A Spatial Approach to Addressing Weather Derivative Basis Risk: A Drought Insurance Example

Nick D. Paulson and Chad E. Hart

Nick Paulson is a graduate student in the Department of Economics, Iowa State University, Ames, IA. Chad Hart is a research scientist in the Center for Agricultural and Rural Development, Iowa State University, Ames, IA.

For questions or comments about the contents of this paper, please contact Nick Paulson, 79 Heady Hall, Iowa State University, Ames, IA 50011; Ph: 515-294-8897; E-mail: *paulson@iastate.edu*.

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Long Beach, California, July 23-26, 2006.

Weather risk markets are among the newest and most dynamic for risk sharing. The market has seen rapid growth with continual emergence of new, weather-based risk management tools over the last decade. The Chicago Mercantile Exchange's (CME) website[1] reports that 20% of the US economy is directly affected by weather, and describes the weather hedging and risk management sector as "today's fastest growing derivative market." Participants come from a broad range of economic sectors including energy, insurance, banking, and agriculture (Varangis, 2001). Exchange traded Cooling Degree Day and Heating Degree Day contracts for 10 U.S. cities were initially opened in 1979 on the Chicago Mercantile Exchange. The CME contracts have since then expanded to 29 cities worldwide, with contracts offered on a variety of weather indexes. Additionally, market survey data indicates that total trading volumes more than quadrupled from 1999 to 2003 while the notional value of trading more than doubled (Van Lennep et al.; Ali; Cao, Li, and Wei). In addition to the exchange-traded weather derivatives, there are many products traded as over-the-counter (OTC) derivatives (Jewson). The Weather Risk Management Association (WRMA) reports that roughly 80% of the traded contracts are for temperature derivatives, comprising 90% of the notional value of the weather risk market. In contrast, the precipitation-based derivative market is still in its infancy stages with only 3% of trade volumes in 2001 and little subsequent growth (Cao, Li, and Wei). Given the success and size of the market for temperature related weather derivatives, a natural area for further development is in precipitation-based weather derivatives.

Figure 1 outlines the major reasons behind crop failures in the United States, illustrating that weather events are the primary source of crop losses in the US. Drought is by far the number one factor, accounting for nearly half of all crop losses. Excess moisture is linked to nearly one-quarter of all crop losses. However, hail remains the only weather event, among named perils, where specific agricultural insurance products have been successfully developed and sold in the US. Moreover, there is a strong correlation between extremes in both temperature and precipitation and catastrophic damages in developing countries (Varangis, Skees, and Barnett).

One of the major problems facing weather-based derivatives is that of weather basis risk. The sellers of weather derivative contracts require high quality data from reputable sources which may exist only in certain locations. The users, or purchasers, of weather derivatives would like to minimize the basis risk involved with the use of weather data collected at a site that does not necessarily correspond with their exposure location (Dischel). Basis risk, in the specific case of rainfall, refers to the relationship between the precipitation measured at the weather station and the production or revenue on the farm. Recorded precipitation may not be highly correlated with actual precipitation at the farm, and production or revenue on the farm may not be highly correlated with precipitation at the farm. Dischel notes that:

> *"Farmers, growers and hydroelectric generators would like to have contracts written on rain falling on their fields, in their groves or over their watersheds. This is generally impossible because the market needs long and accurate measurement records to assess the value of a weather derivative, and unaffiliated parties do not generally compile measurement records at these locations."*

Thus far, weather data has not been used extensively as a basis for crop insurance products. The resulting output of the crop after the weather events (crop yield) can be directly insured for a variety of crops through the multi-peril policies currently offered in the US. Thus, the development of weather-derived insurance products has been limited. This paper develops an insurance policy which would provide coverage for pastureland owners and cow-calf producers against drought risk. Specifically, the policy provides protection against periods of reduced precipitation and drought conditions, which have been outlined as the leading cause of crop losses and are not perils directly covered under any existing FCIC crop policies. The rating methods used in this study could be directly extended to the development of similar products in developing nations.

This study focuses on the first component of basis risk by utilizing a spatial kriging model to interpolate rainfall at locations where actual rainfall is not observed (i.e. the farm). We utilize precipitation data from weather stations administered by the National Oceanic and Atmospheric Administration (NOAA). Since these stations are sponsored by a government agency, all parties involved should have significant confidence in the accuracy of the data. Moreover, extensive time series of historical data are available for multiple locations throughout the state of Iowa. Cross-validation[2] is used to show that the spatial model provides unbiased estimates of unobserved rainfall. The kriging results are also compared to those obtained from a simpler inverse distance weighted (IDW) estimator for rainfall at unobserved locations. Consistent with previous findings, the two methods are shown to be nearly equivalent with respect to the historical rainfall point estimates of interest. The second component of basis risk is addressed through the use of indemnity factors obtained through simple regression analysis relating losses to precipitation shortfalls. The insurance policy is rated as an exotic option on rainfall using the historical rainfall data, the interpolation results, and Monte Carlo analysis assuming that rainfall at a given site follows the Gamma distribution. An historical analysis is included to show the potential performance of the product were it actually marketed. While most authors note the importance of reducing basis risk, this is, to the authors' knowledge, the first direct application of pricing a rainfall insurance policy using the results of a sophisticated interpolation technique.

Background

Insurance

Skees, Barnett, and Hartell outline the necessary conditions for perfect insurability of a risk. First, the loss must be quantifiable and the loss frequency must be calculable to ensure accurate rating for the policy. Second, loss occurrence must be random or unintentional and the potential purchasers of the insurance must be accurately classified by the amount of risk they bring to the overall risk pool to eliminate moral hazard and adverse selection. Finally, there should be a large number of *independent* exposure units to allow the insurer to diversify over the total risk pool. These conditions generally do not hold, and most definitely do not hold in the specific cases of agricultural and weather risk due to their spatial nature. Skees and Barnett classify agricultural risks as “in-between” risks because they generally are not completely independent nor highly correlated. Duncan and Myers explore the impact of catastrophic risk on insurance offerings in a model of risk-averse insurance firms and farmers in a mean-variance framework. They find that unless reinsurance is available and subsidized, an equilibrium where catastrophic insurance is offered may not exist. Weather patterns tend to exhibit positive spatial correlation making losses more volatile from the perspective of the insurer, and thus the cost of maintaining adequate reserves to cover potential losses from systemic events. Thus, insurance may not be the optimal mechanism to provide efficient risk-sharing (Skees and Barnett). However, if the insurer can cover an area large enough to diversify even the systemic risk of weather events, or has access to an adequate reinsurance program, an insurance mechanism should be feasible and implementable. Thus, governments or international organizations may have the potential to play a natural role in providing coverage for weather related risks such as drought. Governments could either offer the insurance directly, or provide reinsurance coverage to existing private insurers similar to the crop insurance program in the US. Natural disaster assistance is an example that is already implemented in both developed and developing nations.

Martin, Barnett, and Coble outline various option structures for precipitation insurance and provide a rating method application for cotton in Mississippi. Skees et al. investigate the development of drought insurance based on a rainfall index in Morocco and find that the product would be both feasible and of significant benefit to Moroccan farmers. Turvey (2001, 1999) also discusses the application of weather derivatives in agriculture by rating various examples of rainfall and temperature options for various locations in Canada. To relate crop yields to weather events, Turvey (2001) examines the correlation of corn, soybean, and hay yields with measures of both rainfall and temperature. Temperature was found to be highly correlated with corn and soybean yields, while precipitation showed more correlation with hay yields. In addition to studies examining the supply-side of product rating, there have also been studies which have explored the demand side for agricultural insurance based on precipitation. Sakurai and Reardon and Gautam, Hazell, and Alderman use household survey data to estimate latent demand for drought insurance in West Africa and Southern India, respectively. Using a set of reduced-form equations resulting from the optimality conditions of a dynamic household optimization problem, both studies estimate a positive latent demand for drought insurance. Additionally, it is estimated that the insurance would be implementable on a full-cost basis. McCarthy estimates the demand for rainfall based insurance contracts for four regions in Morocco, finding that the median willingness to pay for rainfall based insurance was 12-20% above the fair value of the contracts.

Precipitation insurance policies have been explored and utilized in other countries. Argentina, Ethiopia, Mexico, Morocco, Nicaragua, and Tunisia have all tested the feasibility of weather-based insurance products for agriculture (Varangis), while Australia is currently exploring the possibility of developing rainfall insurance (Plate). Two Canadian provinces, Ontario and Saskatchewan, currently have precipitation insurance products on the market offered through Agricorp and Saskatchewan Crop Insurance. The use of precipitation-based insurance in the Canadian provinces is attributed to the high correlation between cattle pasture productivity and rainfall (Varangis, 2001).

The construction of most weather-based insurance policies falls under the larger umbrella of area-based index insurance. Index-based products have a number of advantages. Adverse selection and moral hazard problems are minimized because the underlying index is uncontrollable by the insured. Additionally, individualized product set-up, inspection, and loss adjustment are not required. Thus, weather based insurance for agriculture may be of higher interest for developing countries, as they may lack the required resources needed to develop a crop insurance system such as that of that of the US. Moreover, markets for such policies could be opened up to any interested parties, while private companies may be able to construct "add-on" products that cover the individual risks left outside of the index product's coverage. Specifically, the systemic or catastrophic risk of weather variability could be covered by the indexed product while the poolable risk components could be covered through additional risk management strategies tailored to the individual (Skees and Barnett; Skees, Barnett, and Hartell; Varangis, Skees, and Barnett). These advantages also outline the disadvantages to users of index-based insurance. Index products are such that individual losses can occur without the triggering of payments from the policy or, conversely, it can be that payments flow from the index product even though the individual has not suffered a loss (Skees, Barnett, and Hartell). This problem captures the notion of basis risk, which is unavoidable in index-based products. Basis risk makes index-based weather derivatives more attractive to associations, industries, or institutions whose risk exposure can be spread over larger geographic areas. Conversely, inherent basis risk is more problematic for individual purchasers whose risk exposure is more centralized (Varangis, Skees, and Barnett). Martin, Barnett, and Coble note that weather derivative basis risk may be reduced considerably through a portfolio holding of various weather derivatives based on several surrounding weather stations.

Rainfall Interpolation

There is an extensive literature focused on rainfall interpolation techniques. The simplest method is to set the value of rainfall at out-of-sample locations equal to the rainfall recorded at the nearest observed site (Thiessen). The National Weather Service developed another method in 1972 where rainfall was estimated as a weighted average of surrounding observed values, where the weights were inversely proportional to the squared distances from the unobserved site (Bedient and Huber). More recently, advances in the area of Geostatistics have created more statistically sophisticated interpolation methods through the use of kriging. Kriging, or optimal prediction, refers to the practice of making inferences on unobserved values of a random process given data generated from the same process (Cressie). In practice, kriging techniques form a predictor which is equal to a weighted average of the data in the sample. The weights used in the average are determined from the correlation structure of the process which may be given, assumed, or estimated from the data. Kriging techniques have been rigorously shown to provide predictors which are not only unbiased, but also efficient linear estimators.

Cressie discusses various types of kriging, which differ with respect to the underlying assumptions for the stochastic process. In general, the stochastic process of interest is modeled as the sum of a mean and a spatially correlated error component. Bayesian kriging assumes that the mean and error components are random and independent. Given appropriate priors for the parameters of the mean and error structure components, the optimal predictor for unsampled locations can be found and has been shown to be superior to other kriging methods (Cressie). While point estimates for the conditional means and variances of the process of interest can be derived explicitly given appropriate distributional assumptions (e.g. Kitanidis), an alternative approach is to sample directly from the posterior distribution of interest using Markov Chain Monte Carlo techniques. MCMC methods are often employed when the calculation of interest is that of a complex and high dimensional integral function. When explicit evaluation of these integrals is not possible, MCMC techniques provide an alternative to more traditional numerical or analytic integration methods (Brooks). By specifying starting values and non-informative priors for the variables of interest, MCMC methods use a variety of updating techniques to generate Markov chains of independent samples which converge, at least asymptotically, to the true posterior distribution of the variable of interest. Point estimates of interest are simply computed as sample moments from the sampling distributions. The transition kernels vary, and are defined by the updating schemes employed. Samplers based on the Gibbs and Metropolis-Hastings updating kernels are the most commonly used in MCMC applications. A full discussion of MCMC methods is beyond the scope of this paper. Please refer to Brooks and Gilks, Richardson, and Spiegelhalter for more detailed descriptions of the theory behind MCMC methods and implications for empirical implementation.

While kriging methods provide statistically attractive properties, they can also require a significant amount of computing time and effort. Thus, many studies have focused on the comparison of kriging to the simpler interpolation approaches. While many authors have shown that kriging techniques provide better estimates than simpler methods (Tabios and Salas, and Phillips et. al), others have found that the results depend critically on the density of the sampled locations. Dirks et. al concluded that kriging methods did not provide significantly better estimates than simpler methods, such as inverse distance weighting. In general, studies have shown that kriging dominates the simpler interpolation methods for areas with smaller sampling densities while the methods are fairly equivalent for areas with sampling grids of higher density.

Data

State-level monthly precipitation totals for Iowa were obtained from the NOAA's National Climatic Data Center (NCDC). The historical series of precipitation totals for all sequential combinations of months were compared to historical per-acre hay returns to find the combination of monthly precipitation totals that were the most highly correlated with hay yields. The April through December time period showed the highest correlation between cumulative precipitation and hay yields for Iowa and was adopted as the coverage period for the insurance product. In addition to aggregated state-level data, the NCDC reports data from thousands of individual weather stations located throughout the country. The full data set of Iowa weather stations was condensed to exclude those weather stations that did not have complete precipitation records for the months included in the coverage period (April-December) for the entire 30-year period from 1973-2002. At the time of data collection the last monthly recording was for August 2003, hence the use of 1973-2002 data to calculate the 30-year average precipitation levels guaranteed by the policy. Given the data requirements, the number of usable weather stations was reduced to 67 in the state of Iowa. Figure 2 provides a map of the locations of the weather stations. The grid of 67 weather stations provides a relatively dense sampling grid in comparison to previous studies. The distance between adjacent weather stations averages 20 miles, with a maximum (minimum) distance between weather stations of 50 (7) miles.

Figures 3 and 4 map the means and standard deviations of reported precipitation levels, respectively, for the counties in which the weather stations are located. The weather station data shows that the northwest section of Iowa tends to be the driest with more precipitation, on average, being reported as you move into the southeast section of the state. Precipitation variability, as measured by the standard deviation of reported precipitation, follows a similar pattern across the state with lower variability in the northern section of the state and higher variability in the central and southern regions.

Two additional issues arose with the weather station data. First, for some stations and months only estimated precipitation values were available. These estimated values were assumed to be unbiased and were left unchanged. Second, for some other stations and months, the precipitation values were reported as incomplete. For these incomplete months, the NCDC indicated that somewhere between one and nine days of information were missing from the reported precipitation value. In order to conserve these data points, it was assumed that the incomplete months were missing the average of five days of precipitation information and that the precipitation amount during those five days was equal to the five day average precipitation amount for the month based on the reported total. The adjusted precipitation amount was set equal to the incomplete amount times the sum of one and the ratio of five and the number of days in the months less five. For example, if June was reported as incomplete with 2.5 inches of precipitation, the June precipitation was adjusted to 3.0 inches.

The coordinates of the geographical centers of each county in Iowa, measured in degrees of latitude and longitude, were calculated from a data file created by Giglierano and Madhukar. This yielded 99 county reference points, or sample "farms", where rainfall could be interpolated to rate the insurance policy. The geographic coordinates of each of the 67 weather stations in Iowa were obtained from the NCDC. The distance measure of one degree of latitude is relatively constant across the surface of the earth[3], and equal to roughly 111.3 kilometers (69.1 miles). The distance measure of one degree of longitude varies with location on the earth's surface. At any given point on the earth's surface, one degree of longitude, measured in latitude degrees, is equal to the cosine of the latitude coordinate of the given point. Thus, given the latitude and longitude coordinates of two points on the earth's surface, the distance in degrees of latitude can be computed using a measure of Euclidean distance in a plane[4].

Rainfall Model

Following Cressie and Kitanidis to derive an empirical Bayes predictor for rainfall, let y_i denote observed rainfall at weather station i and assume that the actual rainfall at a given site is determined by the sum of a mean or drift process, μ, and a spatial error process, ε, which are both functions of the site's geographic location, X, and unknown model parameters Θ_μ and Θ_ε, respectively.

$$y_i = \mu\left(X_i, \theta_\mu\right) + \varepsilon\left(X_i, \theta_\varepsilon\right) \tag{1}$$

Using Baye's Rule, and denoting all model parameters by Θ, the posterior distribution for the estimated model parameters is given by

$$p\left(\theta \mid y\right) = \frac{p(y \mid \theta)p(\theta)}{p(y)} = \frac{p(y \mid \theta)p(\theta)}{\int p(y \mid \theta)p(\theta)d\theta} \quad . \tag{2}$$

For any unobserved site *j*, the distribution of estimated rainfall, $\tilde{y}_j$, conditional on observed rainfall at *N* locations is given by

$$\begin{aligned} p(\tilde{y}_j \mid y_1.....y_N = y) &= \int_\theta p(\tilde{y}_j, \theta \mid y)d\theta \\ &= \int_\theta p(\tilde{y}_j \mid \theta, y)p(\theta, y)d\theta \qquad (3) \\ &= E_{p(\theta|y)}\left\lfloor p(\tilde{y}_j \mid \theta, y)\right\rfloor. \end{aligned}$$

Thus, the posterior distribution for $\tilde{y}_j$ given *y* is taken as the expected value of the posterior given *y* and *Θ* with respect to the posterior distribution of *Θ* given *y*. MCMC methods can be used to simultaneously generate Markov Chains of the both model parameters from the posterior distribution in equation 2, and rainfall estimates for any number of unobserved locations from the posterior given in equation 3. Given *J* unobserved sites and *P* model parameters, the order of integration for a given unobserved site is *N+P-1*.

To estimate the model the structure of the mean and error processes must be specified. A linear model was chosen for the mean process due to the relationship between average rainfall and geographic location exhibited in figure 3. After examination of the historical correlations between recorded rainfall against the distance between the weather stations, an exponential correlogram was chosen to represent the error structure.

$$\mu(X_i, \theta_\mu) = \beta_0 + \beta_{lat} lat_i + \beta_{long} long_i \qquad (4)$$

$$\Sigma_{ij} = f(d_{ij}, \varphi, \kappa) = \exp\left(-\left(\varphi d_{ij}\right)^\kappa\right) \qquad (5)$$

The correlation of rainfall at locations *i* and *j*, Σ_{ij}, was specified as a function of the Euclidean distance, d_{ij}, between the two locations. The parameters κ and φ are measures of spatial smoothing and decay, respectively. The smoothing parameter, κ, is bounded between zero and two with larger values indicating higher levels of spatial smoothing. A value of κ equal to two implies the Gaussian correlation function. The decay parameter, φ, is bounded below at zero and indicates the degree of decline in correlation between two locations with distance. A larger (smaller) value of φ indicates a faster (slower) decline in correlation as distance increases (Thomas et. al). Thus, larger estimates for phi indicate a smaller degree of similarity between nearby stations. Given the specifications in equations 4 and 5, the model consists of 5 parameters. Including the 99 sample farms as rainfall estimation points causes the order of integration for the expectation summarized in equation 3 to equal 103. Thus, the computing requirements to carry out the kriging estimation were expected to be quite large.

The IDW method estimates rainfall at an unobserved site as the weighted average of the observed rainfall at the weather stations, where the weights (λ_{ij}) are the inverse distance between the unobserved site and the weather station (d_{ij}), normalized by an appropriate constant. While the IDW method lacks the robust statistical properties of the estimates obtained from kriging, it also requires significantly less computing power and time.

$$\tilde{y}_j = \sum_{i=1}^{N} \lambda_{ij} y_i \text{ , where} \tag{6}$$

$$\lambda_{ij} = \frac{d_{ij}^{-1}}{\sum_{i=1}^{N} d_{ij}^{-1}} .$$

Interpolation Results

Kriging

The WinBUGS software package was used to specify and estimate the rainfall model. For each year in the data a sample from the posterior distributions of each model parameter and rainfall for the 99 sample farms were generated. To save time, the program was set to estimate rainfall for each sample site individually, reducing the order of integration to five[5]. The latitude and longitude coordinates for each of the weather stations and reference points were normalized to make the southwest corner of Iowa the grid origin. The sample autocorrelation plots from initial sample iterations exhibited autocorrelation out to roughly ten lags. To obtain a more independent sample, the chains were thinned to save every tenth iteration. To ensure convergence, three chains were run from different starting values, 5000 "burn-in" iterations were discarded, and the chains were run until the Monte Carlo error for the samples was less than 5% of the sample standard deviation[6]. The process yielded 5000 independent rainfall and parameter samples for each year in the data. The point estimates for unobserved rainfall at the reference points were taken as the sample means from the Markov Chains. The estimated 30-year means and standard deviations of precipitation are illustrated in figures 5 and 6, respectively. The rainfall patterns exhibited in figures 5 and 6 are very similar to those in the actual weather station data illustrated in figures 3 and 4. Average rainfall tends to fall as you move further north and west, while the standard deviation of rainfall is larger in the southern part of Iowa. Furthermore, cross-validation confirmed that the kriging results were statistically unbiased estimates of actual rainfall, while the average standard deviation of the bias estimates was 3.01 inches of rainfall. These results can be interpreted as upper bounds on the performance of the model as cross-validation effectively reduces the grid density.

The parameter estimates for the mean process and the correlogram are summarized in table 1. The complete Markov chains of the model parameters and precipitation estimates are available upon request from the author. The estimates for β_0 can be interpreted as a rainfall estimate for the southwest corner of the Iowa grid, and averaged just under 29 inches of rainfall which is consistent with the true 30-year means from weather stations in that region. The estimates for β_{lat} and β_{long} indicate that, on average, precipitation declines by 1.46 inches for every degree of latitude as you move north, and increases by 0.85 inches for every degree of longitude as you move east. These results are also consistent with the relationship between average rainfall amounts and location in the state of Iowa depicted in figure 3. The smoothing parameter, κ, ranged between 0.56 and 1.66, with an average value of 1.01. The decay parameter, φ, varied within a considerable range from 0.48 to 11.34, with an average value of 3.58. Larger estimates of φ indicate a weaker spatial correlation structure in the rainfall data for the given year. Thus it was expected that the estimation bias would be larger for years with larger φ estimates. Using cross-validation, it was found that the absolute bias was in fact positively correlated with the absolute bias of the rainfall estimates, although the level of correlation was rather low at 0.30.

Figure 7 plots the actual correlations between the rainfall records from the weather stations against the distance between the stations. Also included in figure 7 are the correlograms implied by a smoothing parameter, κ, equal to one and a range of values for the decay parameter. The correlogram with φ set to 0.5 seems to match the historical correlation structure quite well, while the correlograms while larger rates of decay tend to underestimate the historical correlation in the data. However, the parameter estimates for each year are estimated from the cross-section of data for the given year only and can vary depending on the similarity between measurements at differing distances. The correlation structure estimated for any given year is not constrained to match the average historical relationship.

Inverse-Distance Weighting

Again using cross-validation, precipitation estimates were calculated using IDW for the weather station sites for each year over the period 1973-2002 covered in the data set. The precipitation estimates were calculated using from one to the entire set (66) of the nearest weather stations to the station sites and compared to the actual precipitation values recorded at the stations. Figure 8 shows the average and standard deviation of the bias estimates when using one to ten weather stations in computing the IDW rainfall estimate. The minimum bias is achieved when the nearest four weather stations are used to estimate precipitation at the non-sampled site. However, none of the average bias estimates for the individual example sites were statistically different from zero at standard significance levels when any amount of weather stations, from one to the entire sample, were used in computing the precipitation estimates.

There seems to be a fairly significant decrease in the average standard deviation across the example sites as the number of stations used in the estimate increases from one to four or five, with the average standard deviation leveling off as additional stations are included. Thus, while the use of only one weather station may provide an unbiased estimator of the precipitation at a non-sampled site, incorporating additional weather stations seems to provide gains in efficiency for the precipitation estimate.

Given the information in figure 6, the four nearest weather stations were used under the IDW interpolation method. The 30-year means and standard deviations of rainfall for each of the 99 sample farms were computed. A comparison between the kriging results showed that the two methods were nearly identical. The largest difference between the 30-year averages was found to be less than 0.7 inches, while the largest difference in the standard deviation of rainfall was less than 0.7 inches. Moreover, the coefficients of variation implied by the estimates of the two methods differed by less than 2%, implying that the insurance rates calculated from the estimates of either method will be nearly identical.

Contract Structure

The rainfall guaranteed under the policy is computed as the 30-year average of recorded precipitation for the area over the insurance period, which is patterned after the 30-year climate normals used by the NCDC. Equations 7 and 8 outline the liability (L) and indemnity (I) structures adopted for the example policy. The indemnity takes the form of an exotic put option on the 30-year average rainfall guarantee. Indemnities equal the losses resulting from precipitation shortfalls if they occur and are capped at the total liability level insured. The indemnity structure is similar to an example outlined by Martin, Barnett and Coble.

$$L = 0.53 * P_{H,10} * Y_{H,10} * A \tag{7}$$

$$I = Max\left[0, Min\left(\left(L * F * \left(C - \frac{R_A}{R_{30}}\right)\right), L\right)\right] \tag{8}$$

where
- A = total insured acres
- $P_{H,10}$ = 10-year average hay price (\$/ton)
- $Y_{H,10}$ = 10-year average hay yield (tons/acre)
- F = indemnity factor
- C = coverage level ($C \in [0,1]$)
- R_A = actual rainfall
- R_{30} = 30-year average rainfall, or the rainfall guarantee

Indemnities are triggered when actual precipitation is less than a selected percentage (the coverage level, C) of the historical average precipitation. The percentage shortfall in precipitation is translated into a shortfall in liability value, and the indemnities paid are equal to the liability shortfall. The insurance liability was taken as the product of moving averages of hay prices and yields for the state of Iowa[7]. The liability is set equal to the product of 10-year moving averages of Iowa hay prices and yields to establish the expected per-acre value for the product liability. The liability is then multiplied by 0.53 to adjust the liability value for pasture[8]. The data for the hay prices and yields were obtained from the United States Department of

Agriculture, National Agricultural Statistics Service (USDA-NASS). More disaggregated data is available for hay yields at the crop reporting district and county levels, but prices are only consistently reported for the states. The 10-year averages for yields and prices were chosen to provide a representative example of per-acre liability to rate the product. Additionally, the 10-year moving averages mimic current crop yield insurance rules for setting yield guarantees on individualized FCIC insurance products offered in the U.S.

The indemnity factor, *F*, was created to translate precipitation shortfalls into liability value shortfalls. A regression relating precipitation levels to hay yields was estimated for Iowa. To put all variables on a percentage basis, ratios were created for each variable. The precipitation ratio (*RR*) is the ratio of the current year's precipitation to the 30-year average. The hay yield ratio (YR_H) is the ratio of the current year's reported hay yield to the 10-year average hay yield. Table 1 reports the regression estimates.

$$YR_H = \alpha + \beta * RR + \varepsilon \tag{9}$$

The sign of the estimated slope coefficient was as expected, with precipitation shortfalls leading to a reduction in hay yields below the average level. This result is consistent with Turvey's findings of strong relationships between precipitation levels and hay yields. The results exhibit fairly strong yield movements in Iowa, with a one percent drop in precipitation from the 30-year average resulting in a 1.52 percent drop in hay yields below the 10-year average hay yield. The indemnity factor (*F*) was taken as the slope coefficient estimate (1.52). Thus the policy pays 1.52 percent of the liability for every one percent drop in precipitation below the guaranteed historical average.

Insurance Rates

To rate the insurance policy, Monte Carlo analysis was used under three different alternative approaches. Gamma distributions were fit to the historical rainfall means and standard deviations implied by 1) the kriging estimates for the 99 reference points, 2) the IDW estimates for the 99 reference points (IDW_1), and 3) the actual histories for the 67 weather stations (IDW_2). For each method, 5000 random draws were taken from each of the specified gamma distributions. The policy was then rated by taking the average indemnity value over the 5000 rainfall draws for each of the 99 reference points. While the first two methods use estimated histories to generate precipitation draws for the 99 reference points directly, the third method uses the actual rainfall histories to generate random precipitation draws at the weather stations for use with IDW to rate the policy. For the IDW_2 method, IDW was used to evaluate the indemnity value for each reference point over the correlated weather station draws. The historical correlation structure of the historical data was imposed on the set of 67 weather station rainfall draws using a method outlined in Iman and Conover. The Iman and Conover procedure has four attractive properties. First, the procedure works well with any distribution function. Second, the mathematics behind the procedure are not extremely complex as cholesky factorization and matrix inversion are the most exotic steps in the procedure. Third, the procedure can be used under any sampling scheme. Finally, the marginal distributions of interest are maintained throughout the procedure in that the moments of the marginal distributions are not affected by the procedure. A more detailed description of the procedure is included in the Appendix and a MatLab program which implements the algorithm is available from the authors.

The choice of gamma distributions was based on the prevalence of this distributional choice for precipitation in the scientific and agricultural literature (Barger and Thom; Thom;

Ison, Feyerherm, and Bark; Martin, Barnett, and Coble; Groisman et al.). The gamma distribution is bounded from below at zero and can represent skewed data, which makes it appropriate for precipitation modeling. The gamma distribution is defined by two shape parameters which are functions of the distribution's mean and variance.

Using the method proposed by Moschini, nonparametric kernel densities were fit to each of the 30 year precipitation histories for the weather stations and compared with the gamma distributions implied by the sample moments. Although no statistical tests were performed, the gamma density plots were quite similar to the nonparametric estimates and seemed to provide an excellent fit to the data. The gamma distribution and the nonparametric density for the Chariton weather station are illustrated in Figure 9.

The liability value (152 $/acre) was taken as constant and equal to the product of the 2002 10-year average hay yield (3.27 Tons/acre) and price (87.7 $/Ton) for Iowa as reported by NASS, multiplied by a factor of 0.53. The number of insured acres was set to unity to create per-acre premiums that could be scaled to any level of coverage. As expected, the premiums calculated under each of the three methods were nearly identical. Only the premiums calculated from the kriging estimates are reported for convenience. The full set of premiums and rates calculated under each method are available upon request from the author.

Iowa premiums average $18.91 per acre under full coverage, and $13.39 and $5.70 at 95% and 85% coverage, respectively. The average premium across the Iowa reference points is equal to $1.85 per acre for 75% hay coverage, with a standard deviation of $0.65. The premiums seem to be unrestrictive, especially at lower coverage levels which would be expected to be offered to provide drought risk coverage. At 75% coverage, the highest premium is $3.64 in Southeast Iowa at the Taylor county reference point, while the lowest premium is $0.46 in

Northeast Iowa at the Clayton county reference point. These results are expected as the lowest implied coefficient of variation (15.8%) is at the Clayton County reference point, while the largest implied CV (24.6%) is at the Taylor County reference point. Figure 8 maps the premium levels at a 75% coverage level. In general, premium levels are the lowest in the Northeast section of the state, with areas of relatively larger premium levels located in various locations throughout the state. Table 3 reports the premium rates, as a percentage of the liability insured, at 100, 95, 85, and 75 percent coverage levels for each of the 99 county reference points within Iowa for the policy.

Historical Analysis

An historical analysis of the insurance policy was constructed for the 1995-2004 contract years. As an example, the 2003 contract year uses the 2002 30-year average precipitation levels (estimates) as the historical average precipitation level in the indemnity formula. The estimated actual precipitation levels for 2003 were used to calculate indemnity payments based on the 2002 30-year average rainfall levels for each county reference point. Precipitation estimates were taken from the kriging results. Using USDA-NASS data, 10-year averages of hay yields and prices in Iowa were calculated for 1994-2003 to compute liability levels for each year, as outlined in the previous section.

Figure 10 maps indemnity payments at 75% coverage for the 2002 contract year at the county reference points in Iowa. Precipitation was below 75% of the 30-year average for a pocket of counties in Southwestern Iowa, causing the product to trigger indemnity payments. In counties where indemnities were triggered, payments ranged from \$0.94 per acre in Mahaska and Page Counties to \$15.16 per acre in Ringgold County.

Figure 11 maps indemnity payments at 75% coverage across Iowa for the 2000 contract year. Indemnity payments were triggered in a large pocket of counties in the Western and Southwestern portions of the state of Iowa. Indemnities, for triggered counties, ranged from a low of $0.64 per acre in Page County to a high of $36.46 per acre in Carroll County. Historical premium and indemnity levels and maps for other contract years included in the analysis and for coverage levels above 75% are available upon request from the authors.

Table 5 reports the average premiums, indemnity payments, and loss ratios at 75% coverage across the entire state of Iowa for each contract year included in the historical analysis. The loss ratio is the ratio of premium to indemnities and should average one over time if the policy is actuarially fair. Table 6 reports the same information for the counties where losses occurred. No losses were triggered at any of the county reference points in the 1995, 1996, 1998, 2001, and 2004 contract years yielding zero loss ratios. Indemnities were triggered at 16 of the county reference points for the 1997 contract year, with the average indemnity (loss ratio) in the triggered counties equal to $8.36 (5.51). The main loss region for the 1997 contract year was the Northwestern part of Iowa. There was one county reference point (Davis County) with losses in the 1999 contract year with an indemnity payment (loss ratio) equal to $5.15 (2.38). The 2000, 2002, and 2003 contract years yielded 20, 8, and 13 loss counties, respectively. Average indemnities (loss ratios) were $12.98 (6.92) in 2000, $6.50 (3.02) in 2002, and $7.55 (4.82) in 2003. The Northeastern and Northcentral regions of Iowa were the loss areas for the 2003 contract year.

For all contract years in which losses occurred, with the exception of 2003, the counties in which indemnities were triggered were high risk areas relative to the state average as the average premium rates (and thus CV's) were higher than the overall state average for the same contract year. This can be seen by comparing the average premium levels reported for the entire state of Iowa in table 5 and the average premiums in the loss counties in table 6.

Indemnity payments, when triggered, tend to be quite large relative to the per-acre premium rates for the associated area. In general, the policy tends to pay indemnities in concentrated areas and at fairly high loss ratios. At higher coverage levels the loss regions expand to cover larger and more general areas across the state. These results are expected given the spatial nature of weather events. While the policy is theoretically rated to yield a loss-ratio of unity over time for any given location, the systemic nature of weather risk requires a large geographic area of coverage to provide proper risk pooling and insurability for any given year.

Discussion and Conclusions

Markets for weather-based financial tools have realized extensive growth over the past decade. There are currently markets for temperature-based weather derivatives traded on the CME as well as more personal markets for OTC weather derivatives exchanged in the form of weather swaps and options. While the market for weather derivatives based on temperature indexes has grown significantly, the market for precipitation based derivatives is still in its infancy. Weather basis risk and sources of accurate and reliable data for pricing, especially for developing countries, seem to be the largest obstacles to further growth in the weather derivative market.

This paper has outlined a potential drought insurance policy for pasture owners in the state of Iowa. The policy provides coverage against precipitation shortfalls below a coverage threshold of the historical 30-year average rainfall for the area. The rainfall interpolation techniques focused on addressing the first component of weather derivative basis risk. Rainfall at a measurement site may not be the same as rain falling on the farmer's fields. Using a

Bayesian derived estimator, precipitation estimates for 99 sample farms in Iowa were obtained from a spatial kriging model using Markov chain Monte Carlo methods within the WinBUGS software package. The kriging estimates were then compared to a simpler inverse distance weighting estimator. The rainfall interpolation techniques focused on addressing the first component of weather derivative basis risk. Consistent with previous studies, the results were found to be nearly equivalent. Using cross-validation, both interpolation methods were shown to provide unbiased estimates.

Monte Carlo analysis was performed to calculate fair premium rates for the 99 county reference points in the state of Iowa at various coverage levels by specifying the indemnity structure of the policy as an exotic put option on rainfall. The policy was rated using three alternative methods. The first used the estimated rainfall histories from kriging, while the second used the estimated histories from the inverse distance weighting method. Finally, the actual histories from the weather stations were used to specify gamma distributions of rainfall and then inverse distance weighting was used to form the rainfall draws for each sample site. Again, the methods were found to provide nearly equivalent results with respect to premium levels and rates.

An historical analysis was also performed to assess potential product performance if the policy were marketed. The product was shown to successfully trigger losses in regions of abnormally low precipitation reported at local weather stations across Iowa over a 10-year range of contract years. Given the systemic nature of weather events such as precipitation, loss areas over the period analyzed tended to be geographically concentrated exhibiting high loss ratios. While the policy is fairly rated for any given location over time, a sufficiently large geographic coverage area would generally be required for sufficient risk pooling in a given contract. Thus, a drought insurance policy such as this may be more suited for administration under large institutions, associations, or government agencies rather than smaller private companies. Moreover, the extensive weather data available from the NCDC for other states in the U.S. from the NCDC should allow similar derivatives to be developed to cover a variety of weather-related losses in various locations. Furthermore, the rainfall interpolation methods utilitized in this study could be widely applied to data in other areas, to develop other types weather derivatives including insurance for agriculture. The fact that weather-based products fall under the umbrella of index coverage causes the administrative costs to be low relative to other insurance program types, providing even more potential for developing nations. However, the feasibility of weather-based insurance offerings hinges greatly on the reinsurance capacity available to insurers (Duncan and Myers).

An issue which was not addressed is the possibility of an adverse selection advantage created by the use of long-term weather forecasting models, such as the El Nino/Southern Oscillation Index. Over the last 20 years, the ENSO and the weather phenomena associated with it (El Niño and La Niña) have become a common part of the agriculture vocabulary. Farmers around the nation track the ENSO to gauge the likelihood of long-term weather patterns in their area. For Iowa, El Niño seasons typically are wetter than usual (sometimes extremely so), while La Niña seasons can be anywhere from extremely dry to near normal. There are many places producers can find information on the latest ENSO levels and forecasts, including the NOAA Climate Prediction Center. However, the International Research Institute for Climatic Prediction reports that while forecasting El Niño and La Niña episodes from the early part of the summer is not difficult, it is quite difficult to accurately forecast cycles during the months of January through April. This phenomenon is referred to as the "spring barrier" in the Northern

Hemisphere. Thus, forecasting through the use of ENSO forecasts to arbitrage this product would be limited as long as sales closing dates were held in the late winter or early spring periods. Additionally, this problem could potentially be addressed through an adjustment to premium rates, or alternatively an adjustment to the 30-year precipitation average guaranteed, from year to year based on the forecasts of these long term models. However, other existing insurance policies which cover production yield levels face this same type of adverse selection risk as farmers may be more prone to purchasing yield insurance in forecasted drought years. Incorporating seasonal weather forecasts is also an area with potential for further research.

References

Agricorp. "2004 Forage Rainfall Plan." Accessed online at http://www.agricorp.com/uploaddocuments/d601+2004%20Forage%20Rainfall.pdf, March 24, 2004.

Ali, P.U. 2004. "The Legal Characterization of Weather Derivatives." *The Journal of Alternative Investments*, Fall 2004: 75-79.

Barger, G. and H. Thom. 1949. "Evaluation of Drought Hazard." *Agronomy Journal*, Volume 41: 519-526.

Bedient, P.B. and W.C. Huber. 1992. *Hydrology and Floodplain Analysis, Second Edition*. Addison-Wesley, New York, NY.

Brooks, S.P. 1998. "Markov Chain Monte Carlo Method and Its Application." *The Statistician*. Volume 47: 69-100.

Cao, M., A. Li, and J. Wei. 2004. "Precipitation Modeling and Contract Evaluation: A Frontier in Weather Derivatives." *The Journal of Alternative Investments*, Fall 2004: 93-99.

Cressie, N.A.C. 1993. *Statistics for Spatial Data, Revised Edition*. Wiley Series in Probability and Mathematical Statistics. John Wiley and Sons, Inc.

Dirks, K.N., J.E. Hay, C.D. Snow, and D. Harris. 1998. "High-Resolution Studies of Rainfall on Norfolk Island Part II: Interpolation of Rainfall Data." *Journal of Hydrology*. Volume 208:187-193.

Dischel, B. 2000. "Is Precipitation Basis Risk Overstated?" *Weather Risk*, supplement to *Risk Magazine* and *Energy & Power Risk Management*, August 2000.

Duncan, J. and R.J. Myers. 2000. "Crop Insurance under Catastrophic Risk." *American Journal of Agricultural Economics*, Volume 82: 842-855.

Gautam, M., P. Hazell, and H. Alderman. 1994. "Rural Demand for Drought Insurance." Policy Research Working Paper 1383, World Bank, Washington D.C., November 1994.

Giglierano, J.D. and M. Madhukar. "State and County Boundaries of the State of Iowa." A vector digital data file accessed at <ftp://ftp.igsb.uiowa.edu/gis_library/IA_State/Admin_Political_Boundary/county.zip>, November 1st, 1990.

Gilks, W.R., S. Richardson, and D.J. Spiegelhalter (Eds.). 1996. *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London, UK.

Groisman, Pavel, Thomas Karl, David Easterling, Richard Knight, Paul Jamason, Kevin Hennessy, Ramasamy Suppiah, Cher Page, Joanna Wibig, Krzysztof Fortuniak, Vyacheslav Razuvaev, Arthur Douglas, Eirik Forland, and Pan-Mao Zhai. 1999. "Changes in the Probability of Heavy Precipitation: Important Indicators of Climatic Change." *Climatic Change* 42: 243-283.

Iman, R.L. and W. J. Conover. 1982. "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables." *Communication Statistics: Simulation and Computation*, 11: 311-334.

International Research Institute for Climatic Prediction. Accessed online at http://iri.columbia.edu/climate/ENSO/background/prediction.html, August 2005.

Iowa State University Extension. "Cash Rental Rates for Iowa, 2005 Survey." Accessed online at http://www.extension.iastate.edu/Publications/FM1851.pdf, November 2005.

Ison, N., A. Feyerherm, and L. Bark. 1971. "Wet Period Precipitation and the Gamma Distribution." *Journal of Applied Meteorology* 10: 658-665.

Jewson, S. 2004. "Introduction to Weather Derivative Pricing." *The Journal of Alternative Investments*, Fall 2004: 57-64.

Kitanidis, P.K. 1986. "Parameter Uncertainty in Estimation of Spatial Functions: Bayesian Analysis." *Water Resources Research*. Volume 22: 499-507.

Martin, Steven, Barry Barnett, and Keith Coble. 2001. "Developing and Pricing Precipitation Insurance." *Journal of Agricultural and Resource Economics,* Volume 26: 261-274.

McCarthy, N. 2003. "Demand for Rainfall-Index Based Insurance: A Case Study from Morocco." Environment and Production Technology Division Discussion Paper No. 106, International Food Policy Research Institute, Washington, DC.

Moschini, G. 1990. "Nonparametric and Semiparametric Estimation: An Analysis of Multiproduct Returns to Scale." *American Journal of Agricultural Economics*, Volume 72: 589-596.

NOAA-NCDC. 2003. National Oceanic and Atmospheric Administration, National Climatic Data Center. Accessed online at http://nndc.noaa.gov/.

Plate, Alice. "Rainfall Insurance Policy to Offer Better Farmers More Security." Australian Broadcasting Corporation. Accessed online at http://www.abc.net.au/riverland/stories/s1037125.htm, March 24, 2004.

Sakurai, T. and T. Reardon. 1997. "Potential Demand for Drought Insurance in Burkina Faso and its Determinants." *American Journal of Agricultural Economics* 79: 1193-1207.

Saskatchewan Crop Insurance. "Forage Rainfall Insurance Program (FRIP)." Accessed online at http://www.saskcropinsurance.com/programs/2004/forage/FRIP.shtml, March 24, 2004.

Skees, J., B. Barnett, and J. Hartell. 2005. "Innovations in Government Responses to Catastrophic Risk Sharing for Agriculture in Developing Countries." Paper prepared for workshop *Innovations in Agricultural Risk Management in Central America: Challenges and Opportunities to Reach the Rural Poor*, Antigua, Guatemala, May 9-12, 2005.

Skees, J., S. Gober, P. Varangis, R. Lester, and V. Kalavakonda. 2001. "Developing Rainfall-Based Index Insurance in Morocco." Policy Research Working Paper 2577, World Bank, Washington D.C., April 2001.

Skees, J. and B. Barnett. 1999. "Conceptual and Practical Considerations for Sharing Catastrophic/Systemic Risks." *Review of Agricultural Economics* 21: 424-441.

Skees, J., P. Hazell, and M. Miranda. 1999. "New Approaches to Crop Yield Insurance in Developing Countries." International Food Policy Research Institute, Environmental and Production Technology Division Discussion Paper No. 55, November 1999. http://www.ifpri.org/divs/eptd/dp/papers/eptdp55.pdf

Spiegelhalter, D., A. Thomas, N. Best, & D. Lunn. 2003. *WinBUGS User Manual, Version 1.4*. User manual accompanying the WinBUGS 1.4.1 software installation package.

Tabios, G.Q., and J.D. Salas. 1985. "A Comparative Analysis of Techniques for Spatial Interpolation of Precipitation." *Water Resources Bulletin*. Volume 21: 265-380.

Thiessen, A.H. 1911. "Precipitation Averages for Large Areas." *Monthly Weather Review*. Volume 39: 1082-1084.

Thom, H. 1958. "A Note on the Gamma Distribution." *Monthly Weather Review*, Volume 86: 117-121.

Thomas, A., N. Best, D. Lunn, R. Arnold, & D. Spiegelhalter. 2004. *GeoBUGS User Manual*. Available free online at http://www.mrcbsu.cam.ac.uk/bugs/winbugs/geobugs12manual.pdf, accessed January 2006.

Turvey, C.G. 2001. "Weather Insurance and Specific Event Risks in Agriculture." *Review of Agricultural Economics* 23: 333-351.

Turvey, C.G. 1999. "The Essential of Rainfall Derivatives and Insurance." Working Paper WP99/06, University of Guelph, Guelph, Ontario, April 1999.

USDA-NASS. 2003. United States Department of Agriculture, National Agricultural Statistics Service. Accessed online at http://www.usda.gov/nass/.

Van Lennep, D., T. Oetomo, M. Stevenson, and A. de Vries. 2004. "Weather Derivatives: An Attractive Additional Asset Class." *The Journal of Alternative Investments*, Fall 2004: 65-74.

Varangis, P. "Weather-based Index Insurance." Accessed online at http://www.proventionconsortium.org/files/wharton_010801/varangis.pdf, March 24, 2004.

Varangis, P., J. Skees, and B. Barnett. 2002. "Weather Indexes for Developing Countries." In *Climate Risk and the Weather Market: Financial Risk Management with Weather Hedges*. Edited by R. Dischel. London: Risk Books.

Varangis, P. 2001. "Innovative Approaches to Cope with Weather Risk in Developing Countries." *The Climate Report* 2(4).

WinBUGS 1.4.1. 2004. Software package available for free download online at http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml, accessed January 2006.

Appendix

Imposing Correlation

The procedure is based on rank correlations. The rank correlation (r_s), also known as Spearman's rho, for a given set of paired data (x_i, y_i) is calculated by ranking the x's and y's among themselves, from high to low (or low to high), and then substituting into the following formula

$$r_s(x,y) = 1 - \frac{6 * \sum_{i=1}^{n} d_i^{\,2}}{n(n-1)^2}. \tag{A1}$$

where d_i denotes the difference between the ranks assigned to x_i and y_i and n is the sample size. Iman and Conover point out that raw correlation numbers can be misleading when the underlying data is non-normal or contains outliers, which is why the rank correlations are used rather than the simple (Pearson) correlation measure.

The theoretical basis for the procedure is that given a random matrix A whose columns are assumed to have a correlation matrix I (the identity matrix) and a desired correlation matrix B, there exists a transformation matrix C such that the columns of AC' (where C' is the transpose of C) have a positive definite correlation matrix B. Since B is positive definite and symmetric, there exists a lower triangular matrix (the transformation matrix) C such that B = CC'.

Let X be an NXK matrix where each column contains random draws from a specific marginal distribution, N is the sample size, and K is the number of variables. In this setting, X is the matrix of independent random draws from gamma distributions for each weather station. The sample size for this analysis is N = 5000, while the number of variables is K = 67 individual weather stations. Let R be a matrix, of the same dimensionality of X, containing what Iman and Conover refer to as "scores." Iman and Conover suggest using ranks, random normal deviates, or van der Waerden scores ($\Phi^{-1}(i / N+1)$) where Φ^{-1} is the inverse of the standard normal distribution function, N is the number of draws (5000), and $i = 1, ..., N$) as possible scores. Furthermore, the correlation matrix for the columns of R is assumed to be equal to I (the identity matrix), meaning the elements of R are uncorrelated. Following Iman and Conover, van der Waerden scores are used to generate the matrix R in this analysis.

Define T to be the desired rank correlation matrix for a transformation (resorting) of X. In this setting, T is equal to the historical rank correlation matrix of reported rainfall data from the individual weather stations. Given T is positive definite and symmetric it may be written as T = PP', where P' is a lower triangular matrix. P, the transformation matrix, can be found using Cholesky factorization. The transformed matrix of scores, R* = RP', has a rank correlation matrix M which is approximately equal to the target rank correlation matrix T. By rearranging the columns of X into the same ranking as R*, the transformed X matrix has a rank correlation matrix equal to M, which is very close to the target correlation matrix T.

Some of the deviation of M from T is due to correlation among the columns of R, meaning the assumption of the correlation matrix for the columns of R to be equal to I does not hold[9]. Iman and Conover propose a variance reduction procedure to minimize the deviation of M from T. A matrix S is found, such that SDS' = T, where D is the actual correlation matrix associated with the columns of R. Cholesky factorization can then be used to find a lower triangular matrix Q, where D = QQ'. Therefore SQQ'S' = PP'. Obviously, one possible solution is that $S = PQ^{-1}$, where Q^{-1} denotes the inverse matrix of Q. Then the transformed matrix $R^*_B = RS'$ will have a correlation matrix exactly equal to T. Let the rank correlation matrix of R^*_B be equal to M_B. Comparing M_B to M and T, it is shown that M_B is a more accurate approximation to the target rank correlation matrix T. The variance reduction technique proposed by Iman and Conover was also utilized in this analysis.

Transforming the Historical Correlation Matrix

The relationship between the correlation and distance between any two weather stations was examined. It was found that the distance and correlation of reported precipitation between any two weathers stations were inversely related. Figure A1 plots the correlation values from the historical correlation matrix against the distance between the weather stations (in degrees latitude). The correlation and distance values are highly negatively correlated, with a simple correlation coefficient of -0.74. Initially, a linear regression model was fit by regressing the precipitation correlations on the distance between the stations. The linear model was fit in both an unrestricted fashion and also restricting the constant term to equal unity. In the unrestricted (restricted) case, the linear model's fitted correlation values tended to consistently under-predict (over-predict) the correlation values for stations whose distances were less than 1.25 degrees of latitude (roughly 87 miles).

Since information from four "local" weather stations was utilized in the analysis of each county reference point, the values under- or over-predicted by the linear specifications were precisely the correlation values that were the most critical (the maximum distance between any two stations used for the same reference point was found to be just under 90 miles). Therefore a restricted quadratic regression equation[10] was fit to the data. Figure A1 also plots the fitted relationships for the quadratic and both linear specifications. It is evident that the quadratic model (in red) falls within the center of the data for distances under 90 miles, whereas the unrestricted (restricted) model tends to fall below (above) the center of the data for stations less than 90 miles apart. Table A1 summarizes the regression coefficient estimates for each model estimated.

Using the quadratic specification and distance values from a weather station distance matrix, a transformation of the historical correlation matrix was created. This matrix did satisfy the condition of positive definiteness and was used with the 5000 draws for each of the 67 weather stations to impose the transformed target correlation structure. The target correlation matrix can be broken down into 99 *4X4* matrices that are critical to the analysis of each county reference point. To provide an example, the historical correlation and transformed target correlation matrices relevant to the Adair county reference point are provided below in tables A2 and A3 respectively. The transformed correlation values are all within 0.10 of the actual historical correlation values for the Adair county reference point. Table A4 reports the correlation matrix, relevant to the Adair county reference point, of the Monte Carlo precipitation draws after the target correlation structure was imposed using the Iman and Conover method. The largest deviation between the target and actual correlation matrices for Adair County is less than 0.02. The largest deviation between the target and actual correlation matrices over the entire range of entries is only 0.03. Thus the Iman and Conover method provides a very close approximation of the target correlation matrix to the actual correlation structure of the correlated Monte Carlo precipitation draws.

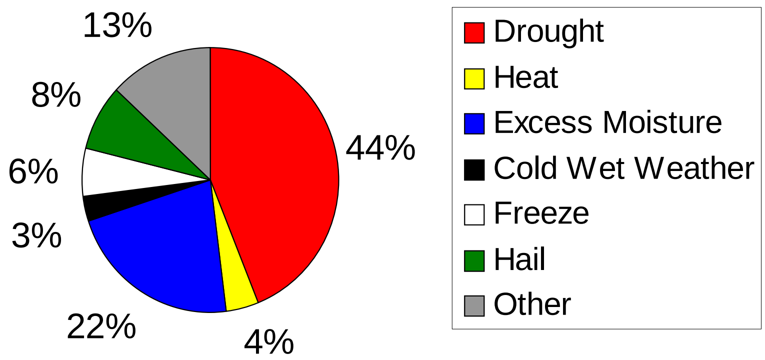

Figure 1. Causes of U.S. Crop Losses
Note: Based on crop loss data, provided by the Risk Management Agency of the USDA, from 1980-1999

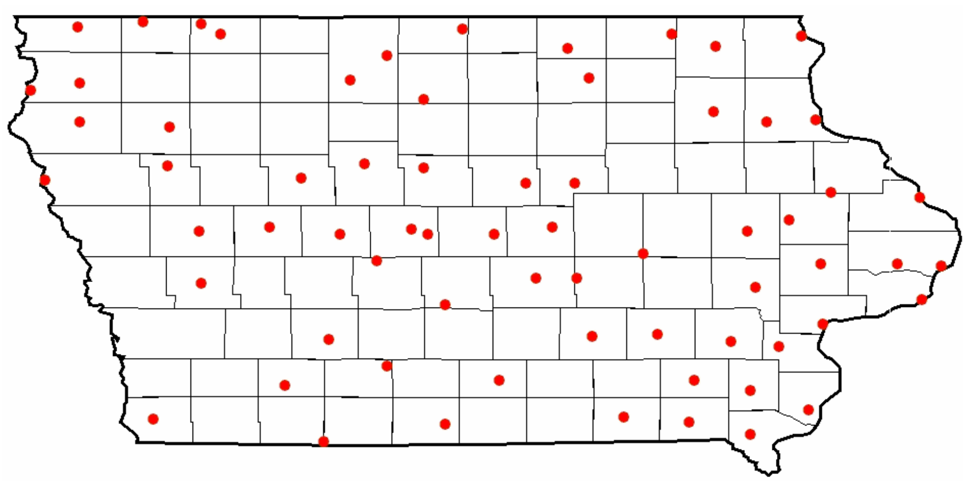

Figure 2. Iowa Weather Station Locations

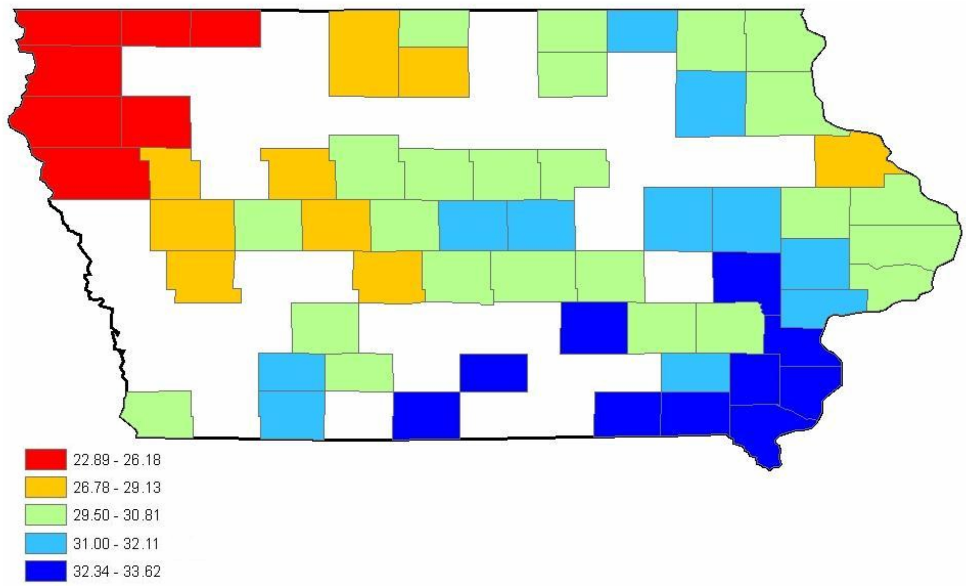

Figure 3. Reported Precipitation Means (Inches)

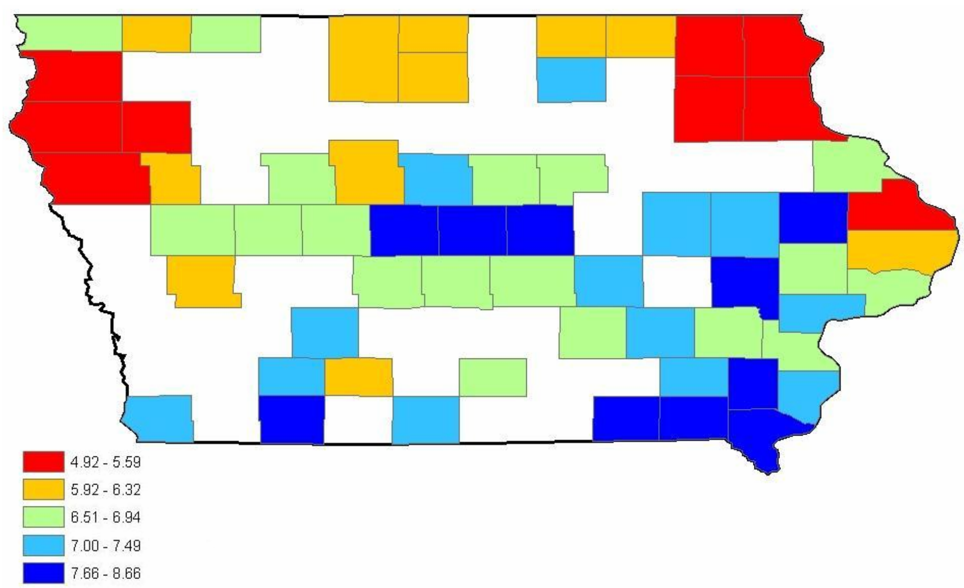

Figure 4. Standard Deviation of Reported Precipitation (Inches)

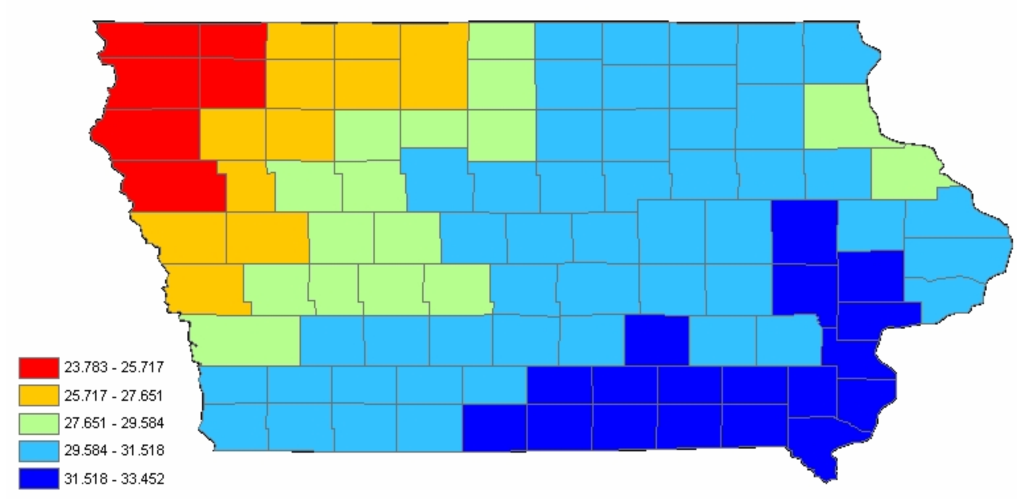

Figure 5. 30-Year Average Precipitation (inches), Kriging Estimates

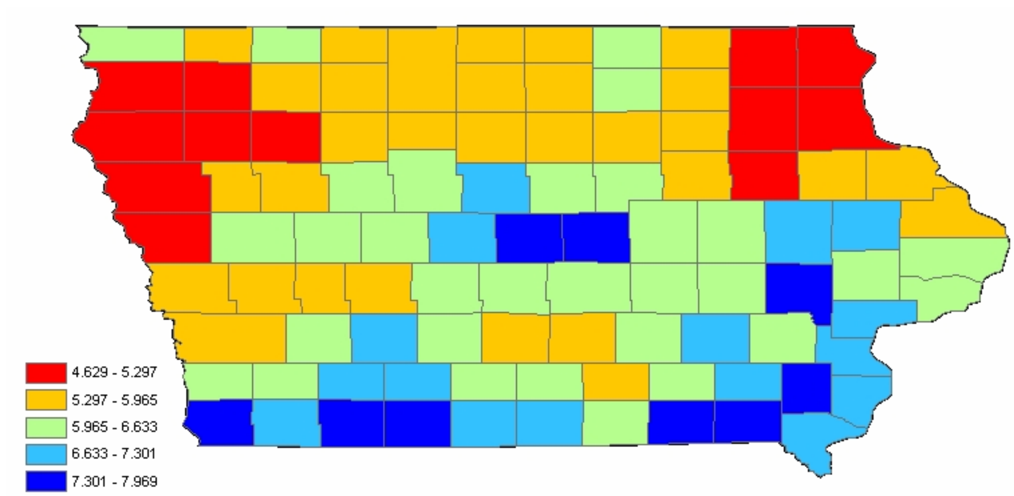

Figure 6. Precipitation Standard Deviation (inches), Kriging Estimates

Table 1. Summary Statistics of Rainfall Model Parameter Estimates

	β_0	β_{lat}	β_{long}	κ	φ
Mean	28.91	-1.46	0.85	1.01	3.58
Median	28.07	-1.89	0.59	0.97	2.03
Standard Deviation	6.89	1.90	0.96	0.24	3.37
Minimum	18.20	-4.70	-0.56	0.56	0.48
Maximum	43.16	2.74	2.85	1.66	11.34

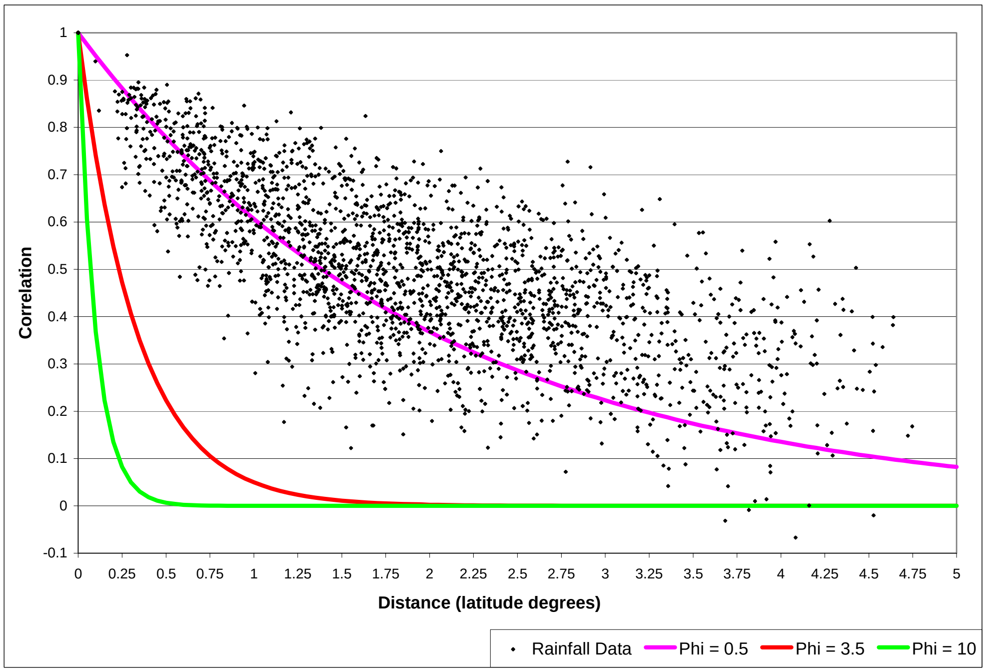

Figure 7. Correlation of Recorded Rainfall Against Distance Between Stations

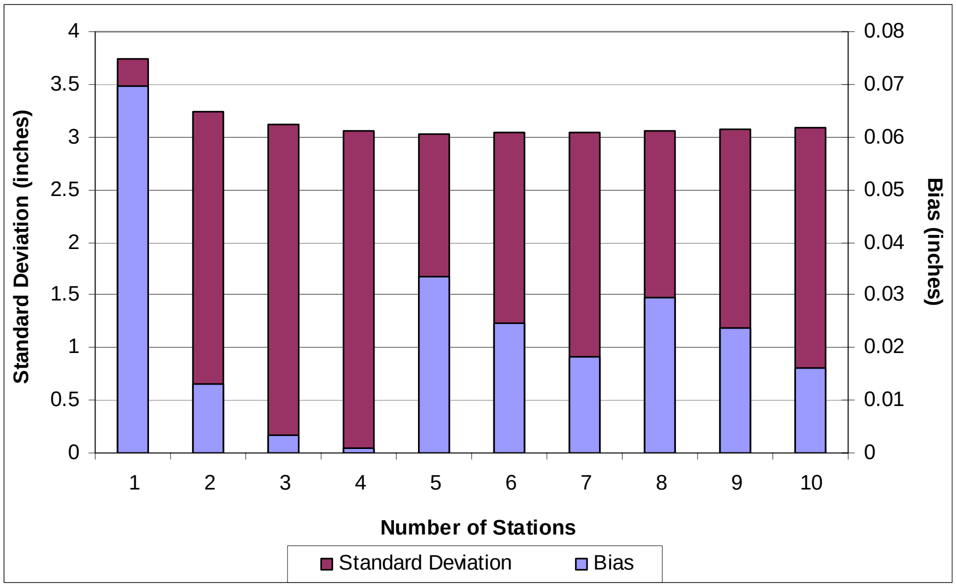

Figure 8. Bias and Standard Deviation of IDW Precipitation Estimates

Table 2. Regression Coefficient Estimates (Standard Errors).

Equation	$\hat{\alpha}$	$\hat{\beta}$	R^2
3	-0.34 (0.20)	1.52 (0.22)	0.85

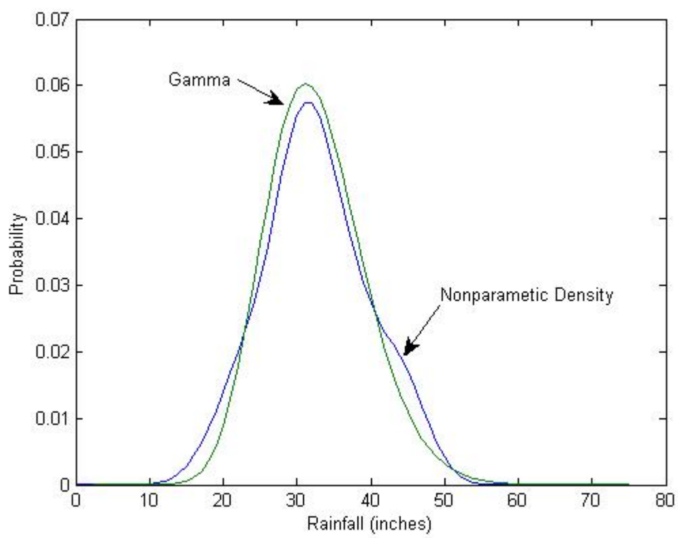

Figure 9. Gamma and Nonparametric Rainfall Densities, Chariton Weather Station

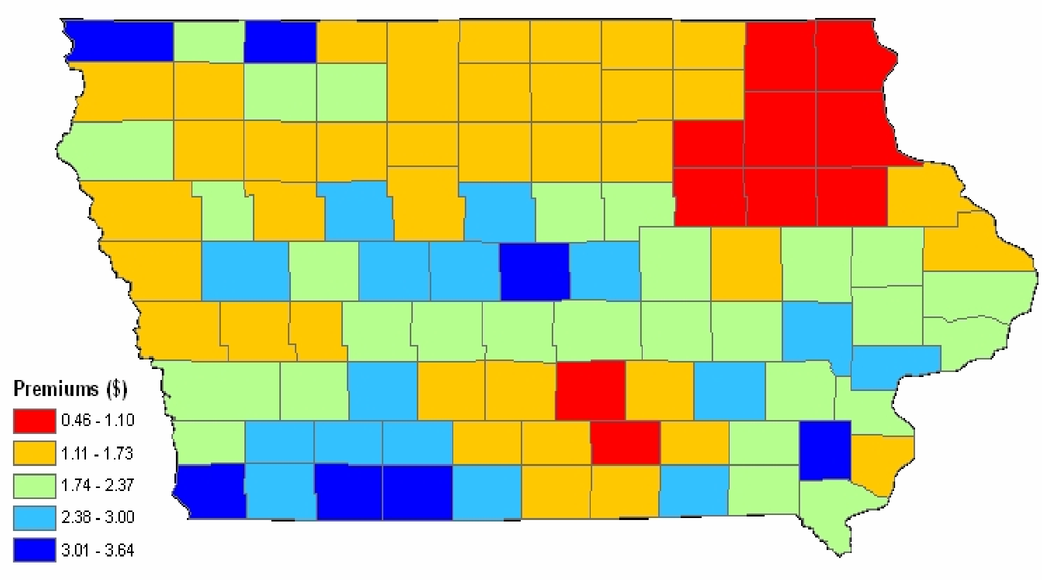

Figure 8. Map of Iowa Premiums ($/acre) at 75 Percent Coverage

Table 3. Iowa County Rates at Various Coverage Levels (%).

County	100%	95 %	85%	75%	County	100%	95 %	85%	75%
Adair	13.92	10.25	4.89	1.90	Jefferson	13.23	9.56	4.28	2.60
Adams	14.07	10.36	4.93	1.87	Johnson	14.49	10.73	5.10	3.24
Allamakee	9.69	6.25	2.01	0.40	Jones	13.24	9.55	4.22	2.53
Appanoose	12.21	8.57	3.53	1.10	Keokuk	14.11	10.44	5.04	3.23
Audubon	11.52	8.00	3.21	0.95	Kossuth	11.95	8.34	3.37	1.91
Benton	11.99	8.33	3.35	1.04	Lee	12.55	8.92	3.85	2.31
Black Hawk	10.56	7.07	2.57	0.65	Linn	12.74	9.09	3.96	2.37
Boone	13.46	9.84	4.57	1.70	Louisa	12.29	8.68	3.67	2.16
Bremer	10.45	6.99	2.48	0.62	Lucas	11.79	8.14	3.21	1.79
Buchanan	10.12	6.62	2.23	0.51	Lyon	14.54	10.85	5.37	3.50
Buena Vista	11.31	7.76	3.02	0.83	Madison	11.96	8.43	3.60	2.11
Butler	10.73	7.27	2.73	0.74	Mahaska	11.93	8.32	3.38	1.93
Calhoun	13.99	10.26	4.75	1.76	Marion	10.61	7.08	2.49	1.30
Carroll	12.97	9.27	4.09	1.39	Marshall	14.27	10.56	4.94	3.09
Cass	13.09	9.40	4.13	1.38	Mills	12.88	9.19	3.97	2.37
Cedar	12.24	8.67	3.70	1.22	Mitchell	12.00	8.36	3.36	1.90
Cerro Gordo	11.32	7.72	3.00	0.87	Monona	11.85	8.27	3.37	1.95
Cherokee	12.10	8.48	3.47	1.04	Monroe	10.70	7.09	2.51	1.29
Chickasaw	11.58	7.92	3.01	0.81	Montgomery	13.41	9.72	4.43	2.73
Clarke	12.23	8.56	3.46	1.06	Muscatine	13.26	9.55	4.33	2.71
Clay	12.09	8.57	3.64	1.20	Obrien	12.11	8.49	3.48	1.99
Clayton	9.38	5.91	1.76	0.30	Osceola	12.84	9.17	3.93	2.34
Clinton	12.52	8.85	3.69	1.19	Page	13.71	10.04	4.71	2.98
Crawford	13.53	9.91	4.64	1.70	Palo Alto	12.39	8.77	3.70	2.18
Dallas	12.57	8.95	3.85	1.22	Plymouth	12.61	8.93	3.79	2.21
Davis	13.82	10.10	4.69	1.75	Pocahontas	11.58	8.00	3.20	1.80
Decatur	13.42	9.74	4.43	1.54	Polk	12.55	8.97	3.90	2.33
Delaware	11.04	7.45	2.73	0.67	Pottawattami	12.56	8.87	3.81	2.22
Des Moines	12.02	8.41	3.44	1.05	Poweshiek	12.66	9.00	3.89	2.31
Dickinson	14.10	10.45	5.06	1.99	Ringgold	14.68	10.90	5.25	3.36
Dubuque	11.19	7.60	2.89	0.71	Sac	11.55	7.97	3.13	1.72
Emmet	12.36	8.72	3.62	1.12	Scott	12.87	9.15	3.86	2.28
Fayette	9.80	6.35	2.09	0.46	Shelby	12.36	8.71	3.63	2.09
Floyd	12.19	8.57	3.57	1.07	Sioux	12.16	8.59	3.63	2.11
Franklin	11.52	7.91	3.04	0.81	Story	14.54	10.82	5.24	3.36
Fremont	14.11	10.44	5.03	1.97	Tama	12.83	9.13	3.95	2.36
Greene	13.50	9.81	4.52	1.59	Taylor	15.30	11.53	5.78	3.80
Grundy	12.68	9.02	3.89	1.26	Union	13.35	9.67	4.38	2.70
Guthrie	12.25	8.71	3.74	1.20	Van Buren	12.97	9.34	4.18	2.52
Hamilton	14.63	10.76	5.04	1.84	Wapello	12.01	8.43	3.47	1.99
Hancock	11.72	8.16	3.25	0.89	Warren	11.66	8.08	3.17	1.74
Hardin	13.05	9.37	4.07	1.34	Washington	12.69	9.11	4.00	2.40
Harrison	12.28	8.59	3.52	1.06	Wayne	12.18	8.52	3.49	2.02
Henry	14.81	11.03	5.36	2.09	Webster	11.93	8.29	3.32	1.90
Howard	11.46	7.91	3.13	0.87	Winnebago	12.19	8.55	3.58	2.05
Humboldt	11.69	8.05	3.14	0.86	Winneshiek	10.78	7.18	2.57	1.32
Ida	12.51	8.90	3.78	1.20	Woodbury	12.13	8.46	3.42	1.93
Iowa	12.60	8.98	3.88	1.30	Worth	11.54	7.96	3.19	1.84
Jackson	11.71	8.04	3.12	0.88	Wright	12.24	8.61	3.50	2.00
Jasper	12.78	9.12	3.95	1.30					

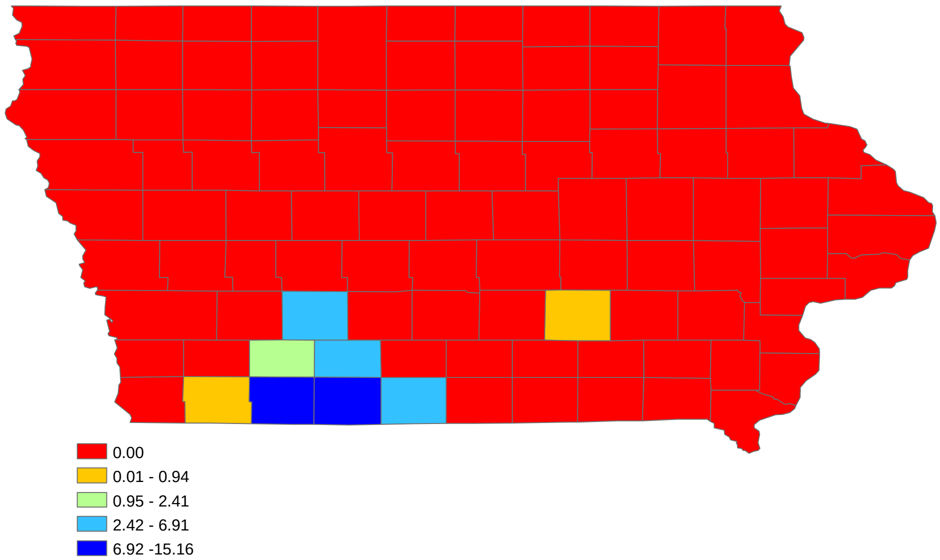

Figure 10. 2002 Contract Year Indemnity Payments ($/acre)

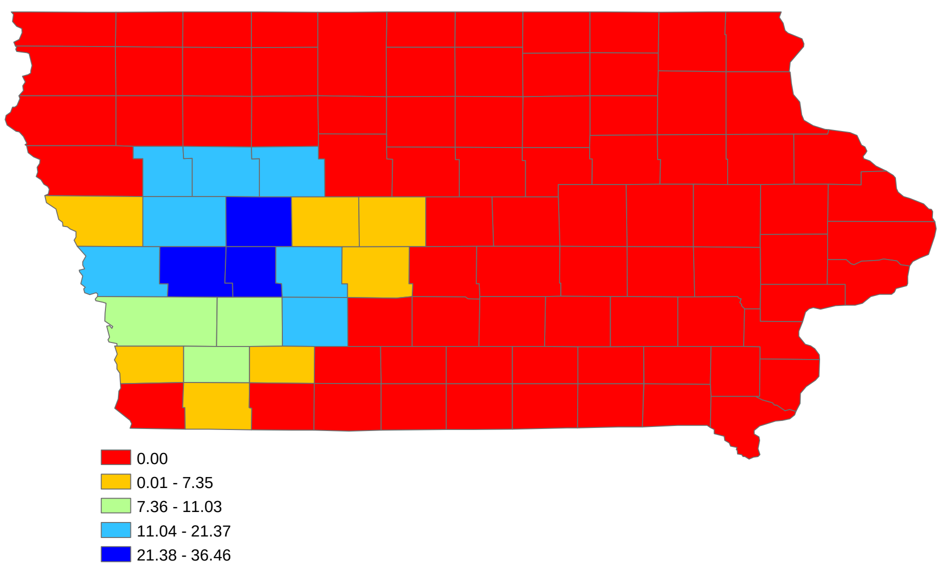

Figure 11. 2000 Contract Year Indemnity Payments ($/acre)

Table 5. Historical Premiums and Indemnities ($/acre), Averages for all Iowa Counties

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Premiums	1.23	1.29	1.41	1.52	1.60	1.63	1.67	1.71	1.73	1.76
Indemnities	0.00	0.00	2.50	0.00	0.10	4.86	0.00	0.97	1.84	0.00
Loss Ratio	0.00	0.00	0.96	0.00	0.03	1.61	0.00	0.31	0.57	0.00

Table 6. Historical Premiums and Indemnities ($/acre), Averages for Loss Counties

	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Premiums	-	-	1.52	-	2.20	1.84	-	2.12	1.57	-
Indemnities	-	-	8.36	-	5.15	12.98	-	6.50	7.55	-
Loss Ratio	-	-	5.51	-	2.38	6.92	-	3.02	4.82	-

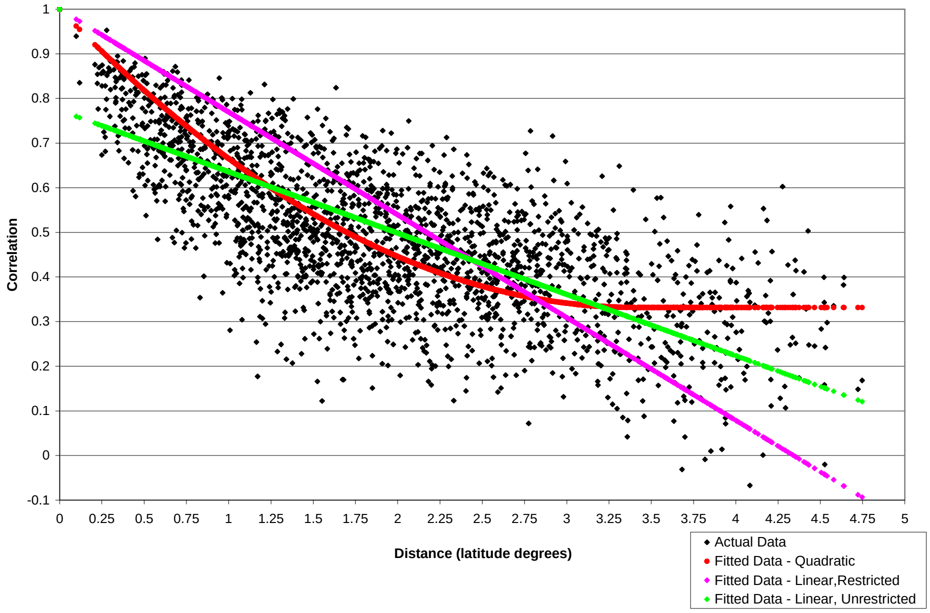

Figure A1. Correlations of Reported Precipitation vs. Distance Between Stations

Table A1. Regression Model Coefficient Estimates

Specification	Coefficient Estimate		
	Constant	$D_{i,j}$	$D_{i,j}^2$
Quadratic	1	-0.392	0.0575
Linear, Unrestricted	0.773	-0.137	-
Linear, Restricted	1	-0.230	-

Table A2. Historical Correlation Matrix for the Adair County Reference Point

	Corning	Greenfield	Lorimor	Perry 1 SE
Corning	1	0.8113	0.7701	0.5800
Greenfield	0.8113	1	0.8466	0.6881
Lorimor	0.7701	0.8466	1	0.7142
Perry 1 SE	0.5800	0.6881	0.7142	1

Table A3. Transformed Target Correlation Matrix for the Adair County Reference Point

	Corning	Greenfield	Lorimor	Perry 1 SE
Corning	1	0.8552	0.8038	0.6720
Greenfield	0.8552	1	0.8677	0.7894
Lorimor	0.8038	0.8677	1	0.7471
Perry 1 SE	0.6720	0.7894	0.7471	1

Table A4. Correlation Matrix of Correlated Draws for the Adair County Reference Point

	Corning	Greenfield	Lorimor	Perry 1 SE
Corning	1	0.8452	0.7871	0.6521
Greenfield	0.8452	1	0.8589	0.7752
Lorimor	0.7871	0.8589	1	0.7295
Perry 1 SE	0.6521	0.7752	0.7295	1

Endnotes

1. www.cme.com

2. Cross-validation refers to removing an observation from the data and estimating the variable of interest for the removed site using the remaining data. The estimated value is then compared to the true value to evaluate accuracy of the interpolation method.

3. Kirvan and Foote note that while a degree of latitude is not exactly equal to the same distance at all points due to the earth's curvature, the maximum variation in distance between degrees of latitude is only 1.13 kilometers. Therefore, the use of the average length of a degree of latitude, roughly 111.3 kilometers, is acceptable.

4. Note that we disregard issues of curvature of the earth and elevation in our analysis. While these issues may play a crucial role in examining larger distances over hillier terrain, we feel the approximation of distance in two-dimensions is adequate for the state of Iowa and small distances, relative to the size of the earth, employed in the analysis.

5. While this approach saved considerable time, the correlation structure of the rainfall distributions across space for any give year were lost. While this information was not critical to this specific application, the spatial structure of the rainfall distributions would be of definite interest for reinsurance purposes.

6. The Monte Carlo error is a measure of the deviation of the sampled mean from the mean of the true posterior distribution. See Gilks, Richardson, and Spiegelhalter and Brooks for further discussion on convergence criterion in MCMC applications.

7. It is assumed here that pasture and hay production are directly proportional. Hay prices and yields were used due to a lack of data for production value on pastureland.

The feed value of an acre of pastureland will most likely be lower than the production value of an acre of hay. Alternatively, the liability value could be multiplied by a factor which accounts for the differential in value of an acre of pastureland versus an acre of land used for hay production.

8. The factor of 0.53 is meant to adjust the liability computed from hay yield and price data for a better proxy of pasture value. It is equal to the average ratio of pasture rental rates to rental rates for land used for hay production across Iowa counties, as reported in the 2005 Iowa Rental Rate Survey published by Iowa State University Extension.

9. While computers generate random numbers in an independent fashion theoretically, there is always some level of sample correlation among the draws that creates bias in the Iman and Conover procedure.

10. The constant term for the quadratic specification was restricted to equal unity and the formula was further restricted to ensure that the marginal effect of distance was bounded above by zero. The full specification was $corr_{i,j} = 1 - 0.3921*\min(3.41, dist_{i,j}) + 0.0575*[\min(3.41, dist_{i,j})]^2$. This effectively sets the correlation between any two weather stations further than 235 miles (3.41 degrees latitude) apart equal to 0.33.