A combination of insufficient model resolution and limited process understanding and representation means that the large-scale mechanisms influencing monsoon rainfall are not very well represented in these GCMs. An RCM generates output which inherits these errors from the driving GCM. The disagreement between the ECHAM5 and HadCM3 model output is likely indicative of wider disagreement amongst GCMs.

Additional sources of error may also contribute to the discrepancies. Whilst RCMs typically provide much better representation of the land surface and orography than GCMs, many of the subgrid-scale features are still not adequately represented by a 25 km HadRM3 horizontal resolution. Enhancing model spatial resolution further may reduce biases resulting from errors in the specification of orography and the land surface (McGregor, 1997) although others will inevitably remain. There are also uncertainties associated with the observations and the interpolation of rain-gauge station data to a gridded dataset. Wiltshire et al. (2010) states, “the uncertainties associated with observed rainfall are potentially large primarily due to known gauge undercatch errors (which are often corrected in post-processing) as well as the high spatial variability of rainfall which is often statistically undersampled”.

Given these multiple sources of uncertainty and the lack of out-of-sample verification for climate models (Stainforth et al., 2007a) it is not clear which data should be considered more or less reliable. The aim of this paper is not, however, to provide definitive results regarding the best source of information or even the optimum pricing strategy for index insurance in Kolhapur, but rather to demonstrate how a BN tool might be utilised to investigate the sensitivity of pricing assumptions to different input climate information. The errors and uncertainties are simply part of this information; the basis for a judgement about the viability of WII from an insurance perspective and the usefulness of WII as a climate change adaptation strategy. The next section presents the results of the BN analysis.

4. Results

In this section results are presented for multiple BNs with different underlying data inputs and different methods of data processing and interpretation. Section 4.1 presents a BN that utilises observational data only, Section 4.2 presents BNs that combine observations with raw climate model simulation data, and Section 4.3 presents BNs with bias corrected model output.

4.1. BN using observational data

The BN developed in this section incorporates observational data only to illustrate how WII premiums might be determined if reliant solely on a single observed time series. A three tier payout profile is adopted (see Fig. 4); a step-wise function is chosen because existing WII products are often structured in this way (MicroEnsure, 2010). Within each tier a fixed
proportion of the maximum indemnity (total insured value) of a policy is paid. The tiers correspond to specific rainfall thresholds, though defining these thresholds is itself subject to considerable uncertainty. Rajagopalan (2009) states that the rice crop requires about 1200 mm of rain over the monsoon season whilst Bhuiyan (1992) cites a suitable range as being between 700 and 1500 mm. Veeramani et al. (2005) state that the rice crop requires between 250 and 350 mm each month during the monsoon and highlight that it is more sensitive to drought than flooding. Guided by these studies, a maximum payout for rainfall deficit is chosen to correspond to a rainfall total below 150 mm in July, while a maximum payout for excess rainfall is chosen to correspond to a rainfall total above 600 mm in July. Intermediate thresholds are selected for smaller payouts in the event of partially decreased but not failed yields.

While the focus here is on the month of July, one could apply the approach to consider other coverage periods. In addition, a BN could be developed to inform a more complicated product structure combining multiple coverage periods with different triggers and additional features such as conditional payouts. However, extending the methodology would only be worth considering if the available data proved informative and fit for purpose. By concentrating on a simple idealised product structure, it is possible to investigate the adequacy of input climate data; if it is inappropriate for a simple idealised product, it cannot be assumed to be any more appropriate for a more complicated product. Indeed, Giannini et al. (2009) show that more complicated product structures are often more sensitive to methodological choices and assumptions.

The BN shown in Fig. 5 consists of seven nodes linking observed July rainfall to WII payouts (i.e., losses for the insurer) and the required premium. Given this pricing structure and input data, the BN shows that 75% of the observed years would have been associated with no losses; the insurer would not have been required to payout on any of the policies. The remaining years would have incurred losses; 9% of years (i.e., four years) resulting in a maximum payout to each policyholder. Administration costs and the expenses of marketing and maintaining a policy are taken into account in the Policy Loading node. This node is assigned a constant percentage of the maximum indemnity. The Loss node combines the probabilities in the parent nodes to produce probabilities for each payout tier. The combined probabilities are multiplied by the payout values assigned to each tier, shown in Table 2, to produce an expected loss value representing the expected payout per year as a percentage of the maximum indemnity. In this case it is 13.8%. The premium necessary to render the policy potentially viable can then be calculated using Eq. (1). This is done in the premium calculation node and presented in the premium required node. In Fig. 5, the policy loading costs are assumed to be 4% of the maximum indemnity which leads to a premium value of ~18% for marginal profitability to the insurer.

\[
\text{Premium} = \text{Expected Loss} + \text{Policy Loading}
\] (1)

4.2. BNs combining observational and raw climate model data

To investigate how climate model data might be used to aid pricing decisions, 30 year time intervals centered on the decade of interest are selected from the available data; i.e., the 2030s are represented using the data from 2020 to 2049. In Fig. 6 the BN is made more complicated by allowing for a number of different sources of input climate data. These may be used individually or in combination as controlled by the Climate Information Source node. Different time periods may be selected using the Climate Time Scale node. The two parent nodes for the HadRM3/HadCM3 July Rainfall node and the two parent nodes for the HadRM3/ECHAM5 July Rainfall include frequency distributions for the model rainfall data over Kolhapur for the relevant periods.

In Fig. 6 the different sources of data are given equal weights and the period assessed is the past (1961 to 2004); weighting the data raises fundamental challenges in interpretation (see discussion in Section 5) but is illustrated here for demonstration purposes. The resulting probabilities show that, with a policy loading of 4%, a premium of 53.4% of the maximum indemnity is required to yield the policy structure potentially viable given the assumptions in the BN. According to the BN, there is more than a 40% probability of rainfall during July being below 150 mm. Given the need for such a high premium to account...
for the large probability of a rainfall deficit event, it would be highly unlikely that any insurer would consider the product viable under these data and pricing assumptions. The high probability of a payout for drought conditions is, however, mostly attributable to the model data. The BN structure enables us to explore the sensitivity of this conclusion to different assumptions regarding the information content of the various sources of data. It is a matter of professional judgement as to which, if any, of these data contain relevant information for a particular purpose.

Fig. 7 shows the BN when the HadRM3/HadCM3 model alone is selected and the time period is set to the 2030s. Recall from Fig. 3 that the rainfall totals from the HadRM3/HadCM3 model are significantly lower than the observational data during July but greater than HadRM3/ECHAM5. In Fig. 7 all of the probability density remains below the 540 mm threshold for excess rainfall payout. There is, however, density below the rainfall deficit threshold values with 3.33% (one year out of the 30 year sample) below 150 mm. In this scenario, for this particular GCM forcing this particular RCM, a premium of 11.5% would ensure the long term viability of the product, conditional on the given set of network assumptions.

If one believed the future rainfall values associated with the raw HadRM3/HadCM3 model output in the 2030s, then climate change would appear to pose little threat to the long-term viability of WII in the Kolhapur region; at least it is no less viable than observations suggest it would have been in the past. However, selecting the raw model output from the HadRM3/ECHAM5 model in the BN implies a very different conclusion. Table 3 shows the sensitivity of the required premium to the alternative sources of climate information. In the HadRM3/ECHAM5 case, again for the 2030s, 70% of the rainfall data falls below the 150 mm threshold associated with the maximum rainfall deficit payout. If this data were reliable then the WII product would clearly not be viable based on the policy structure and thresholds selected.
These results show that the source of climate information significantly alters the assessment of viability for the WII product structure in Kolhapur under both past and projected future climatic conditions. However, given the evidence of systematic model errors described in Section 3, an obvious next step is to examine the effect of applying bias corrections to the model output. In the following section, a simple bias correction is applied to the model data and the impacts on the BN results analysed.

4.3. BNs incorporating bias-corrected climate model data

In this section, the BNs presented show how the viability of the pricing structure in the idealised WII product can be assessed using bias corrected climate model output. The introduction of statistical bias corrections to model simulations of future climate at this scale has no physical justification (see discussion in Section 5) but is nevertheless widely used (Christensen et al., 2008; Piani et al., 2010) and is therefore worthy of inclusion in this analysis. The bias correction method applied in the BNs here is therefore purely demonstrative of the approach one could adopt if there was reason to believe that model biases are insensitive to the underlying climate forcings.

Acknowledging the systematic errors in the model results described in Section 3, a two moment bias correction is applied to the HadRM3 model output. Each of the HadRM3 distributions, driven by HadCM3, ECHAM5 and ERA-Interim, are offset to be consistent with the observed mean and scaled according to the different observed and RCM hindcast standard deviations. Other methods are possible but this is sufficient to show how users may incorporate bias corrected data in a BN framework; adjusting data to higher order moments or using alternative methods is superfluous for the purposes of this analysis.

Table 4 shows the mean biases of the RCM output from the observed data for the period 1961–2004 (1990–2008 for ERA-Interim data) as well as the ratio of the observed data standard deviation to the model data standard deviations. Eq. (2) is applied to adjust the hindcast model precipitation time series \( P_h \), using values of the mean model hindcast precipitation \( \bar{P}_h \), the mean observed precipitation \( \bar{P}_o \), and the standard deviations in the model hindcast and observed data, \( \sigma(P_h) \) and \( \sigma(P_o) \), respectively.

\[
P_{h\text{bias adjusted}} = (P_h - \bar{P}_h) \times (\sigma(P_o)/\sigma(P_h)) + \bar{P}_o
\]

Eq. (3) is used to bias correct the future model projections; the subscript \( f \) refers to the future model projection data. For the future projections, the bias corrections applied reflect the change in the mean between the two model periods and uses the mean and standard deviation bias corrections from the hindcast period.
\begin{equation}
P_f(\text{bias adjusted}) = \left( P_f - P_f \right) \times \frac{(\sigma(P_o)/\sigma(P_h)) + \bar{P}_o + (\bar{P}_f - \bar{P}_h)}{18}
\end{equation}

Fig. 8 shows how the bias corrected HadRM3/ECHAM5 hindcast data affects the threshold exceedance probabilities of the WII product for the past climate period, 1961–2004. The bias corrected model data is no longer constrained to low precipitation values and spans the range consistent with the observations. As a consequence, the BN shows that a premium of 19% is required (assuming a policy loading of 4%); compare this with the 95% for the non-bias corrected situation (see Tables 3 and 5). The BN shows a 13.6% probability of a maximum Tier 3 payout with lower probabilities of a Tier 1 or Tier 2 payout, resulting in an expected loss of 15.2%.

When the bias correction is applied to the future projection of the HadRM3/ECHAM5 model (see Fig. 9) the probabilities suggest an increase in the risk of a loss event. Here the BN indicates a slight increase in the density in both the higher and lower range of rainfall values corresponding to increased probabilities of payouts in all tiers compared to the BN in Fig. 8. There is now a 20% probability of incurring a Tier 3 loss resulting in an expected loss of 25%; the required premium is thus 29% compared with 83% in the non-bias corrected case (see Tables 3 and 5).

The bias corrections applied to the model data dramatically alter the probability distributions but it is important to again emphasize that the methodology applied here is simply demonstrative. As in Section 4.2, the BN can be used to test the sensitivity of the required premium to various pricing assumptions and climate information sources. Table 5 contains the required premiums to yield the pricing structure viable using the bias corrected model output. The results using this method suggest that the premium will need to increase in the future to maintain profitability. The HadRM3/HadCM3 output shows a 7 percentage point increase in the required premium while the HadRM3/ECHAM5 output shows a 10 percentage point increase. The HadRM3/ERA-Interim bias corrected output (for the period 1990–2008) suggests a premium of 23%.

---

**Table 3**

<table>
<thead>
<tr>
<th>Time period</th>
<th>Data source</th>
<th>Required premium (% of maximum indemnity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1961–2004</td>
<td>Observations</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>HadRM3/HadCM3</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>HadRM3/ECHAM5</td>
<td>95</td>
</tr>
<tr>
<td>1990–2008</td>
<td>HadRM3/ERA-Interim</td>
<td>86</td>
</tr>
<tr>
<td>2020–2049</td>
<td>HadRM3/HadCM3</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>HadRM3/ECHAM5</td>
<td>83</td>
</tr>
</tbody>
</table>

---

**Fig. 7.** BN with HadCM3 data selected – the 2030s. BN expanded from Fig. 5 to inform the viability of WII for rice production in Kolhapur, India using multiple sources of climate data. The HadRM3/HadCM3 model is selected in the **Climate Information Source** node, the **Climate Time Scale** set to the 2030s (model period 2020–2049) and the policy loading is 4% of the maximum indemnity.
By incorporating this bias corrected model output and reanalysis data into the pricing decisions, an index insurer may be cautious in their assessment of current and future WII viability; perhaps more so than if they simply used past observations.

5. Discussion

In providing WII to low-income farmers, the objective of an insurer is to design a product which is attractive to the policyholder, represents the underlying risk and is ultimately profitable. To sustain profitability in light of altered risks, the most obvious option for the insurer is to alter premiums to reflect the underlying risk but insurers have many other levers available. For example, thresholds at which payouts occur, the value of the payouts and limits on coverage can all be adjusted in the WII policy design. While this study focuses on the required premiums under climate change, it is important to note that BNs might be used to guide policy decisions related to the entire policy structure. Crucially, however, the success of the BN approach is conditional on the accuracy and relevance of the input data.

The results presented demonstrate that in the context explored in this study, the primary application of the BN tool would not be to obtain an optimal pricing structure but rather to assess the sensitivity of different policy structures to a variety of sources of climate data. The BNs presented in Section 4 show how the tool might be used to stress-test pricing assumptions and guide the design of index insurance products in a changing climate but in applying the tool to inform decisions, insurers must be wary of over-interpreting the probabilistic output. Section 4.1 shows how the required premium might be derived if reliant on observed rainfall measurements only. Given a policy with a three tiered payout structure, and an assumed policy loading of 4% of the maximum indemnity, it is found that a premium of 18% is required. However, with decadal and multi-decadal climate variability and climate change, it is unlikely that the temporal distribution of observed rainfall in the past is the same as the probability distribution of expected rainfall for the present climate; even less so for future climate. Fig. 3 shows that there has been

Table 4
Mean biases and ratio of standard deviations for model output compared to observed July rainfall data over Kolhapur from 1961 to 2004 (1990 to 2008 for ERA-Interim data).

<table>
<thead>
<tr>
<th>Data source</th>
<th>Observed $\mu$ minus model $\mu$ (mm)</th>
<th>Observed $\sigma$/model $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HadRM3/HadCM3</td>
<td>171.05</td>
<td>2.97</td>
</tr>
<tr>
<td>HadRM3/ECHAM5</td>
<td>293.49</td>
<td>4.00</td>
</tr>
<tr>
<td>HadRM3/ERA-Interim</td>
<td>284.16</td>
<td>4.28</td>
</tr>
</tbody>
</table>

Fig. 8. BN with bias corrected ECHAM5 data selected – the past. BN expanded from Fig. 5 to inform the viability of WII for rice production in Kolhapur, India using multiple sources of climate data including bias corrected model data. The HadRM3/ECHAM5 model is selected in the Climate Information Source node, the Climate Time Scale set to the past (observed period 1961–2004) and the policy loading is 4% of the maximum indemnity.

By incorporating this bias corrected model output and reanalysis data into the pricing decisions, an index insurer may be cautious in their assessment of current and future WII viability; perhaps more so than if they simply used past observations.
a downward trend in observed rainfall during July suggesting that the probability of a year with a rainfall deficit has increased in recent years whilst the probability of a year with excess rainfall has decreased. Without a better understanding of internal climate variability, and given only a limited sample of observations, we must be careful not to over-interpret the trend statistics but we can use current understanding to urge caution in basing WII policies on the statistics of past observed climate alone.

To illustrate how model reanalysis datasets and forecasts might be used as an additional source of information to aid the policy design process, Section 4.2 presents BNs which combine both model and observed data. Table 3 shows the sensitivity of the required premiums to each climate information source. The BN in Fig. 6 shows how providing equal weights to each of the information sources might influence the probability of a loss. The resulting required premium is 53% which is considerably higher than the value of the premium calculated using observations alone (18%). However, the decision to weight different sources of climate model information is contentious; it is not clear that any suitable weighting strategy exists (Stainforth et al., 2007b; Tebaldi and Knutti, 2007). Numerous methods have been developed to establish different metrics and measures of model skill (Tebaldi et al., 2005; Greene et al., 2006) yet the decision to weight models based on their ability to capture past observations and key climatic processes has been questioned (Stainforth et al., 2007b). In a recent study, Weigel et al. (2010) state:

“When internal variability is large, more information may be lost by inappropriate weighting than could potentially be gained by optimum weighting...results indicate that for many applications equal weighting may be the safer and more transparent way to combine models”.

<table>
<thead>
<tr>
<th>Time period</th>
<th>Data source</th>
<th>Required premium (% of maximum payout)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1961–2004</td>
<td>Observations</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>HadRM3/HadCM3</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>HadRM3/ECHAM5</td>
<td>19</td>
</tr>
<tr>
<td>2020–2049</td>
<td>HadRM3/HadCM3</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>HadRM3/ECHAM5</td>
<td>29</td>
</tr>
</tbody>
</table>

![BN with bias corrected ECHAM5 data selected – the 2030s. BN expanded from Fig. 5 to inform the viability of WII for rice production in Kolhapur, India using multiple sources of climate data including bias corrected model data. The HadRM3/ECHAM5 model is selected in the Climate Information Source node, the Climate Time Scale set to the 2030s (model period 2020–2049) and the policy loading is 4% of the maximum indemnity.](image-url)
One of the main barriers to the formal uptake and use of the BN approach to incorporate climate model output in aiding WI product design is likely to be model error. Analysis presented in Section 3 reveals that the model rainfall output in the case study focus region is associated with large model biases and the results in Section 4.2 show how using this output, without accounting for model error, can change the implied premiums to unviable values. The HadRM3 model output data, driven by the HadCM3 and ECHAM5 GCMs and the ERA-Interim reanalysis data, significantly underestimates the observed rainfall occurring in Kolhapur in the monsoon season (see Fig. 3).

Capturing the complex interactions that control the monsoon rains is challenging, perhaps impossible given limited computing capacity, current process understanding and imperfect models. A number of studies have attempted to model the monsoon rainfall under altered climate forcings (Meehl and Washington, 1993; Kitoh et al., 1997; Meehl et al., 2000; Lal et al., 2001; Lucas-Picher et al., 2011). These studies are inconsistent in their conclusions; probably reflecting the limits of our current understanding of the relevant processes. Nevertheless, incorporating results from a wide range of models into a BN could provide a greater understanding of the implications of current model uncertainty and, of course, the BNs presented in Section 4 could be updated with new model information when it becomes available. The question would remain, of course, as to whether model uncertainty reflects the “true” range of uncertainty which is a function of our imperfect understanding.

The results in Section 4.3 show how model biases might be taken into account. Incorporating the bias corrected model output in the pricing decisions leads to substantial increases in the implied premiums in the future (to 29–31% of the maximum indemnity in the 2030s, see Table 5). Such methodologies are however purely demonstrative and there are good reasons for not bias correcting rainfall data. Statistical bias corrections to rainfall data have a questionable physical basis, not least because systematic decreases or increases in the rainfall values compromises the conservation of total moisture content in the model. Haerter et al. (2011) note that all bias correction methods make assumptions about the applicability of statistical transfer functions from observed climate to future climate and caution that such techniques can obfuscate the nonlinear impacts of climate change in model projections.

An additional source of climate information, not explored in this study, is the use of subjective judgements about future trends for rainfall in the Monsoon region. This information could be obtained using expert elicitation (Granger Morgan et al., 2001; Arnell et al., 2005) and incorporated into a BN. The use of expert judgements is particularly useful in situations where model performance is poor but where physical reasoning and other empirical indicators can inform expert opinions. There is evidence, however, that experts may not hold probabilistic understanding of climatic processes under climate change (Miller et al., 2013). Other sources of quantitative data could also be incorporated into a BN to inform WI design. Paleoclimate data and remote sensing data, particularly satellite data, may help to further explore the discrepancies in the observations and additional model runs under different greenhouse gas emissions scenarios (an issue not examined in this study) could be analysed. Finally, the BNs presented could be extended to include different assumptions about socio-economic variables such as willingness to pay. Complimentary research on the demand-side could establish thresholds at which the premiums become unaffordable for a particular group of potential policyholders. Combining supply-side and demand-side studies could enable substantial progress to be made in understanding the viability of WI mechanisms in a changing climate.

In considering ways to extend the applications of this study, it is important to bear in mind the purpose of examining the BN tool in the context of informing WI design under climate change. As noted in Section 1.2, the BN should only be used as an instrument for consultation and not as a means of informing a mathematically optimal WI policy design. Creating a more complicated BN may be desirable from the point of view of system understanding but cannot be expected to necessarily generate more reliable output. Furthermore, the BN method is not the only way to consider different sources of climate information. Giannini et al. (2009) examine the long-term viability of index insurance in Central America using climate data from historical observations, knowledge of teleconnections and IPCC model projections without using a specific tool to aid the process. However, their approach involved analysis of each data source separately, which can be both time-consuming and fail to address the sensitivities of products to multiple uncertainties. In order to inform estimates of probability distributions for the relevant variables, and provide practical support to index insurance decisions, such methods should allow for a comparison between different sources of climate information. Further study and implementation of the method presented here will help to ascertain whether the tool is beneficial in practice.

6. Conclusions

As the WI sector matures and the use of insurance as a climate change adaptation measure gains support, the need to develop products that account for changing climatic conditions becomes ever more important. Relying on limited observations of past climate to design WI products risks undermining the value of insurance in providing small-holder farmers with protection against meteorological hazards. As evident in the findings presented in this study, using multiple sources of climate information in the design phase may increase our perceived uncertainty but more information is surely better and, used appropriately, could lead to a more robust product. Our findings are also relevant to other areas of the insurance industry which utilise climate information to inform decisions. Without consideration of multiple sources of climate information, and an acknowledgement of the associated biases and errors, insurers are likely to miscalculate and misrepresent the underlying climate hazard risks.

The BN tool could be used as a framework for combining climate model information with observational data to inform WI pricing decisions. This paper demonstrates how index insurers might utilise BNs to consider multiple sources of climate
information and test the sensitivities of loss estimates to different combinations of climate data inputs in a transparent manner. With further input from insurers BNs can be tailored towards specific needs and purposes, thereby excluding irrelevant information. While the BN approach is only one method to combine multiple sources of climate information it has the ability to incorporate further sources of climate information, including paleoclimate data, remote sensing data and expert judgements of possible changes in relevant climate drivers. Crucially however, the user of the BN must be willing and able to interpret probabilistic information conditional on assumptions.

In the Kolhapur case study the existence of model errors and uncertainties in the observations means that any assessment, from a climate perspective, of the long-term viability of WII pricing assumptions for rice crop products in Kolhapur are unlikely to be accurate. In the context of climate change, the absence of reliable, robust, quantitative model projection data means that choices considered optimal based on today’s available model information may lead to maladaptation and risk significant losses to insurers. Combining different sources of data and, crucially, accounting for the associated uncertainties and contradictions in the information remains a significant challenge to index insurers in assessing product viability. Yet by further investigating the implementation of novel methodologies, such as the BN approach, scientists and index insurers can make substantial progress in reducing the risks associated with basing WII decisions on incomplete and imperfect climate information.

Acknowledgements

This work was supported by the EPSRC and Lloyd’s of London, as well as the LSE’s Grantham Research Institute on Climate Change and the Environment and the ESRC Centre for Climate Change Economics and Policy, funded by the Economic and Social Research Council and Munich Re. We would particularly like to thank Trevor Maynard and Nigel Ralph for their valuable input in the initial phases of the work. The UK Met Office provided data for the analysis and special thanks go to Doug McNeill and Andy Wiltshire for their help. We would also like to acknowledge Ross Blamey, Mark Tadross, Olivier Crespo, Hayley McIntosh and Brendan Argo for their helpful comments and input.

References


