PURE RISK PREMIUM FOR CROP INSURANCE

ABINAYA P.

A project report submitted to the

Madras School of Economics

In partial fulfillment of the requirement for the award of the

Post Graduate Diploma

in

ACTUARIAL ECONOMICS

MADRAS SCHOOL OF ECONOMICS
Gandhi Mandapam Road, Kotturpuram,
Chennai – 600 025, INDIA

May 2010

Degree and Branch	: **Post Graduate Diploma (Actuarial Economics)**
Month and Year of Submission	: **MAY 2010**
Title of the Project Work	: **Pure Risk Premium for Crop Insurance**
Name of Student	: **ABINAYA P.**
Roll Number	: **AE-2008-01**
Name and Designation of Supervisor	: **Dr. Lalitha Subramanian** Visiting Associate Professor, Madras School of Economics Chennai- 600025

BONAFIDE CERTIFICATE

This is to certify that this project report titled "**Pure Risk Premium for crop insurance"** is the bonafide work of **Ms. Abinaya P** who has carried out the research under my supervision. Certified, further that to the best of my knowledge, the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

Dr. D.K Srivastava
Director,
Madras School of Economics
Anna University,
Chennai- 600025

Dr. Lalitha Subramanian
Professor,
Madras School of Economics
Anna University,
Chennai- 600025

Abstract

Government run crop yield insurance scheme is a major instrument being used to protect the Indian farmer from crop failures. However, these crop insurance schemes suffer from adverse selection and moral hazard problems because of the usage of flat rates. Given the current levels of yield and rainfall variability the actuarially fair premium rates are likely to be high and in many cases unattractive or unaffordable. So, instead of adopting the easy and unsustainable route of large subsidies, in the long term the government should consider risk mitigation through improvements in the irrigation and water management infrastructure. However, actuarial premium rates will show a clear picture of the reserves that will be available in the long term, for addressing the problem of crop failure and the vulnerability of farmers to this problem. When these rates cross a certain level, serious steps could be taken by the government towards investment in irrigation infrastructure and other agricultural risk management strategies .We follow a parametric approach to the estimation of probability densities for yield distribution and apply the formula $Prob[x < \lambda\mu]\ [\lambda\mu - E(x \mid x < \lambda\mu)]$, where x is the realized yield, λ is the coverage level and μ is the guaranteed yield to find the premium in terms of yield (kg/hectare) for a fixed guaranteed yield.

Acknowledgements

I am grateful to Dr. Lalitha Subramanian, Visiting Associate Professor for her guidance and resourceful insights from time to time. I gratefully appreciate Dr. K.R. Shanmugam, my panel member for his support and advice. I thank Dr. Shyamala, Head of the Department of Ethiraj College for her technical assistance. I would like to thank my seniors (Andal Vedanarayanan, Aishwarya and Kamleshwari) for their help on data procurement and for giving me a helping hand time and again. Lastly, I would like to thank all my faculty, parents and friends for making this dissertation possible.

Table of contents

List of Tables

List of figures

Pure risk premium for crop insurance

Chapter 1

1.1 Introduction

The main objective of this study is to design a single-crop yield insurance plan with a particular emphasis on production risk modeling and arrive at the pure risk premium for the same. We show how the data available shapes the insurance scheme and ratemaking procedure. Though there are well established public crop insurance programs, their overall financial experience has been expensive for the government and far from popular with farmers. This is mainly due to the lack of sound actuarial rate making procedure followed in fixing the premiums. Premium rating procedures that more accurately reflect producer risk mitigate adverse selection and moral hazard problem. If premium rates are actuarially set, thought they may be higher than the subsidized premium rates of the government crop insurance plan, we will have a clear idea of the reserves that will be available in the long term, for addressing the problem of crop failure and the vulnerability of farmers to this problem.

Specifically yield distributions should be estimated on the basis of farm level dataset, but due to the lack of such detailed information, we resort to district level data, using a yield distribution fitting approach, where the yield distributions are used to derive the probabilities of the yield falling below a certain level.

The crop insurance plan is an individual yield guarantee, where the yield is measured in kg/hectare. Cover is only for the coming season, after which it has to be renegotiated. Claims are not of fixed amount. Settlement of claim may take some time till the actual loss is assessed. The policy can give rise to utmost one claim only. Under an individual yield guarantee, a farmer receives an indemnity whenever his actual yield falls below his yield guarantee. The yield guarantee is different for every insured farmer, as it varies with the recent yield realized in his farm. It is based on the average of the past five years of yield data available for that particular farm. So, it can be seen that the yield guarantee is neither chosen by the farmer nor the insurer. The farmer is however given the option of choosing a certain percentage of his yield guarantee called coverage level between 60%-90%, as a way of reducing his burden of premium payment. The coverage level brings down the premium but also has an impact on the level of indemnity payments made at the time of loss. If the yield that is realized by a farmer in the next insured season is less than this figure, he receives an indemnity equal to the multiplicative product of the market price of the rice per kg and the difference between his yield guarantee and his actual yield, per hectare. This reduces the problem of adverse selection that arises when flat rates are set for an entire state. Flat rates encourage high risk producers to take out insurance and become part of the insurance pool. This inadvertently results in wrong predictions of losses because of a high proportion of high risk producers being part of the pool, thus collapsing the system. An index based system may suffer from the same problem, when there is low correlation between the index and the farmers' yield.

We propose to address the problem of moral hazard by introducing coinsurance, which is an insurance policy provision under which the insurer and the insured share costs incurred according to a specific formula. More generally, it is a sharing of risk between the insurer and the insured also called copay. Since coinsurance affects the indemnity payments and we concern ourselves only with premium calculation, we do not make any calculations with regard to coinsurance. Moreover, moral hazard problem is inherent in a situation where the insurer has to rely on the insured itself for the data, it is highly recommended to obtain such information from a third party who is involved in assessing the crop yield of the insured lands, even though it might take some time to compile such detailed information.

1.2 Parametric approach

This approach is essentially used due to data availability constraints in terms of all the independent variables affecting the yield. Using time series data for yield modeling requires some care due to temporal correlation that affects estimation and inference (Wooldridge, 2003). For this reason, the tests for heteroskedasticity and unit-root are addressed prior to estimating the yield distributions (see section 3). Therefore the series is required to be detrended. Though other types of statistical procedures for modeling crop yields like non-parametric and semi-parametric are available, parametric approach performs relatively well in small sample applications and when adequate information on the possible explanatory variables are not available. Another advantage of this approach is that once a distribution has been fitted, a variety of statistical tests and hypothesis testing could be done.

Parametric approach to the estimation of probability densities generally involves using an observed series of yield realizations to estimate specific parameters that describe the probability density and distribution function. Given a set of iid (independent and identically distributed) yield realizations, one can estimate the parameters of the distribution using maximum likelihood method, method of moments or method of percentiles estimation procedures. Since crop yields are affected by systemic components like erratic weather, pests and natural phenomena, only when the effect of such components are removed from the data, will the observations be independent. This way detrending helps estimate the distributions by addressing the issue of systemic elements present in the yield data.

The scheme formulated here provides 'comprehensive risk insurance' against loss in farm yield i.e. fall in actual yield over guaranteed yield in a notified area arising out of adverse fluctuations in yield due to occurrence of any one or combination of non-preventable natural perils such as flood, inundation, storm, cyclone, hailstorm, landslide, drought, dry spells and large scale outbreak of pests/diseases.

Estimating the premiums actuarially will show the quantum of risk involved in insuring the production risk in agriculture for every district or unit of area covered. The extent to which government subsidy is causing the adverse selection problem, by keeping the premiums to a low level, could be reckoned from the results. Moreover, once the government knows the amount of money that is being spent on crop insurance in the form of subsidies, it will take steps to invest in some risk management strategies that will help the farmers in the long run like setting up good irrigation infrastructure, restoring the soil fertility through natural means, educating the farmers on good agricultural practices and addressing other basic issues that are ravaging their livelihood.

1.3 Literature review

S.S. Raju and Ramesh Chand in a study on the "Performance of National Agricultural Insurance Scheme and Suggestions to make it More Effective" have stated that crop insurance is the only mechanism available to safeguard against production risks. They have examined the features and performance of National Agricultural Insurance Scheme (NAIS) operating in the country and have suggested some modifications to make it more effective.

"If crop insurance programme is to be made an important tool in agricultural risk management, the present level of coverage will have to be improved, at least by 3-4 fold. Such an expansion can occur only with improvements in and broad-basing of the insurance scheme. Every suggested improvement has financial implications and affects the concerned insurance practices. It requires renewed efforts by the government in terms of designing appropriate mechanisms and providing financial support to agricultural insurance. Providing of similar support to the private sector insurers would help in increasing the insurance coverage and improving the viability of insurance schemes over time."

1.3.1 Main Features of NAIS

The National Agricultural Insurance Scheme (NAIS) was introduced in the country from the rabi season of 1999-2000. Agricultural Insurance Company of India Ltd (AIC), which was incorporated in December, 2002, and which started operating from April 2003, took over the implementation of NAIS. This scheme is available to both loanees and non-loanees. It covers all food grains, oilseeds and annual horticultural / commercial crops for which past yield data are available for an adequate number of years.

The premium rates applicable on the sum insured are:

- Bajra and oilseeds : 3.5 %
- Other kharif crops : 2.5 %
- Wheat : 1.5%
- Other rabi crops : 2.0%
- Annual commercial / horticultural crops : Actuarial rate

Initially, the premium in the case of small and marginal farmers was subsidized @ 50 per cent, which was shared equally by the Government of India and the concerned State/UT. NAIS premium rates are now proposed to be actuarially fixed in the next five years time.

The number of loanee farmers covered under NAIS averaged around 19 lakhs in the rabi-winter season (October to March) during 2000-01 and 2002-03. This number showed a significant increase during the next three rabi seasons (2003-04 to 2005-06) and reached the figure of 32.75 lakh. The number of non-borrower farmers showed a wide year-to-year fluctuations. There was a big jump in the non- loanee farmers opting for insurance in the year after 2002-03, which was a very severe drought year. The compensation received by those who had insured, induced a large number of other farmers to take the benefit of insurance in the adverse event. This shows a strong tendency towards adverse selection problem. Further, the non-borrower farmers' participation had come from those areas and crops which were most likely to report high crop losses. Their participation was predictably the highest, during adverse seasons. On an average, 1.63 hectare area was insured per farmer under NAIS during rabi 1999 through rabi 2005-06. The average sum insured per household ranged from less than Rs 5000 in Goa, Himachal Pradesh and Jharkand to more than Rs 15000 in Gujarat, Tamil Nadu and Pondicherry. The average amount insured per farmer under NAIS at the aggregate level was Rs.9573. The average premium paid by the individual farmer on per hectare basis varied between Rs.44 (Goa) and Rs.513 (Meghalaya) .

1.3.2 Suggestions that were made to make the National Agricultural Insurance Scheme more effective

For the scheme to become more popular, the unit for determining claim should be reduced to the level of 'village' in the case of large villages and to 'cluster of villages' in the case of small villages. Ideally, "Individual approach" would reflect crop losses on a realistic basis, and has been regarded most desirable (Dandekar, 1985). However, under the Indian conditions, implementing a crop insurance scheme at the "individual farm unit level" is beset with problems, such as:

- Non-availability of the past records of land surveys, ownerships, tenancy and yields at individual farm level
- Small size of farm holdings
- Remoteness of hamlets and inaccessibility of some farm-holdings
- A large variety of crops, varied agro-climatic conditions and package of practices, and
- Inadequate infrastructure.

Data being the lifeline of insurance, the actuarial rating of the product at GP Gram Panchayat) level would be possible only if the historical yield data at GP level is available for a reasonably long period. Such data at the GP level are not available and therefore, it would be difficult for the insurer to work out premium rates on sound actuarial principles (Planning Commission, 2007).

At present, the levels of indemnity are 60 per cent, 80 per cent and 90 per cent corresponding to high, medium and low risk areas. Suggestion was made that instead of three levels of indemnity, there should be only two levels of indemnity, viz. 80 per cent and 90 per cent. But, these higher levels of indemnity may escalate the premium rates, and would increase the subsidy burden of the government.

The processing of claims in NAIS begins only after the harvesting of the crop. Further, claim payments have to wait for the results of CCE (crop cutting experiments) and also for the release of funds from the central and state governments. Consequently, there are a few months' gap between the occurrence of loss and indemnity payment. To expedite the settlement of claims in the case of adverse seasonal conditions, and to ensure that at least a part-payment is made to the farmer, it is suggested to introduce 'on-account' settlement of claims, without waiting for the receipt of yield data, to the extent of 50 per cent of likely claims, subject to adjustment against the claims assessed on the yield basis.

Crop insurance helps in the flow of credit to crop production in the event of crop failures, since loans can be repaid with the help of the indemnity payments received.

Sidharth Sinha in his paper on 'Agriculture Insurance in India' has focused on the use of agriculture insurance schemes to protect farmers from agricultural production variability. According to Mr. Sidharth Sinha, "The experience of government supported and subsidized crop insurance and the recent entry of private insurers raise questions about the co-existence of government and private agriculture insurance. One view is that the private sector will be unable to compete with government insurance, given the subsidies and access to the administrative machinery for delivering insurance. An alternative view is that given the less than 10% coverage by government insurance the private sector can carve out a reasonable market for itself based on improved efficiency, better design and superior services. An alternative to public-private competition is public-private partnership in providing agriculture insurance."

He emphasizes that private participation will be possible with appropriate risk sharing between the government and private insurers. The US and Spanish systems are examples of such private participation. Private sector can be expected to improve efficiency and quality of service. Given the magnitude of agriculture risks, the actuarially fair premium for insurance may not be affordable for farmers. Government initiatives are necessary to reduce the risks of agriculture through expansion of canal and other irrigation facilities and development of an integrated market for agricultural products. Rainfall is the major yield risk factor in Indian Agriculture. This is because the irrigation system is inadequate and unreliable.

In October 1965, the Government of India decided to draw up a Crop Insurance Bill and a model scheme of crop insurance in order to enable the States to introduce crop insurance. In March 1970 the draft bill and Model Scheme was referred to an expert committee with Dr. Dharam Narain as the Chairman. Prof V.M. Dandekar examined in detail the arguments of the Expert Committee and strongly advocated the introduction of crop insurance (Dandekar 1976).

Dandekar also states that “the principal actuarial aspect of a crop insurance scheme is the year to year variation in crop yields”. Thus, “in many years the amount of premia received will nearly balance the amount of indemnities paid, though in some years the premia received will exceed the indemnities paid out and vice-versa”.

Diversification across time ensures that over a sufficiently long period of time the insurer breaks even, but still has to withstand year to year fluctuations in profits. Even though the farm level approach is the best from the perspective of reducing the risk, the area approach is the preferred alternative in terms of the administrative costs of risk assessment and loss estimation, as well as being less susceptible to the moral hazard problem.

1.3.3 Adverse selection problem and compulsory insurance

According to Dandekar, “There is no such danger (of adverse selection) in a scheme based on the area approach. In such schemes all participants are exposed to the same risk, which is determined by an independent chance system and in which the individual experience does not count”. However, it was felt that participation would be limited and premium collection difficult if the insurance were not made compulsory. “It is feared, with considerable justification, that collection of premium will prove a formidable if not an impossible task. The reasons are two fold. Firstly, there is the general difficulty of collecting any cash payments from the farmers on a regular basis because, for a large majority of the farmers, the cash position is usually nil if not minus. Secondly, in relation to the somewhat better-off farmers, as the Expert Committee says ‘even if, in the initial stages, the farmers are attracted towards the Crop Insurance Scheme in the hope that they would get, in bad years, indemnities in excess of the premium payments, it would be difficult to sustain their interest in the Scheme once they realise that they would, by large be getting back what they pay over a period.’ Collection of premium from the farmers is undoubtedly a difficult task and may prove impossible if this to be done on a voluntary basis".

On this basis Dandekar recommended that the "crop insurance scheme should be linked, on a compulsory basis, with the crop loan system.... The entire amount of the crop loans should be insured. Premium should be deducted while advancing the loan. Indemnities when they become payable should be adjusted against the recovery of the loan". The main advantage of this approach is that, "Not only the scheme can immediately get off the ground but there will be hardly any administrative costs involved". This was also expected to solve the problem of loan recovery since, "the entire agricultural credit structure is in urgent need of protection from the hazards of agriculture and this can be done only by means of an appropriate crop insurance scheme suitably linked to the agricultural credit structure". A non-borrower farmer could also take the insurance on a voluntary basis.

While Dandekar proceeded largely on the basis of a self-supporting scheme he did not rule out "legitimate grounds for a certain amount of subsidy". Dandekar suggested that less risky areas should be charged "slightly higher, but only slightly, higher premium than warranted" to subsidize more risky areas. This implies that while more risky areas would be charged higher premium than less risky areas, the difference would be less than the actuarial amount. Dandekar also provided for direct subsidy of high risk areas and of small and marginal farmers.

Chapter 2

2.1 Data source

We have obtained the following districts' time series data on the half-yearly yield of rice from 1966 to 2004 - Chengalpattu, Coimbatore, Madurai, Salem, Trichy where rice is grown in winter and groundnut in summer from SDDS-DES (Special Data Dissemination Standards - Directorate of Economics and Statistics, Ministry of Agriculture, GOI). It is to be noted that groundnut is not insured.

Source : Cropping Pattern (Agricultural and Horticultural) in Different Zones, their Average Yields in Comparison to National Average/Critical Gaps/Reasons Identified and Yield Potential- P.Das).

Table 1 Data on yield of rice for the five chosen districts from 1966 to 2004

Year	**Coimbatore**	**Madurai**	**Salem**	**Trichy**
1966	1846.15	1552.80	1659.42	1558.70
1967	1989.47	1509.32	1593.98	1452.91
1968	2348.99	1444.14	1497.28	1670.48
1969	2424.63	1675.74	1753.41	1475.60
1970	3103.38	2501.71	2154.16	1655.80
1971	2426.50	2262.63	2028.26	2009.37
1972	2296.27	2231.69	2073.12	1791.30
1973	2179.12	2246.77	2099.69	1903.52
1974	2522.20	1693.61	1635.59	1267.63
1975	2406.35	2475.02	2115.35	1883.05
1976	2580.22	2300.00	1975.00	2006.25
1977	2254.22	2686.62	2092.28	2142.80
1978	2229.68	2467.88	1963.97	2168.76
1979	2445.70	2536.21	2107.69	1746.48

Year	Coimbatore	Madurai	Salem	Trichy
1980	2552.17	2571.43	1954.81	1594.62
1981	2578.13	2799.25	2298.13	2147.32
1982	2890.11	1950.50	1953.85	1693.75
1983	2630.76	2386.78	2381.90	1780.43
1984	2929.82	2237.41	2376.47	2420.98
1985	3754.33	3022.01	2814.61	2732.43
1986	4109.89	2980.58	3131.31	2826.59
1987	3872.34	3323.74	3373.74	3036.59
1988	3124.29	3966.72	3299.00	3316.48
1989	3378.85	3803.92	3204.32	3565.63
1990	4125.33	3473.49	3270.10	2993.46
1991	3713.47	3566.58	3284.59	3400.79
1992	3554.24	3808.10	2985.31	3240.56
1993	3854.17	3620.00	3175.57	3112.07
1994	4160.72	4324.60	3562.28	3351.61
1995	4135.20	3183.93	3565.21	2491.63
1996	4068.78	3641.50	3438.89	2900.66
1997	4103.53	3929.41	3641.23	3040.40
1998	4512.25	4721.88	4171.11	3518.25
1999	4442.27	4452.19	4086.07	3831.59
2000	4307.36	4350.16	3937.81	3631.53
2001	4442.97	4188.47	4030.68	3213.27
2002	4152.13	3006.84	3004.15	2365.82
2003	3327.37	3077.22	2881.58	2350.04
2004	3845.66	3446.76	3082.50	2918.23

2.2 Methodology

Goodwin and Ker (1998) suggest an approach to normalizing yields. Assuming that the following functional relationship is appropriate for detrending a temporal series of crop yields and has been estimated for a series of yields ranging from t = 1966, …, 2004 $x_t = \beta\ y_t + e_t$, where y_t represents some linear or nonlinear function of time and x_t the actual crop yield for every farmer. Estimates of such a relationship will yield trend-predicted yields (y_t) and deviations from the trend (e_t). If one believed that the deviations from the trend tended to be proportional to the level of yields, one might consider constructing normalized yields as:

$$(\text{normalized yield})_t = x_{2004}\ (1 + e_t / x_t)$$

One must balance the complexity of empirical techniques that may be more appropriate against practical considerations associated with modeling what is often a very large collection of yield trends. We have not assumed that yield trends are linear in nature. In many cases, the exact functional form for trend effects is unknown and thus nonlinear functions including higher-ordered polynomials may be used to represent time trends.

Many researchers have observed that crop yields tend to be negatively skewed, with yields near the maximum being observed more frequently than yields near the minimum. In this light, a frequent choice for modeling yield distributions is the four-parameter beta distribution. The beta distribution can accommodate the skewness so commonly observed for crop yields and can assume a variety of shapes. This distribution does suffer from a number of shortcomings, however. Although the distribution has four parameters (two shape parameters and a maximum and minimum possible yield), it is commonly applied with only the shape parameters being estimated. The maximum and minimum yields are generally set by assumption or in some ad hoc manner. Other parametric distributions that have been used to model crop yields include the Weibull, the log-normal, the gamma, the logistic, versions of the Burr distribution, and mixtures of parametric distributions. These parametric distributions vary in terms of their flexibility and ability to capture intrinsic properties of various crop yields and thus they differ in terms of their appropriateness for modeling crop yield densities. For example, the log-normal distribution imposes positive skewness on the distribution, a characteristic that is not typically expected for crop yields.

In the case of a yield insurance contract that pays indemnity if realized yields fall beneath some threshold that defines a guarantee, two fundamental components are inherent in such an insurance contract. First, one must establish the guarantee that determines the conditions under which indemnity payments will be paid. The yield guarantee establishes total liability (i.e., the maximum possible indemnity or, equivalently, the amount of indemnity paid in the event of a total loss). Second, one must establish the appropriate price (premium) that should be charged for the coverage offered under the contract. An error in either of them can be costly for the insurer and/or the insured. Both the premium, which reflects the probability and expected level of loss, and the insurance guarantee, which is generally set to reflect the expected yield, are elements that depend upon the underlying yield distribution. Thus, accuracy in the determination of insurance premium is linked to a representation of the distribution of yields. Most crop insurance contracts establish the yield guarantee to reflect a proportion of the expected yield. In the U.S. crop insurance program, growers are able to select to insure from 50-85% of their expected yields. Expected yields are generally measured by taking an average of previous realized yields. The insurance premium is usually expressed as a "premium rate" which represents the rupees paid in premium for per hectare of cultivated area. In this dissertation, we have calculated the premium in terms of yield. An actuarially fair premium will equal the expected insured loss (expected indemnities).

We consider an insurance contract that will pay indemnities anytime yields fall beneath a certain proportion λ of the expected insured yield μ (guaranteed yield). Expected insured loss will be given by the product of the probability that a loss will be realized times the expected loss, given that a loss occurs. In other words, Expected insured loss (x) = $E\max[\lambda\mu - x, 0] = \text{Prob}[x < \lambda\mu]\,[\lambda\mu - E(x|x < \lambda\mu)]$.

2.3 Procedure

We start with the detrending process where we regress the yield series (x) of each district against a function of time-years (y). The time functions we have considered are as follows:

1) Linear : $x_t = a + by_t + e_t$ t = 1966, 1967 ….. 2004
2) Quadratic : $x_t = a + by_t + cy^2_t + e_t$
3) Cubic : $x_t = a + by_t + cy^2_t + dy^3_t + e_t$
4) Fourth degree polynomial: $x_t = a + by_t + cy^2_t + dy^3_t + fy^4_t + e_t$
5) Fifth degree polynomial : $x_t = a + by_t + cy^2_t + dy^3_t + fy^4_t + gy^5_t + e_t$
6) Log-linear : $\ln(x_t) = a + by_t + e_t$
7) Linear-log : $x_t = a + b\ln(y_t) + e_t$
8) Log-log : $\ln(x_t) = a + b\ln(y_t) + e_t$
9) Reciprocal : $x_t = a + b(1/y_t) + e_t$

After performing these regressions, we choose the best model by choosing the model with the least 'square root of the sum of squared residuals' followed by the normalization of yield to 2004 level which is the latest available data.

Table 2 Square root of sum of squared residuals for the time trend models considered for the five chosen districts

District	Linear	Quadratic	Cubic	Reciprocal	Log-log
Chengalpattu	2254.6	2227.3	2226.8	2251.3	2466
Coimbatore	2443.9	2429.8	2429.4	2441.6	2557.5
Madurai	2946.2	2820.4	2819.2	2939.4	3193
Salem	2289.6	2252.7	2251.9	2285.9	2457.7
Trichy	2766	2637.4	2635.8	2760.5	2936.9
District	**4th degree polynomial**	**5th degree polynomial**	**Log-linear**	**Linear-log**	**Minimum deviation**
Chengalpattu	2226.3	2006.6	2472.1	2252.9	**2006.6**
Coimbatore	2428.9	2041.6	2562.1	2442.8	**2041.6**
Madurai	2818.1	2577.8	3199.6	2942.8	**2577.8**
Salem	2251	1793	2463.4	2287.7	**1793**
Trichy	2634.2	2199.9	2941.8	2763.2	**2199.9**

We find that the fifth degree polynomial of time gives the best fit model for every district.

The next step would be to apply the normalizing technique suggested by Goodwin and Ker as below:

$(\text{normalized yield})_t = x_{2004} (1 + e_t / x_t)$

The normalized yield is as follows:

Table 3 Normalized yield for the five chosen districts

Year	Chengalpattu	Coimbatore	Madurai	Salem	Trichy
1966	2205.0	2694.1	2826.9	2613.6	2569.8
1967	2557.0	3217.8	2778.3	2633.4	2501.2
1968	1378.6	3987.9	2601.1	2513.2	3022.8
1969	3360.9	4134.5	3147.1	3080.2	2681.8
1970	3566.6	4949.2	4366.9	3696.8	3074.8
1971	3099.2	4179.8	4040.5	3538.5	3551.7
1972	3418.3	3941.2	3915.8	3559.0	3228.7
1973	3920.4	3728.3	3876.7	3582.7	3371.5
1974	3116.6	4197.2	2746.6	2761.7	2021.7
1975	3781.1	3952.7	3957.1	3463.1	3191.5
1976	3592.0	4102.5	3603.1	3162.9	3250.7
1977	3384.3	3463.7	3965.1	3226.8	3318.2
1978	2710.0	3260.3	3561.6	2890.7	3223.7
1979	2742.8	3524.6	3525.5	2995.6	2458.2
1980	2277.3	3545.4	3425.9	2586.8	1951.1
1981	2495.1	3450.7	3587.1	2999.0	2830.7
1982	1395.1	3784.4	1975.1	2295.0	1872.3
1983	2307.1	3205.1	2685.5	2818.1	1850.3
1984	2074.5	3567.1	2285.9	2718.4	2828.1
1985	2936.2	4325.9	3326.7	3097.8	3021.6
1986	3129.9	4521.1	3162.6	3308.8	3019.2

Year	Chengalpattu	Coimbatore	Madurai	Salem	Trichy
1987	2769.2	4223.0	3445.2	3420.7	3124.0
1988	3205.3	3291.1	3936.4	3284.6	3293.2
1989	3314.0	3505.3	3724.4	3103.2	3397.9
1990	3268.9	4158.4	3323.6	3062.2	2836.9
1991	3368.2	3712.5	3371.8	3032.2	3170.5
1992	3028.5	3435.0	3523.6	2628.0	2973.8
1993	2933.1	3696.5	3293.0	2777.7	2812.3
1994	3102.1	3953.2	3854.6	3116.3	3017.3
1995	2819.6	3866.3	2707.9	3065.2	1995.4
1996	2718.4	3777.7	3218.3	2934.0	2541.9
1997	3212.2	3776.6	3468.1	3089.8	2690.1
1998	3311.9	4118.7	4041.9	3476.8	3136.1
1999	3238.7	4074.9	3888.9	3440.0	3399.5
2000	3081.9	3977.2	3842.1	3360.1	3314.1
2001	3197.9	4141.0	3779.5	3479.5	3083.4
2002	2766.8	3920.6	2607.2	2606.3	2194.0
2003	2660.1	3104.1	2827.0	2581.0	2367.0
2004	2896.3	3815.7	3365.9	2923.7	3193.4

Once we have obtained the normalized yield, we can fit the distribution to them and estimate the parameters for each district, after which we apply the formula

Prob $[x < \lambda\mu]$ $[\lambda\mu - E(x|x < \lambda\mu)]$ to find the premium in terms of yield (kg/hectare).

The following distributions were considered as candidate distributions for the crop yield model- Pareto, generalized Pareto, Weibull, gamma, Burr and other versions of the Burr distribution. On the basis of the chi-square goodness of fit test (see appendix section 1) we select the best fitting distribution. On estimation of distributions, we found that Weibull distribution suits four of the five the crop yield distributions and four-parameter Beta for the fifth district.

2.4 Distribution fitting results

Chengalpattu district:

The distribution of crop yield in this district was found to be negatively skewed.

Table 4 Distribution results of Chengalpattu district

Distribution	Parameters	Chi-Squared Statistic
Burr	k=253.44 α=6.7769 β=7111.9	2.65
Gen. Gamma	k=0.96913 α=25.172 β=104.99	9.55
Gen. Gamma (4P)	k=16.454 α=19.783 β=31871.0 γ=-35227.0	3.88
Lognormal	σ=0.22117 μ=7.9618	12.84
Lognormal (3P)	σ=0.02361 μ=10.061 γ=-20476.0	5.23
Normal	σ=554.82 μ=2931.8	4.22
Pareto	α=1.3644 β=1378.6	19.48
Pareto 2	α=78.499 β=2.3496E+5	118.91
Weibull	α=4.9696 β=3173.8	3.05
Weibull (3P)	α=26.899 β=11929.0 γ=-8757.9	0.60

Figure 1 Weibull distribution fitting for Chengalpattu district

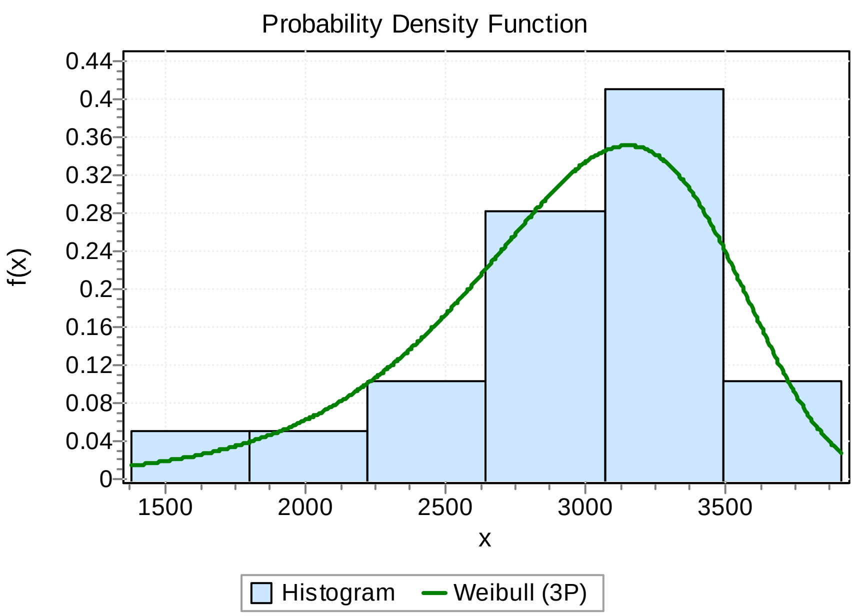


Table 1 shows that Weibull distribution gives the least chi-square test statistic and therefore suiting as the best fit for Chengalpattu district.

Coimbatore district:

Figure 2 Weibull distribution fitting for Coimbatore district

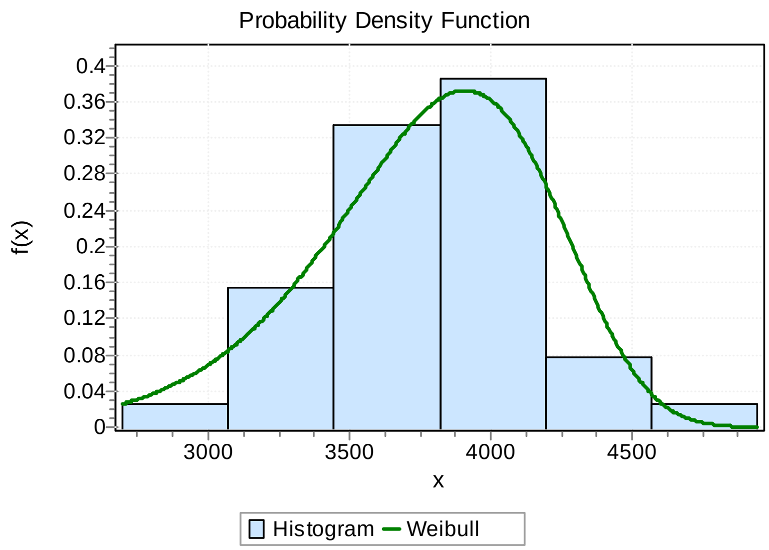


Table 5 Distribution fitting results for Coimbatore district

Distribution	Parameters	Chi-Squared Statistic
Burr	k=2.0949 α=13.179 β=4095.5	0.99
Gen. Gamma	k=1.0005 α=78.635 β=48.459	0.52
Gen. Gamma (4P)	k=2.9407 α=54.409 β=2361.2 γ=-5370.1	0.18
Lognormal	σ=0.11395 μ=8.2369	0.19
Lognormal (3P)	σ=0.03156 μ=9.4984 γ=-9541.3	0.23
Normal	σ=429.24 μ=3802.1	0.19
Pareto	α=2.9576 β=2694.1	21.24
Pareto 2	α=58.054 β=2.1680E+5	145.85
Weibull	α=10.583 β=3948.4	0.38

We find that though generalized gamma gives the least Chi-square test statistic, we choose Weibull for this district too for the sake of uniformity and because the chi-squared statistic is not too high being at 0.38.

Madurai district:

Figure 3 Weibull distribution fitting for Madurai district

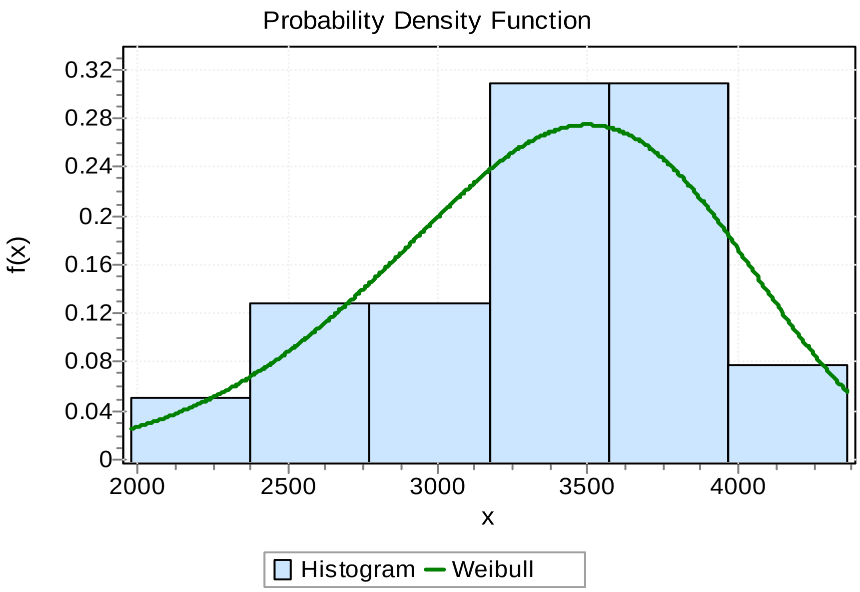


Table 6 Distribution fitting results for Madurai district

Distribution	Parameters	Chi-Squared Statistic
Burr	k =215.53 α=7.664 β=7249.5	2.72
Gen. Gamma	k =0.98932 α=36.652 β=88.517	6.37
Gen. Gamma (4P)	k =2.8561 α=137.04	5.94
	β =3232.3 γ=-14713.0	
Lognormal	Σ=0.17262 μ=8.1097	6.36
Lognormal (3P)	σ=0.03157 μ=9.7452 γ=-13701.0	6.01
Normal	Σ=546.48 μ=3373.9	6.83
Pareto	α =1.9181 β=1975.1	10.96
Pareto 2	α =123.49 β=4.3060E+5	64.66
Weibull	α =6.6298 β=3584.4	1.48
Weibull (3P)	α =13.448 β=6149.9 γ=-2539.9	1.93

Weibull distribution gives the least chi-square test statistic, proving to be the best fit for Madurai district's yield realizations.

Salem district:

Figure 4 Weibull distribution fitting for Salem district

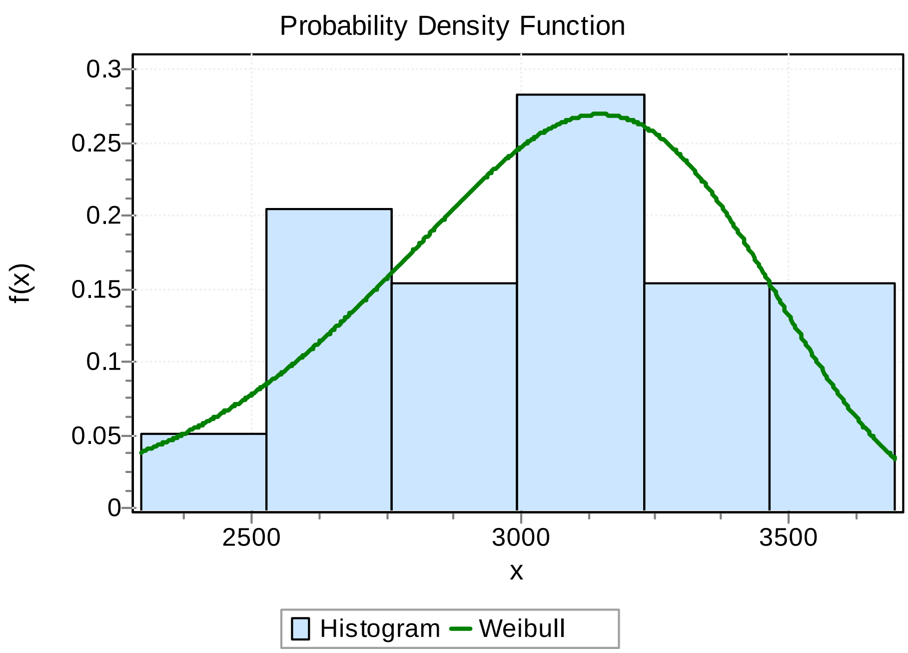


Table 7 Distribution fitting results for Salem district

Distribution	Parameters	Chi-Squared Statistic
Burr	k=3.3308 α=11.11 β=3505.6	0.64
Gen. Gamma	k=1.0015 α=75.439 β=40.687	2.56
Gen. Gamma (4P)	k=2.2003 α=69.803 β=930.55 γ=-3348.2	2.24
Lognormal	σ=0.11582 μ=8.016	1.80
Lognormal (3P)	σ=0.03948 μ=9.0892 γ=-5815.7	1.29
Normal	σ=352.23 μ=3049.3	2.19
Pareto	α=3.6031 β=2295.0	7.99
Pareto 2	α=68.837 β=2.0348E+5	143.30
Weibull	α=9.9158 β=3182.7	0.47

Weibull again gives the best fit for Salem district which can be seen from the least chi-square test statistic.

Trichy district:

Figure 5 Beta distribution fitting for Trichy district

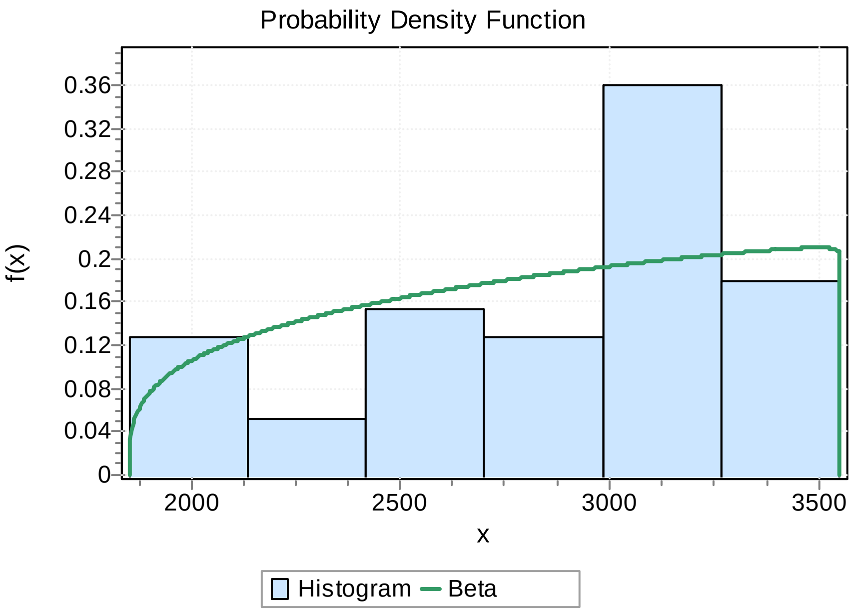


Table 8 Distribution fitting results for Trichy district

Distribution	Parameters	Chi-Squared statistic
Beta	$\alpha_1=1.2992$ $\alpha_2=1.0095$ $a=1850.3$ $b=3551.7$	0.77
Burr	$k=11.334$ $\alpha=8.4688$ $\beta=4046.1$	2.67
Burr (4P)	$k=0.20271$ $\alpha=0.78001$ $\beta=1.4327$ $\gamma=1850.3$	29.74
Gamma	$\alpha=36.384$ $\beta=78.493$	4.27
Gamma (3P)	$\alpha=139.36$ $\beta=41.703$ $\gamma=-2957.1$	7.56
Gen. Gamma	$k=0.98571$ $\alpha=34.551$ $\beta=78.493$	4.50
Gen. Gamma (4P)	$k=7.3929E+7$ $\alpha=3.1133$ $\beta=5.4251E+10$ $\gamma=-5.4251E+10$	2.56
Lognormal	$\sigma=0.17928$ $\mu=7.942$	6.47
Lognormal (3P)	$\sigma=0.03169$ $\mu=9.6063$ $\gamma=-12001.0$	3.02
Normal	$\sigma=473.47$ $\mu=2855.9$	3.05
Pareto	$\alpha=2.3868$ $\beta=1850.3$	9.04
Pareto 2	$\alpha=78.499$ $\beta=2.3496E+5$	87.15
Weibull	$\alpha=6.237$ $\beta=3049.9$	3.81
Weibull (3P)	$\alpha=6.4478E+7$ $\beta=2.2604E+10$ $\gamma=-2.2604E+10$	2.55

For Trichy district, beta distribution (4P) establishes itself as the best fit as can be seen from the least Chi-square test statistic.

The next step would be to apply the formula - Prob $[x < \lambda\mu]$ $[\lambda\mu - E(x|x< \lambda\mu)]$

For this, we need the yield guarantee for each district, and for each district we consider five coverage levels – 60%, 70%, 80%, 90%, 100%. Though 100% will not be a viable option neither for the farmer nor the insured, we still show the same so as to understand the extent to which the premium is reduced as a result of applying these coverage levels.

Chapter 3

3.1 Premium calculation:

For every district, we mostly find that the actuarial rates are higher than NAIS rates as expected. It should be noted that premium rates are only expressed in terms of yield and not in market price since we concern ourselves primarily with the comparison of them with NAIS rates which we have also calculated in terms of yield. The way we compare them with NAIS rates is that we consider the guaranteed yield as the proxy for sum insured and take the percentage stipulated for 'other rabi crops' which is 2% on guaranteed yield.

Chengalpattu district:

Table 9 Comparison of Actuarial premium and NAIS flat-rate premium for Chengalpattu district

Coverage levels (λ)	100%	90%	80%	70%	60%
Guaranteed yield	3149.77	2834.79	2519.82	2204.84	1889.86
Prob[x < λμ]	0.61	0.37	0.20	0.10	0.05
E(x|x< λμ)]	2923.40	2669.71	1773.42	833.59	310.37
Premium	670.76	350.02	285.11	193.65	96.48
NAIS premium	63.00	56.70	50.40	44.10	37.80

Coimbatore district:

When 60% coverage level is provided for Coimbatore district, the actuarial premium is actually lower than that of the subsidized NAIS premium rate for the rice crop.

Table 10 Comparison of Actuarial premium and NAIS flat-rate premium for Coimbatore district

Coverage levels (λ)	**100%**	**90%**	**80%**	**70%**	**60%**
Guaranteed yield	4015.10	3613.59	3212.08	2810.57	2409.06
Prob[x < λμ]	0.70	0.32	0.11	0.03	0.01
E(x|x< λμ)]	1060.79	357.46	95.80	20.71	3.49
Premium	1779.47	937.89	297.58	67.80	11.57
NAIS premium	72.28	65.05	57.82	50.59	43.37

Madurai district:

We find that when 100% coverage is provided for this district the actuarial premium rate is almost 12 times of that of the NAIS rate.

Table 11 Comparison of Actuarial premium and NAIS flat-rate premium for Madurai district

Coverage levels (λ)	**100%**	**90%**	**80%**	**70%**	**60%**
Guaranteed yield	3613.89	3252.50	2891.11	2529.72	2168.33
Prob[x < λμ]	0.65	0.41	0.21	0.09	0.04
E(x|x< λμ)]	1969.41	1132.81	532.39	206.89	66.00
Premium	1072.35	865.88	504.19	219.42	73.75
NAIS premium	90.35	81.31	72.28	63.24	54.21

Salem district:

The actuarial premium rate is almost half of the subsidized NAIS rate when 60% coverage is provided for Salem district.

Table 12 Comparison of Actuarial premium and NAIS flat-rate premium for Salem district

Coverage levels (λ)	**100%**	**90%**	**80%**	**70%**	**60%**
Guaranteed yield	3387.34	3048.61	2709.87	2371.14	2032.40
Prob[x < λμ]	0.84	0.48	0.18	0.05	0.01
E(x|x< λμ)]	2475.95	1305.89	449.58	112.92	21.45
Premium	768.80	835.28	415.19	118.70	23.41
NAIS premium	67.75	60.97	54.20	47.42	40.65

Trichy district:

Table 13 Comparison of Actuarial premium and NAIS flat-rate premium for Trichy district

Coverage levels (λ)	**100%**	**90%**	**80%**	**70%**	**60%**
Guaranteed yield	2895.78	2606.20	2316.62	2027.05	1737.47
Prob[x < λμ]	0.53	0.35	0.19	0.05	NA
E(x|x< λμ)]	1304.88	800.22	397.04	103.94	NA
Premium	850.83	634.71	360.65	102.58	NA
NAIS premium	57.92	52.12	46.33	40.54	NA

3.2 Key conclusion and further suggestions

Table 14 Comparison of actuarial rates and NAIS rates of premiums

Chengalpattu district	Premium	670.76	350.02	285.11	193.65	96.48
	NAIS premium	63	56.70	50.40	44.10	37.80
	Ratio	**10.65**	**6.17**	**5.66**	**4.39**	**2.55**
Coimbatore district	Premium	1779.47	937.89	297.58	67.80	11.57
	NAIS premium	72.28	65.05	57.82	50.59	43.37
	Ratio	**24.62**	**14.42**	**5.15**	**1.34**	**0.27**
Madurai district	Premium	1072.35	865.88	504.19	219.42	73.75
	NAIS premium	90.35	81.31	72.28	63.24	54.21
	Ratio	**11.87**	**10.65**	**6.98**	**3.47**	**1.36**
Salem district	Premium	768.80	835.28	415.19	118.70	23.41
	NAIS premium	67.75	60.97	54.20	47.42	40.65
	Ratio	**11.35**	**13.70**	**7.66**	**2.50**	**0.58**
Trichy district	Premium	850.83	634.71	360.65	102.58	NA
	NAIS premium	57.92	52.12	46.33	40.54	NA
	Ratio	**14.69**	**12.18**	**7.78**	**2.53**	**NA**

As expected, actuarial rates are higher than NAIS premium rates in all the districts for all coverage levels except in the case of Salem and Coimbatore at 60% coverage level. This shows that sometimes actuarial rates could prove to be cheaper to the government to employ if the loss is assessed objectively and the premium rate fixed accordingly. This could be analyzed in more detail if we know the indemnity payment scheme followed by them. In most cases, the NAIS scheme proves to be costlier to the government. In Coimbatore the actuarial rate is as high 24 times the NAIS rate. This scenario could become worse in the near future, forcing the government to let the farmers fend for themselves. A wiser strategy would be to invest in agricultural risk management practices and educating the farmers about it which would be useful and resourceful to the farmers in the long run. This could also be useful in preventing the farmers from stop cultivating altogether because of the huge risk and expenses involved in it.

Though we could have been able to work with farm level data if available, it might not work out to be the best option for the insurer because of the problem of moral hazard involved in it. Farmers might become lethargic if the indemnity payment that they would receive in the case of a loss, is dependent on their own effort put in on their farm. Instead, if like in an area approach, all insured lands could be fragmented into homogeneous regions, we might be possible to lessen this problem of moral hazard. However, there is a lot of variability across regions in rainfall distribution, soil conditions, land characteristics (like slope of the land causing water logging), fertility, and even more with regard to cultivation methods, fertilizers and pesticides used making it difficult to form homogeneous regions. Hence, a trade-off between fragmentation and usage of farm level data should be made to get the best results.

References

- Goodwin, Barry K. and Mahul, Olivier. 2004. "Risk Modeling Concepts Relating to the Design and Rating of Agricultural Insurance Contracts." Document Of The World Bank,World Bank Policy Research Working Paper 3392

- Sinha, Sidharth. June 2007 ."Agriculture Insurance in India." Centre for Insurance and Risk Management,Working Paper Series

- Raju, S.S. and Chand, Ramesh. 2008. "A Study on the Performance of National Agricultural Insurance Scheme and Suggestions to Make it More Effective," National Centre for Agricultural Economics and Policy Research (Indian Council
- of Agricultural Research), New Delhi - 110 012, (January-June 2008), Agricultural Economics Research Review, Vol. 21, pp 11-19

- Das, P. "Cropping Pattern (Agricultural and Horticultural) in Different Zones, their Average Yields in Comparison to National Average/ Critical Gaps/Reasons Identified and Yield Potential," Indian Council of Agricultural Research, New Delhi

Appendix

Section 1

Chi-square ($\chi 2$) tests is based on comparing the frequencies actually observed for the observations with the frequencies expected under some hypothesis, using the test statistic

$\sum (f_{i-} e_i)^2 / e_i$ where f_i and e_i are the observed and expected frequencies respectively in the ith category/cell, and the summation is taken over all categories/cells involved. This statistic has, approximately, a $\chi 2$ distribution under the hypothesis on the basis of which the expected frequencies were calculated.

This is investigating whether it is reasonable to regard a random sample as coming from a particular specified distribution, ie whether a particular model provides a "good fit" to the data.The test statistic is only approximately, not exactly, distributed as $\chi 2$.

Section 2

Graphical Illustration of Time trend fitting and normalization of yield realizations

Figure 6 Graph of actual yield, trend-fitting and normalized yield for Chengalpattu district

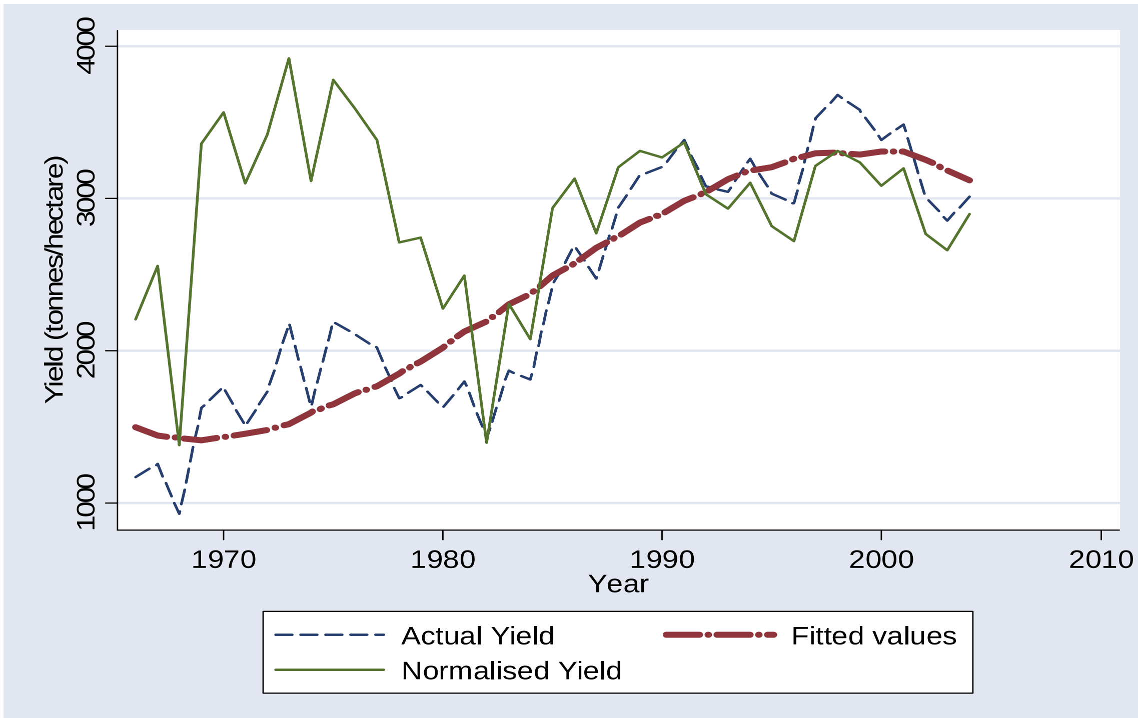


Figure 7 Graph of actual yield, trend-fitting and normalized yield for Coimbatore district

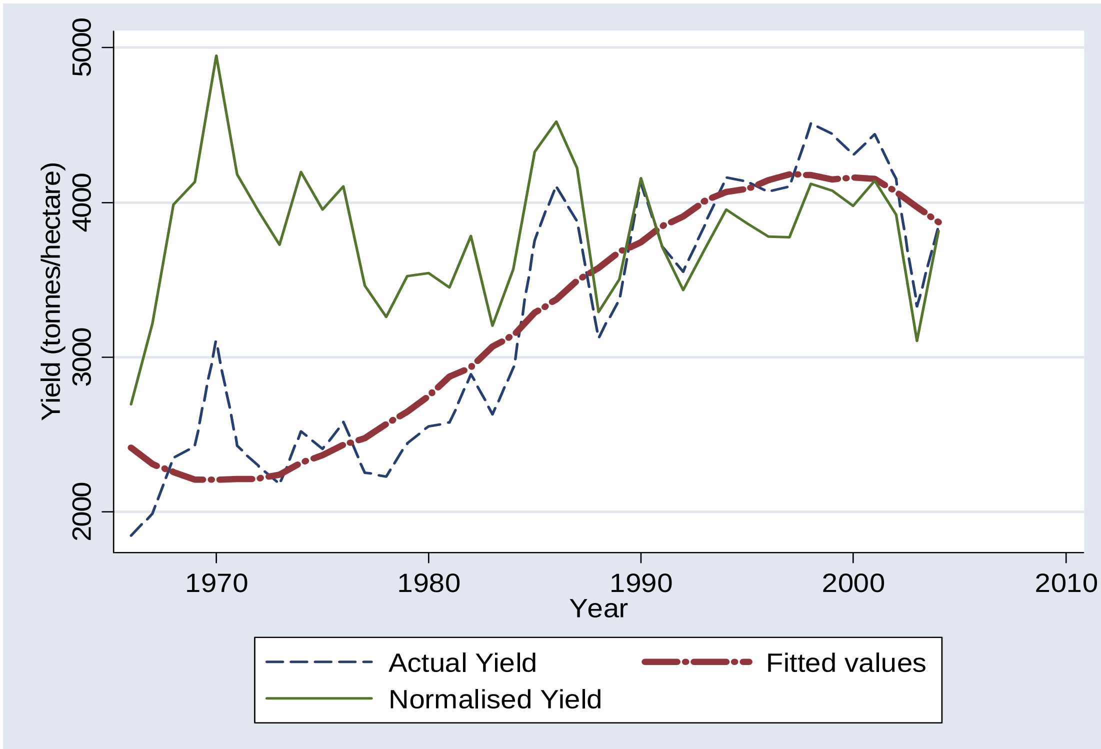


Figure 8 Graph of actual yield, trend-fitting and normalized yield for Madurai district

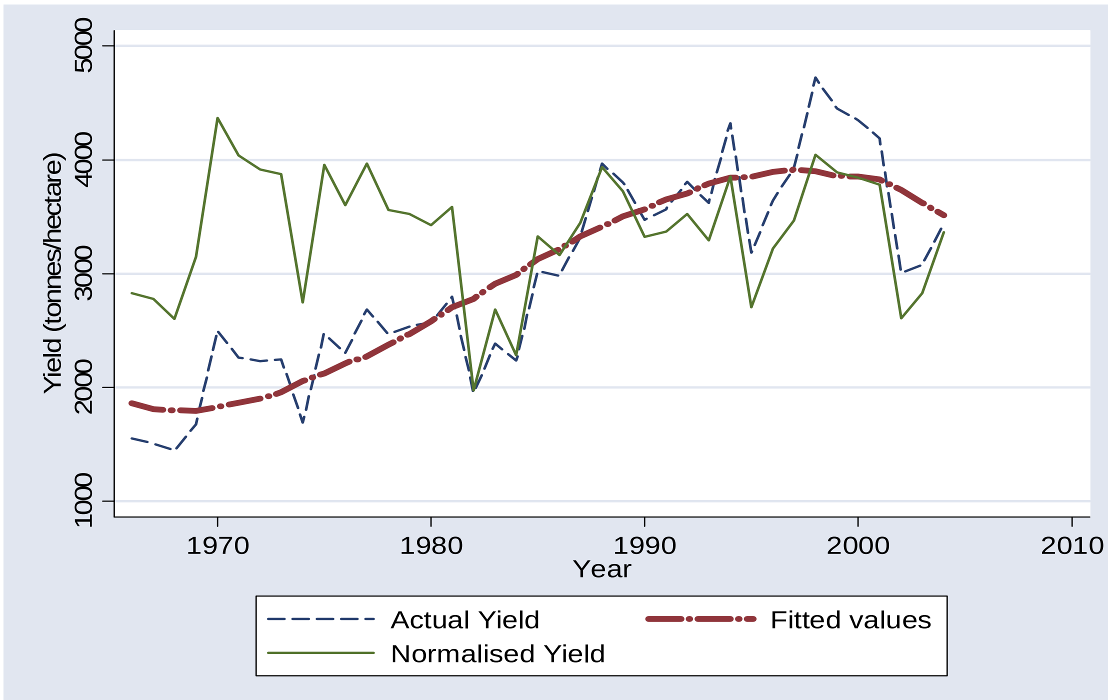


Figure 9 Graph of actual yield, trend-fitting and normalized yield for Salem district

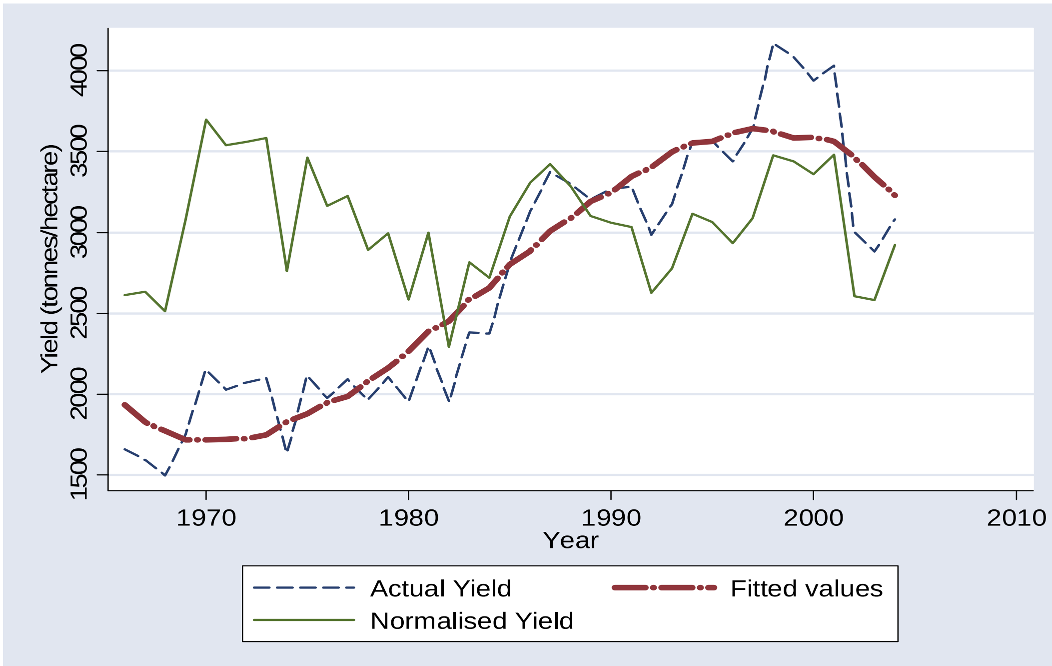
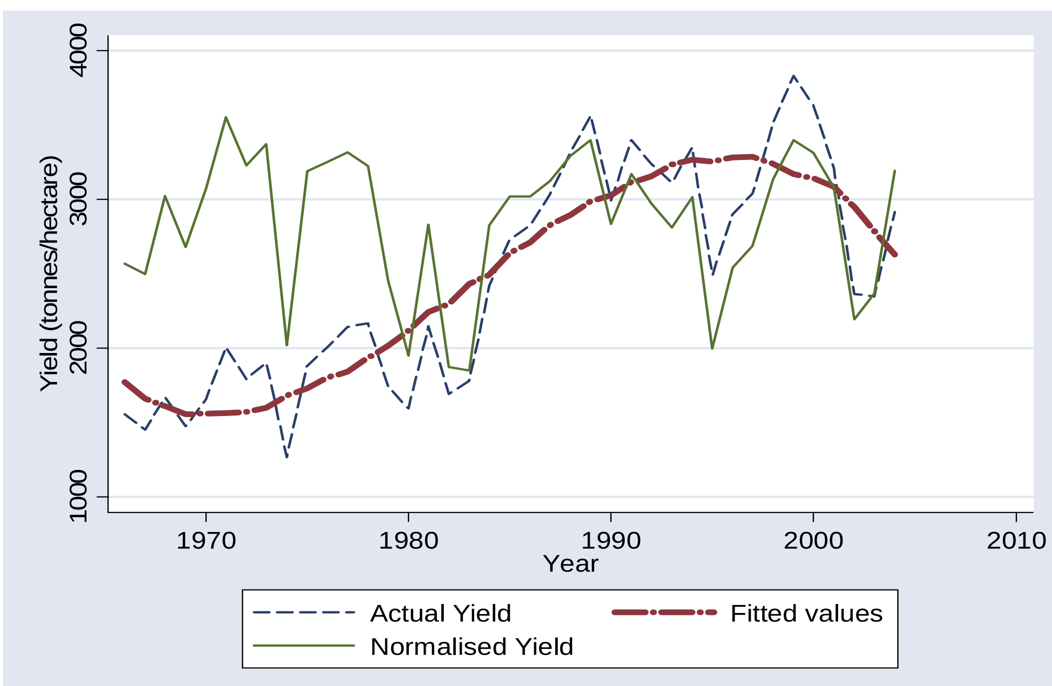


Figure 10 Graph of actual yield, trend-fitting and normalized yield for Trichy district



Section 3

Heteroskedasticity test

A random variable is said to be homoskedastic if it has the same potential distribution in all observations. If this condition is not satisfied, it is said to be heteroscedastic, and clearly the possible types of heteroskedasticity are endless.

However, in one particularly common type the standard deviation of the distribution is proportional to the size of one of the explanatory variables.

The Goldfeld–Quandt test is a test for this type of heteroskedasticity. The sample is divided into three ranges. Then comparison is made between their variances – var1 and var2. If the random variable is homoskedastic, there should be no systematic difference between the two variances.

However, if the standard deviation of the distribution of the random variable is proportional to the size of the variable, one of the variances might tend to be larger than the other. If it is different, the question is whether it is significantly different. The test statistic is the F statistic.

$F(n_2, n_1) = (var_2/n_2) \div (var_1/n_1)$

n_1 and n_2 are the numbers of observations in the lower and upper ranges of observations. We reject the null hypothesis of homoskedasticity at the 0.5% level in the case of Coimbatore and Madurai only. In the case of Chengalpattu, Salem and Trichy we fail to reject the null hypothesis of homoskedasticity.

Table 15 F-test for Heteroskedasticity

	Chengalpattu	Coimbatore	Madurai	Salem	Trichy
Variance of first 15 observations	133833.71	81859.68	192452	49459.5	69809.9
Variance of last 15 observations	66513.49	113455.56	278026	182946	195509
Mean of first 15 observations	1678.52	2373.67	2143.7	1913.6	1755.15
Mean of last 15 observations	3234.22	4049.7	3786.08	3474.47	3090.66
F test statistic	2.01	0.72	0.69	0.27	0.36
F critical one-tail	2.48	0.4	0.4	0.4	0.4

Unit root test for testing whether the series needs to be detrended or not

A stationary process is a stochastic process whose joint probability distribution does not change when shifted in time or space. As a result, parameters such as the mean and variance, if they exist, also do not change over time or position.

Stationarity is used as a tool in time series analysis, where the raw data are often transformed to become stationary, for example, economic data are often seasonal and/or dependent on the price level. Processes are described as *trend stationary* if they are a linear combination of a stationary process and one or more processes exhibiting a trend. Transforming these data to leave a stationary data set for analysis is referred to as de-trending.

In time series models in econometrics, a linear stochastic process has a unit root if 1 is a root of the process's characteristic equation. The process will be non-stationary. If the other roots of the characteristic equation lie inside the unit circle, then the first difference of the process will be stationary. The Dickey–Fuller test tests whether a unit root is present in an autoregressive model. It is named after the statisticians D. A. Dickey and W. A. Fuller, who developed the test in 1979.

A simple auto regressive model of order 1 is

$$y_{t} = \rho y_{t-1} + u_t$$

where y_t is the variable of interest, t is the time index, ρ is a coefficient, and u_t is the error term. A unit root is present if $\rho = 1$. The model would be non-stationary in this case. Since the test is done over the residual term rather than raw data, it is not possible to use standard t-distribution to provide critical values. Therefore this statistic has a specific distribution simply known as the Dickey–Fuller table.

The intuition behind the test is as follows. If the series y is (trend-)stationary, then it has a tendency to return to a constant (or deterministically trending) mean. Therefore large values will tend to be followed by smaller values (negative changes), and small values by larger values (positive changes). Accordingly, the level of the series will be a significant predictor of next period's change, and will have a negative coefficient. If, on the other hand, the series is integrated, then positive changes and negative changes will occur with probabilities that do not depend on the current level of the series; in a random walk, where you are now does not affect which way you will go next.

In this case, we clearly see that unit root is present in the yield realizations of all the districts, requiring the series to be detrended.

Table 16 Dickey-Fuller test for unit root

Districts	**Chengalpattu**	**Coimbatore**	**Madurai**	**Salem**	**Trichy**
Test Statistic Z(t)	-1.639	-1.826	-2.037	-1.503	-1.823
5% Critical Value	-2.964	-2.964	-2.964	-2.964	-2.964
P-value	0.463	0.3675	0.2706	0.532	0.369

Section 4

Loss distribution and estimation

The gamma distribution

It is possible to use the method of maximum likelihood (ML) or the method of moments to estimate the parameter of the gamma distribution. The moments have a simple form and so the method of moments is very easy to apply. The MLEs for the gamma distribution cannot be obtained in closed form (ie in terms of elementary functions) but the moment estimators can be used as initial estimates in the search for the maximum likelihood estimates. It is more convenient to obtain maximum likelihood estimates for the gamma distribution using a different parameterization based on the invariance property of maximum likelihood estimators.

The normal distribution

The method of moments and maximum likelihood estimation are both straightforward to apply in this case.

The Pareto and generalized Pareto distributions

The method of moments is very easy to apply in the case of the Pareto distribution, but the estimates obtained in this way will tend to have rather large standard errors, mainly because S^2 , the sample variance, has a very large variance. However, the method does provide initial estimates for more efficient methods of estimation that may not be so simple to apply, like maximum likelihood, where numerical methods may need to be used. For the generalized Pareto, things are not quite so easy. As for estimation, the CDF does not exist in closed form, so the method of percentiles is not available. ML can be used, but again suitable computer software is required; the method of moments can provide initial estimates for any iterative scheme.

The lognormal distribution

Estimation for the lognormal distribution is straightforward since μ and σ^2 may be estimated using the log transformed data. Alternatively the method of moments can be used.

The Weibull and Burr distributions

Neither the method of moments nor maximum likelihood is elementary to apply if both c and γ are unknown. In the case where γ has the known value γ , maximum likelihood is easy enough. The distribution function of the W(c, γ) distribution is an elementary function, and a simple method of estimation of both c and γ can be based on this fact. The method involves equating selected sample percentiles to the distribution function; for example, equate the quartiles, the 25th and 75th percentiles, to the population quartiles. This corresponds to the way in which sample moments are equated to population moments in the method of moments. This method will be referred to as the method of percentiles.In the method of moments the first two moments are used if there are two unknown parameters, and this seems intuitively reasonable (although the theoretical basis for this is not so clear). In a similar fashion, the median would be used if there were one parameter to estimate. With two parameters, the best procedure is less clear, but the lower and upper quartiles seem a sensible choice.

Totally different estimates could be obtained using the method of moments. It turns out that for the Pareto distribution this method is very unreliable unless you use very large samples on which to base your estimates. In this particular case the method of percentiles is unlikely to give you reasonable estimates unless you use samples of, say, 1,000 or more. We now turn to the Burr distribution. Since the CDF exists in closed form, it may be possible to fit the Burr distribution to data by using the method of percentiles; ML will certainly require the use of computer software that allows non-linear optimization.