IDENTIFYING THE SUB-COUNTY ZONE FOR GROWERS: A MODEL-BASED CLASSIFICATION FOR CROSS SECTIONAL AND TIME SERIES DATA

H. HOLLY WANG

HAO ZHANG

June 14, 1999

The authors are assistant professors in the Department of Agricultural Economics, and the Program in Statistics, respectively, Washington State University.

ABSTRACT

County-based group crop insurance is not effective to growers in counties with heterogenous production practice. A model-based cluster and classification method is developed to facilitate the information needed in the study of sub-county-zone-based crop insurance. Wheat yield data from Whitman, WA is applied in the empirical analysis.

Introduction

The Risk Management Agency (RMA), former Federal Crop Insurance Corporation (FCIC) has provided Multiple Peril Crop Insurance (MPCI) for most major crops to US farmers since 1980s. The indemnity of this insurance is based on each insured's farm yield, i.e., if one's farm yield falls below his preselected coverage level, the difference will be paid. However, moral hazard and adverse selection that often occur with MPCI are believed to cause insurers great financial losses. MPCI also requires high administration costs, because the yield loss of each farm has to be individually evaluated and the management practice individually monitored. All these problems have prevented FCIC from providing MPCI at a low cost. Indeed, the government has paid 4.2 billion dollars to support this program between 1981 and 1990 (US GAO, 1995).

An alternative crop insurance program, area yield crop insurance, was studied by Miranda (1991) with the indemnity based on the average yield of an area. Currently, the area is chosen to be the county. One reason for this choice is that county level average yields for major crops are kept for decades by USDA NASS. In any particular year, an insured will receive an indemnity payment only when the county average yield of that year is lower than his preselected coverage level. Under this insurance, not only moral hazard and adverse selection have no basis, but the administration cost is also reduced greatly. RMA is currently providing Group Risk Plan (GRP) for a certain number of crops to farmers in certain areas, in an attempt to reduce financial losses.

The risk management effectiveness of GRP to a particular farmer depends heavily on the correlation between the individual farm's yield and county average yield (Wang et al., 1998). Only when all farmers have similar production practices to the insured crops and the natural conditions (such as precipitation and soil type) are homogeneous for the whole county, the county level GRP can be a meaningful and effective insurance to farmers.

However, geographical and natural conditions may vary from one area to another in a county, which results in different farming practices and farm yields. For example, Whitman in eastern Washington, an important wheat production county with an area several times larger than a typical county in the Midwest, has three distinct precipitation zones: 1) a low precipitation zone receives 9 - 14 inches annually; 2) an intermediate precipitation zone receives 15-18 inches; and 3) a high precipitation zone receives 19 - 24 inches. The cropping systems in the three zones are also different. Crops are grown once every two years in the low precipitation zone with winter wheat - summer fallow as the primary rotation, twice every three years typically in the intermediate precipitation zone with winter wheat - spring barley - summer fallow as the primary rotation, and annually with wheat rotated with peas or other crops in the high precipitation zone (USDA, 1978). All these result in different wheat yield levels and risks across the county. The country level GRP is thus not an effective risk management instrument to farmers. This is one of the most important reasons that no Whitman farmer participated in GRP in 1997.

In a county like this, reducing the size of the area on which GRP average yield is based helps improve its effectiveness (Wang, et al, 1998). Therefore, a sub-county zone based GRP (ZGRP) will be more effective to farmers than the county based GRP without losing the advantages over MPCI. There are two essential problems in the design, evaluation or any research on ZGRP: to identify different clusters each of which consists homogeneous farms, and to classify each farm into an appropriate cluster. These two problems are complicated by the fact that farm identifications are removed from databases available to the public. For example, RMA keeps yield records of individual farms in each county for 10 years. But farm identifications such as names, addresses, etc., are not available for the sake of confidentiality of farmers.

We consider a model based approach to these two problems when each farm's yields are

available for some consecutive years without farm identifications. Farm yields are correlated across both time and sections (farms). The following mixed effects model is employed that catches the correlations:

$$X_{ijt}=\mu+\mu_i+f_{ij}+y_t+\epsilon_{ijt},\ i=1,2,...,I,\ j=1,2,...,n_i,\ t=1,2,...,T, \qquad (1)$$

where X_{ijt} denotes the yield of farm j in cluster i at year t, μ_i is the fixed cluster effect with $\sum \mu_i=0$, f_{ij} is the random farm effect with a mean 0 and a standard deviation $\sigma_{i,f}$, and counts for the effect caused by the unmeasured natural conditions and farming practices, y_t is the random year effect with mean 0 and standard deviation σ_y caused by precipitations, temperature and other yearly factors, and ϵ_{ijt} is the random error with a mean 0 and a standard deviation σ_ϵ. I is the number of clusters, n_i is the number of farms in cluster i, and T is the total number of years. We assume all random variables are independently and normally distributed.

Our goal here is to define a cluster, identify all appropriate clusters for the given data set, and then classify all farm yields. With the linear mixed effects model, a cluster can be defined as a group of farms with the same mean $\mu+\mu_i$, denoted by m_i. We will first describe the statistical methods for clustering and classification in next section, and apply these methods to Whitman County wheat yield data in the following section. Conclusions are drawn in the final section.

Statistical Analysis

Since we do not know which cluster a farm belongs to, the likelihood function for a sample from model (1) can not be evaluated. For Whitman County data, there are 2,945 farm observations and each farm has as many as 10 years of yield. For these reasons, we do not use the likelihood function of the observed x_{ijt} values. Instead, we will use the average yield of each farm for clustering and classification. Let $\bar{X}_{ij.}, \bar{\epsilon}_{ij.}, \bar{y}_{.}$ be the averages over the omitted indices. Then

$$\bar{X}_{ij.}=m_i+f_{ij}+\bar{y}_{.}+\bar{\epsilon}_{ij.}$$

are normal random variables with the following covariance structure

$$Cov(\bar{X}_{ij\cdot},\ \bar{X}_{kl\cdot}) = \begin{cases} \sigma_{i,f}^2 + \frac{1}{T}\sigma_y^2 + \frac{1}{T}\sigma_\epsilon^2 & \text{if } i = k,\ j = l \\ \frac{1}{T}\sigma_y^2 & \text{otherwise.} \end{cases} \qquad (2)$$

We see that the correlation coefficient between $\bar{X}_{ij\cdot}$ and $\bar{X}_{kl\cdot}$ approaches 0 when T approaches infinity. Therefore, they are approximately independent for a large T.

Clustering

The T-year average yields $\{\bar{X}_{ij\cdot}\}$ consist of I sets of identically (not independently) distributed random variables: $\{\bar{X}_{1j\cdot},\ j=1,\cdots,n_1\}$, $\{\bar{X}_{2j\cdot},\ j=1,\cdots,n_2\}$, $\cdots$, $\{\bar{X}_{Ij\cdot},\ j=1,\cdots,n_I\}$, that are normally distributed with a mean m_i and a standard deviation σ_i, $i=1, 2, ..., I$, respectively, where $\sigma_i^2=\sigma_{i,f}^2+(\sigma_y^2+\sigma_\epsilon^2)/T$. If the variables are all independent, as they approximately are when T is large, then $\{\bar{X}_{ij\cdot}\}$ can be viewed as an i.i.d. sample from the following mixture distribution:

$$f(x;\theta)=\sum_{i=1}^{I} p_i \phi(x|m_i,\sigma_i) \qquad (3)$$

where the parameter $\theta=(p_1,p_2,\cdots,p_{I-1},m_i,\sigma_i,i=1,2,\cdots,I)$, $p_i=n_i/n$, $p_I=1-\sum_{i=1}^{I-1} p_i$ and $\phi(\cdot|m,\sigma)$ denotes the probability density function of the normal distribution with mean m and standard deviation σ.

The mixture distribution of form (3) has been extensively studied and applied, which also provides an important model based approach to clustering (Everitt and Hand,1981; Titterington, Smith and Makov,1985; McLachlan and Basford, 1988). Parameters in the mixture distribution (3) can be estimated via the maximum likelihood method and the likelihood function for an i.i.d sample $x_1,,x_2,\cdots,x_n$ is

$$\prod_i f(x_i;\theta). \qquad (4)$$

However, in our problem, the average yields are not exactly independent, and consequently, (4) with x_i being replaced by $\bar{X}_{ij\cdot}$ is not the exact but an approximation of the likelihood function of our sample from model (1). Nevertheless, we still maximize (4) for parameter estimation. Note that it does not give estimates for $\sigma_{i,f}$, σ_y and σ_ϵ in model (1), but these parameters are not needed in our approach to clustering and classification. We have done some simulations to evaluate this approximated maximum likelihood estimators.

In the simulations, we first generate sample data from model (1) for a given set of parameters, then estimate the parameters by maximizing the approximated likelihood function. We compare these estimates with the maximum likelihood estimates obtained from an i.i.d. sample from the corresponding mixture distribution (3), and with the true values. If these sets of estimates are reasonably close, the approximation will be considered acceptable for parameter estimation.

We first considered I=2, i.e., two components in the mixture. We have generated i.i.d. samples from the mixture distribution (3) with various parameter values of $p_1, m_1, m_2, \sigma_1, \sigma_2$ and found the MLE performs well when the two normal components are well separated. The MLE generally performs well when the two normal components are separated by 1.5 standard deviations from each of the two means, i.e., the cross point (the value at which the two densities equal) is at least 1.5 standard deviations from each of the two means. We used those parameter values for which the MLE preforms well to generate samples from the mixed model (1) for comparison purpose. Findings for the $I=2$ case helps us choose the parameter values to be used for $I=3$ since the three components should not be less separated for the MLE to perform comparably to the I=2 case.

The y_t term in the mixed model is simulated from the normal distribution with mean 0 and standard deviation σ_y and the beta distribution with shape parameters 6 and 2. The transformed distribution is skewed to the left with mean 0 and standard deviation σ_y. We used beta distribution

to see how sensitive the estimates are to the distribution asymmetry because yield is usually asymmetric. Note that even though y_t is non-normal, the average yield $\bar{X}_{ij\cdot}$ is approximately normal by the Central Limit Theorem. We keep the other two random terms in the mixed model normal since it is less justifiable to assume their non-normality than the y_t's. Rarely occurred bad weather can greatly affect crop yields while more frequently occurred good weather can only improve yield by a small margin. Therefore y_t should be skewed to the left.

From the simulation results, the MLE from the approximated likelihood function (hereafter called AMLE) gives competitive estimates for p_i, in fact the AMLE for p_i has smaller variances and seems unbiased. The AMLE for the mean m_i has a larger variance than the MLE, and stronger correlations (i.e., larger σ_y and σ_ϵ) increase the variance. For a component with a smaller p_i, the variance of AMLE for m_i increases by a smaller amount than the component with a larger p_i.

The AMLE persistently underestimates the standard deviations σ_i. This has to do with the positive correlations among X_{ijt} . We can see this when there is only one cluster, i.e., I =1. We see from (2) that $\boldsymbol{X} = (\bar{X}_{1j\cdot},\ j=1,2,\ldots,n_1,\ t=1,2,\ldots,T)$ has a covariance matrix with diagonal elements all equal to σ^2 and off diagonal elements all equal to $\rho\sigma^2$, where ρ is the correlation coefficient between any two distinct elements of $\boldsymbol{X}$. We simply write X_j for $\bar{X}_{1j\cdot}$ If X_js are viewed as independently normally distributed, the MLE of σ^2 is $\sum (X_j-\bar{X})^2/n_1$, which will be the AMLE of σ^2. Note that

$$\begin{aligned}\frac{1}{n_1}E\sum_{j=1}^{n_1}(X_j-\bar{X})^2 &= \frac{1}{n_1}E\sum_{j=1}^{n_1}((X_j-\mu)^2-\frac{1}{n_1}(\sum_{j=1}^{n_1}(X_j-\mu))^2)\\ &= \frac{1}{n_1}E\sum_{j=1}^{n_1}\left(\frac{n_1-1}{n_1}(X_j-\mu)^2-2\sum_{i<j}(X_i-\mu)(X_j-\mu)\right)\\ &= (1-\frac{1}{n_1})(1-\rho)\sigma^2.\end{aligned}$$

Therefore, it is biased and underestimates σ^2. Since we do not have an analytic expression for MLE variances σ_i^2 in the mixture distribution, it is hard to give the biases of the AMLE for the variances.

Non-normality of the y_t term has no obvious effects on the estimates. We see that increasing the sample size *n* decreases the variances of the estimators, and increasing *T* (therefore decreasing correlations among the average farm yields) makes σ_i^2 less underestimated.

We conclude that the AMLE is acceptable when the correlations among the individual farm yields are not too high. For crop yield data, there are often missing values due to crop rotation and fallow. The average yields, nevertheless, can still be regarded as a sample from the mixture distribution. Therefore, the parameters can be estimated by AMLE.

The number of clusters, *I*, can be estimated from the histogram of the pooled sample $\{\bar{X}_{ij\cdot}\}$. Actually, *I* is first picked from the histogram and then other parameters are estimated. Akaike's information criterion has been applied to determine the proper number of clusters *I* by Scolve (1983), and Bozdogan and Scolve (1984). For the mixture model (3), AIC is to choose the *I* that minimizes

$$AIC(I) = -2L(\hat{\theta}) + 2N(I)$$

where *L* is the log likelihood, $\hat{\theta}$ is the MLE of the parameter θ, and *N(I)* is the number of free parameters in the *I*-component mixture model. AIC is used to determine the number of clusters.

Classification

We now consider how to classify a farm into one of the clusters. We will employ the following three classification methods. The first one is the Bayesian classification that maximizes the posterior distribution. When an observed value *x* (average yield) is from the mixture distribution, the probability that it belongs to cluster *i*, according to the Bayesian formula, is

$$P_i = \frac{p_i f_i(x)}{\sum_{s=1}^{I} p_s f_s(x)},$$

where $f_i(x)$ is the probability density function of cluster *i*, which is $\phi(x|m_i,\sigma_i)$ in our model. Therefore, it is classified to cluster *i* if $p_i f_i(x) = \max_j\{p_j f_j(x)\}$.

The second method is similar to the minimum distance classification. But we use the

standardized distance here. More specifically, the farm is classified to cluster i if $\frac{|x-m_i|}{\sigma_i} = \min_j \left\{ \frac{|x-m_j|}{\sigma_j} \right\}$, where m_i and σ_i are the mean and standard deviation of cluster i.

The third method is to maximize the probability density function, i.e., to classify x to cluster i if $f_i(x) = \max_j f_j(x)$.

As with any classification method, misclassification can occur. Misclassification rates for each of the three methods can be calculated for a given set of parameters as illustrated in next section. Therefore, the methods can be compared and evaluated for a given set of parameters.

The Application to Whitman County Wheat Growers

The Data and Exploratory Analysis

Whitman is a county in eastern Washington where dryland wheat production is a prominent industry, and produces one of the highest yields in the world. It accounts for 20% of wheat production in Washington. FCIC has recorded dryland winter wheat farm yields for each MPCI participant for the maximum of 10 production years from 1981 to1995. We obtained the yields for 2945 farms and plotted the annual average yields of the farms in Figure 1. No obvious trend is present. Figure 2 shows the histogram of the temporal average yield $\bar{X}_{ij\cdot}$ of each of the farms. The distribution seems to have three modes, with the lowest one clearly differentiated from the other two, but the two higher ones are close to each other.

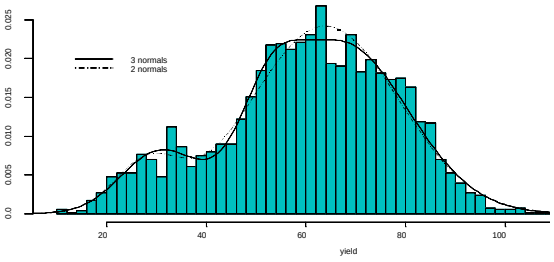

Figure 1 Histogram of the average yields and fitted pdf curves

As discussed earlier, because crop yield distributions are usually skewed to the left, beta distributions are used by some agricultural economists when studying yield risk and crop insurance (Nelson, 1991; Hennessy, Babcock and Heyes,1997). However, the average yields over the years will be approximated normally distributed based on the Central Limit Theorem.

Parameter Estimation and Clustering

We, first, need to determine the number of normal components in the mixed distribution. The likelihood ratio test is inappropriate for this test, because the parameter value in the null hypothesis lies on the boundary of the unconstrained parameter space, as noted in Titterington, *et al.* (1985). We take the number of components I to be 2 and 3 and estimate the corresponding parameters. We then use Akaike's Information Criterion (AIC) and some diagnostic analysis, such as quantile - quantile plot (QQP) to determine the proper number of components.

The approximated MLEs are presented in Table 1. For the three-component mixture model, 15% of the farms are from the low precipitation zone, 11% from the intermediate zone, and the rest 74% from the high zone. The mean yields of the three zones are 30.36 bu/ac, 52.56 bu/ac and 67.01 bu/ac, respectively, and the standard deviations are 7.64 bu/ac, 5.63 bu/ac and 13.53 bu/ac. The last two zones have closer mean yields, while the low zone mean yield is away from the other two. For the two-component mixture model, 10% are from the component with mean 28.09 and standard deviation 7.33, and 90% from another component with mean 63.85 and standard deviation 14.77. Even though the 2-normal mixture distribution also fits the data satisfactorily well, it makes it hard

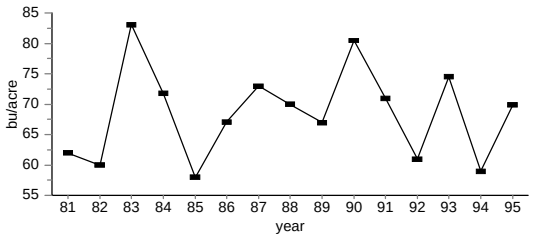

Figure 1 Annual average winter wheat yields of Whitman County from 1981 to 1995

to interpret the components. It seems that the 2-normal mixture combines two components in the three-normal mixture into The AIC suggests choosing the 3-normal mixture, so does QQP.

Classification

We use the 3-component mixture distribution for classification. The Bayesian classification is unsatisfactory since no farm is classified in cluster 2. The reason is that, according to the clustering result in the previous subsection, the second and the third clusters have means close to each other, but cluster 2 has only 11% of all farms while cluster 3 has 74%. The huge difference in the proportions has most farms classified into cluster 3 by the Bayesian method.

Table 1 Mixture Model Parameters for Whitman County Wheat Yields

	p	μ	σ	Log-likelihood
Mixture of Three Normal Distributions				-12581
1	0.146	30.36	7.64	
	(0.019)*	(1.025)	(0.551)	
2	0.107	52.56	5.63	
	(0.078)	(1.01)	(1.72)	
3	0.747	67.01	13.53	
	-	(2.10)	(0.91)	
Mixture of Two Normal Distributions				-12586
1	0.10	28.09	6.61	
	(0.010)	(0.688)	(0.448)	
2	0.90	63.85	14.77	
	-	(0.394)	(0.301)	

* Values in the parentheses are the estimated standard deviations of the estimators.

Presented in Table 2 are the classification results by the minimum distance method and the maximum probability density method. For the minimum distance classification, 17.4% were classified into cluster 1, 22.3% into cluster 2 and 60.3% to cluster 3. For the maximum probability density classification, 16.9% are classified into cluster 1, 31.1% to cluster 2 and 52% to cluster 3. The proportions of farms classified to the three clusters do

not agree well with the proportions in the clustering result or in the mixture model. This is due to the misclassification. As we will see below, these classification results are really what one should expect.

We now briefly discuss the misclassification rates of the three methods. For simplicity, we assume the sample is from the mixture distribution with three normal components with parameters as in Table 1, i.e., the estimates are the true parameter values. The Bayesian classification does not perform well due to the huge difference between p_2 and p_3. In fact, $p_3 f_3(x) > p_2 f_2(x)$ for all x. Therefore, no data is classified into cluster 2. We will therefore focus on misclassification rates of the other two methods.

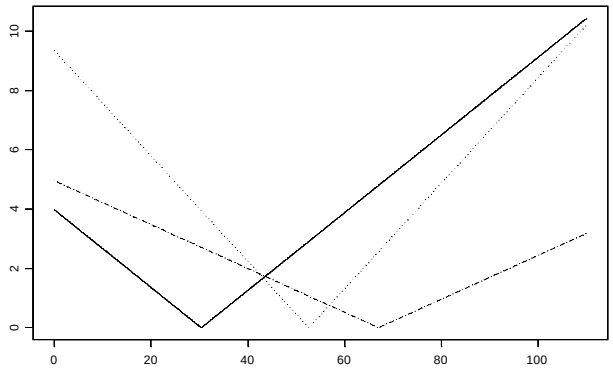

Figure 3 Plots of the distance functions. Solid line for cluster 1, dotted for cluster 2 and dashed for cluster 3.

It is easy to see that the minimum distance classification classifies x into cluster 1 if $x <$ 43.1426, to cluster 2 if x is between 43.1426 and 56.8072, and to cluster 3 if x > 56.8072 (The three distance functions are plotted in Figure 3). Let α_{ij} be the conditional probability that x is classified into cluster j given that it is from cluster i. α_{ij} can be directly calculated from the normal distributions:

Table 2 Classification of Farms

Cluster	Minimum Distance	Maximum Density
1	17.4	16.9
2	22.3	31.1
3	60.3	52.0

$$\alpha_{i1} = \Phi\left(\frac{43.1426 - m_i}{s_i}\right),$$
$$\alpha_{i2} = \Phi\left(\frac{56.8072 - m_i}{s_i}\right) - \Phi\left(\frac{43.1426 - m_i}{s_i}\right),$$
$$\alpha_{i3} = 1 - \Phi\left(\frac{56.8072 - m_i}{s_i}\right)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. For example, $\alpha_{12} = 0.046744$, $\alpha_{13} = 0.000266$, $\alpha_{21} = 0.047011$, $\alpha_{23} = 0.225302$, $\alpha_{31} = 0.038817$, $\alpha_{32} = 0.186486$. Then the probability that x is classified into cluster j is $\sum p_i \alpha_{ij}$. For j = 1, 2, 3, the probability is 0.1732, 0.2240, and 0.6028 respectively. Note that the classifications results in Table 1 are very close to these theoretical probabilities.

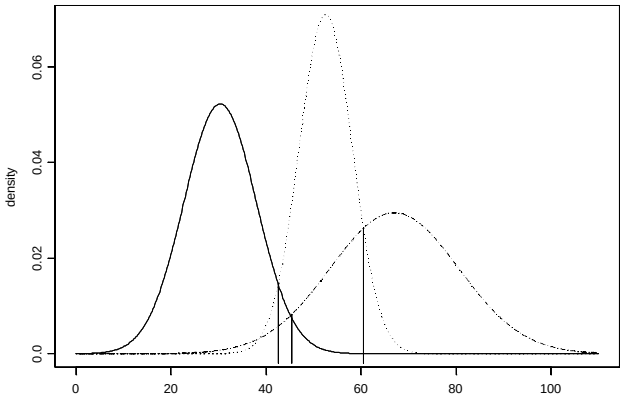

Figure 4 Pdfs for the three normal components. Solid - cluster 1, dotted - cluster 2, and dashed - cluster 3.

The maximum probability density method classifies x into cluster 1 if x<42.6, to cluster 2 if x is between 42.6 and 60.52, and to cluster 3 if x > 60.52 (See Figure 4). Analogously, the probability that x is classified into cluster 1, 2, or 3 is 0.1687, 0.3117, or 0.5196. And these results agree well with the classification results in Table 1.

The minimum distance classification outperforms the other two method for the Whitman County wheat yield data. Different proportions in the clustering results and classification results are due to misclassification and such differences are what we should expect theoretically. Using the individual yields instead of the average yields might reduce misclassification rates. For example, each farm in cluster i has mean yield m_i with variance $\sigma_{1f}^2 + \sigma_y^2 + \sigma_\epsilon^2$, therefore, X_{ijt} will be in the interval $m_i \pm 1.96(\sigma_{1f}^2 + \sigma_y^2 + \sigma_\epsilon^2)^{0.5}$ with a probability 0.95 assuming all the random terms are normal. However, because of the dependence of yields, estimation for $\sigma_{1f}^2 + \sigma_y^2 + \sigma_\epsilon^2$ will be difficult. Missing yield values further add to the difficulty.

Summary and Conclusion

In this paper, a statistical approach is developed to cluster and classify each subject into an appropriate category via the mixed effect model when data are correlated across both time and section. Averages over time are used, which have approximately a mixture distribution. Parameters can be estimated by AMLE. Simulation results show that this approximation provides reasonably good estimates of the parameters when the correlations among individual subjects are not too high.

This approach is particularly appropriate for clustering and classifying farms since yields are correlated across both time and sections. Also, due to the crop rotation and fallow, crop yield data have a lot of missing values. Using the average yields instead of the original yield data takes care of the missing values. This model is applied to wheat yields in Whitman County of Washington to identify clusters of farms and classify farms into different clusters for the purpose of design, evaluation and implementation of a sub-county zone based crop insurance.

Reference

Bozdogan, H. and Scolve, S.L. (1984) "Multi-sample cluster analysis using Akaike's information criterion". *Ann. Inst. Statist. Math.* 36, 163-180

Everitt, B.S. and Hand, D. J. (1981) *Finite Mixture Distributions.* Chapman and Hall, London.

Hennessay, David A., Bruce A. Babcock, and Dermot J. Hayes. (1997) "Budgetary and Producer Welfare Effects of Revenue Insurance." , *Amer. J. of Agri. Econ.*, 79:1024-1034

McLachlan, Geoffrey J. and Basford, Kaye E. (1987) *Maximum Models: Inferences and Applications to Clustering*. Marcel Dekker, Inc., New York.

Miranda, M. J. (1991) "Area-Yield Crop Insurance Reconsidered", *Amer. J. of Agri. Econ.*, 73:233-242

Nelson, C. H. (1991) "The Influence of Distributional Assumption on the Calculation of Crop Insurance Premia", *North Central J. of Agri. Econ.*, 12:71-78

Scolve, S.L. (1983) "Application of the conditional population-mixture model to image segmentation" *IEEE Trans. Pattern Anal. Machine Intell.* PAMI-5, 428-433

Titterington, D. M., Smith, A.F.M. and Makov, U.E. (1985) *Statistical Analysis of Finite Mixture distributions*. Wiley, London.

Wang, H. Holly, Steven D. Hanson, Robert J. Myers, and J. Roy Black. (1998) "The Effects of Yield Insurance Designs on Farmer Participation and Welfare." *Amer. J. of Agri. Econ.* 80: 806-820

WASS (1996) *Washington Agricultural Statistics*, Compiled by Washington Agricultural statistics Service, 1995-1996

U.S. Department of Agriculture (1978) "Palouse Cooperative River Basin Study." Cooperative study by Soil Conservation Service, Forest Service, and Economics, Statistics, and Cooperatives Service of Whitman County, WA, Washington D.C.: Government Printing Office

U.S. General Accounting Office. (1995) "Crop Insurance Additional Actions Could Further Improve Programs's Financial Condition" Report to the Ranking Minority Member, Committee on Agriculture, Nutrition, and Forestry, U.S. Senate Washington DC GAO/RCED-95-269 September 1995