Risk in Agriculture Credit Applications: A New Approach

For most farmers in developing countries, access to finance remains difficult despite agriculture's economic importance. The causes are manifold, but usually boil down to the following reasons:

1) Small loan sizes with high expenses per monetary unit lent out;
2) Low creditworthiness due to lack of stable income, limited assets, and absence of clear financial data, and;
3) Systemic risks in agriculture, such as droughts, crop failures, animal diseases, etc.

In this paper, we will present the Agri-Risk Analyser (ARA) which is a powerful tool to assist banks and MFIs in quickly assessing the credit risk on a per-farmer level, as well as for the credit portfolio, based on a practical, hands-on method.

The ARA-approach is based on the following principles:

1) A farmer's creditworthiness can be assessed by predicting the future incoming and outgoing cash flows;
2) The cash flows are based on generic, standardised data regarding crop yields and prices. This circumvents elaborate questionnaires aimed at grasping the farmer's financial situation;
3) The uncertainty surrounding each of these parameters is expressed in probability distributions linked to the inputs, and picked up by the ARA;
4) The probabilistic approach allows for running Monte Carlo simulations. The simulation result will not only reveal the expected default rates, but also the risk expressed in standard deviations as well as other interesting breakdowns of the risk factors.

In the following paragraphs we will zoom in on these principles. The examples given will be geared to the situation in Kenya but the principles are, of course, equally applicable to other economies.

Cash flows

Clearly, an assessment of the famer's future incoming and outgoing cash flows would be instrumental in developing a credit score.

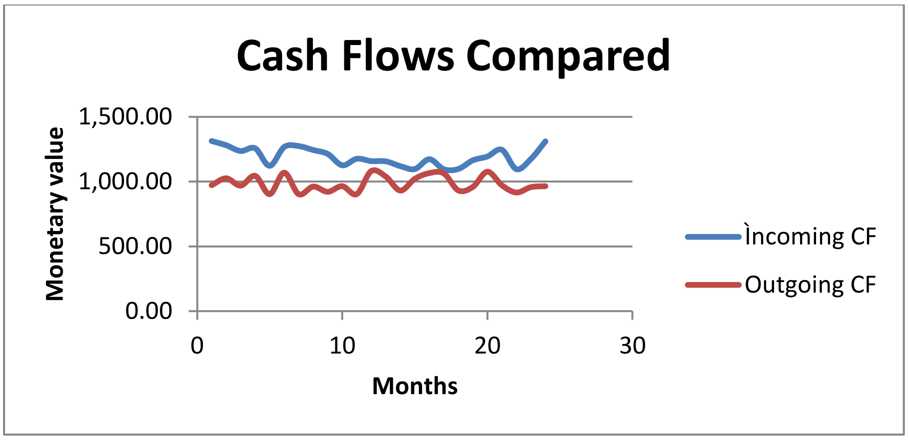

Figure 1: Projection of future cash flows

As long as the projected incoming cash flow is higher than the outgoing cash flow, the farmer is financially healthy, while the reverse indicates future liquidity problems. Farmers' savings and other financial buffers should be included in this model. A Monte Carlo simulation can generate a thousand different developments and versions of the predicted cash flows. If the incoming cash flow exceeds the outgoing flow in 980 of these simulated cases, we may conclude that the farmer remains liquid with a certainty of 980/1,000 or 98%. The probability of default (PD) would then be its complement, so 100% minus 98% equals 2%. Intuitively, we understand that the further apart these cash flows are, the more financially robust the farmer is. Furthermore, it is understood that the fact that the incoming cash flow exceeds the outgoing flow most of the time will not be good enough.

Standardised data

Interviewing farmers and trying to pin down precise financial data can be a challenge. Lack of a proper financial administration, literacy issues, and the dominant role of cash stymie the process. To a large extent, we can sidestep these problems by reverting to commonly known data sets. For instance, crop yield data for the various Kenyan regions are publicly available. This means that once we know the surface tilled for each of the various crops we can arrive at a good estimate of the future harvest yields.

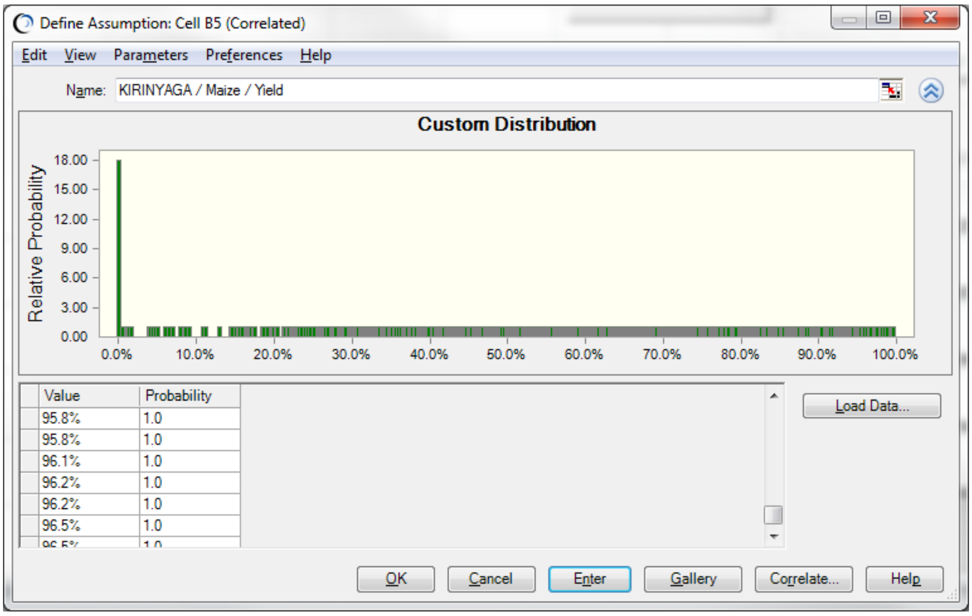

Figure 2: Custom distribution reflecting expected crop yield in % of maximum yield (example: Maize, Kenya, Kirinyaga)

More importantly, we get a better picture concerning the risk enveloping these yields, which will include crop diseases, climatological disasters, and so on. Typically, these risks are substantial. Useful data are available for most crops, for every region.

Yields differ from crop to crop, and growing various crops helps to reduce risk through the creation of a portfolio. This portfolio effect might be dulled by correlations. For example, potato- and sweet potato yields are highly correlated, leaving little room for portfolio benefits. On the other hand, the correlation between tea and potato yields is a lot less and offers the possibility of diversification. These effects should be considered.

Correlations	KIRINYAGA / Maize / Yield	KIRINYAGA / Wheat / Yield	KIRINYAGA / Dry Beans / Yield	KIRINYAGA / Potatoes / Yield	KIRINYAGA / Sweet Potatoes / Yield	KIRINYAGA / Tea / Yield
KIRINYAGA / Maize / Yield	1.0	0.926659555	0.789979071	0.797356554	0.858897318	0.94041499
KIRINYAGA / Wheat / Yield	0.906940398	1.0	0.819913516	0.686765139	0.753074325	0.982368643
KIRINYAGA / Dry Beans / Yield	0.739622066	0.758950694	1.0	0.891164475	0.909283724	0.7968262
KIRINYAGA / Potatoes / Yield	0.7546389	0.672914452	0.782999073	1.0	0.990081066	0.688119178
KIRINYAGA / Sweet Potatoes / Yield	0.793523494	0.718817098	0.839028545	0.898204013	1.0	0.755691475
KIRINYAGA / Tea / Yield	0.786699761	0.825426542	0.700504313	0.62061136	0.656268334	1.0

Figure 3: Correlations between crop yields in various regions

In brief, by asking the farmer three simple questions we can assess his incoming cash flow:

1) Where is your farm?
2) How many acres?

3) What is the percentage breakdown for the land use for the various crops?

A similar approach can be taken for the outgoing cash flow. Here, we base the expenses on typical cost levels considering the family composition. Age, gender, illness, health costs, and mortality rates for the various age rates are included.

	Annual Probability of Mortality		
Age	Male	Female	Avg
20	0.29%	0.05%	0.17%
21	0.27%	0.08%	0.17%
22	0.25%	0.11%	0.18%
23	0.23%	0.12%	0.18%
24	0.22%	0.11%	0.16%
25	0.20%	0.05%	0.13%
26	0.19%	0.07%	0.13%
27	0.18%	0.09%	0.14%
28	0.18%	0.10%	0.14%
29	0.18%	0.11%	0.14%

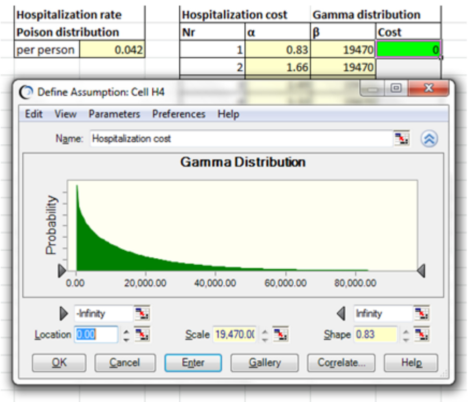

Figure 4: Mortality rates

Three simple questions will again suffice:

1) Gender and age?
2) Gender and age of partner?
3) Gender and age of children or other dependents?

Of course, leeway should be given to include specific alternative income and expenses, but, all in all, the approach described here gives a fair estimate of the average expenses.

Risk and standard deviation

For all of the inputs, not only the average values are considered but also the standard deviations or the risk.

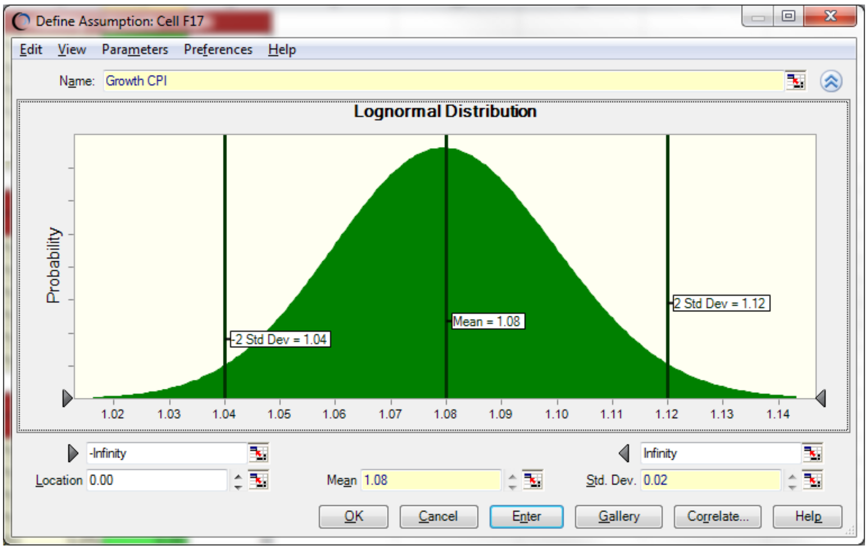

Figure 5: Modelling uncertainty regarding input parameters, e.g. growth of consumer price index

This approach avoids the pitfalls of single-point estimates that hide the inherent risks behind an illusion of accuracy. The ARA champions the view that the risk surrounding the variables should be fully integrated in a model. As bankers and risk managers, we are particularly interested in these uncertainties and would like to quantify these. Once the risk has been acknowledged and measured we can consider mitigation measures or adjust the pricing accordingly.

Monte Carlo Simulations

Now, it's time to start spinning the roulette wheel: using a Monte Carlo simulator, we will generate random numbers for each of the input variables. This process is governed by the distributions associated with these input parameters. As the input variables take on different values, the output parameter and the sum of the provisions will change as well. Thus, if we were to run 1,000 trials in a simulation, we would end up with close to 1,000 different results, while our original calculation model would produce just one outcome. We can then display those 10,000 results in a histogram with the results grouped in bins. Let's take a closer look at the results of a cash flow simulation for a sharecropper, "Farmer 1". The histogram used is basically a frequency diagram with the various amounts for the expected consolidated cash flow on the horizontal axis and the rate of occurrence on the right-hand vertical axis. It is important to realise that the higher bars in the graph correspond with an increased relative probability. So, we can conclude that the most likely cash flow projections are around 700,000.

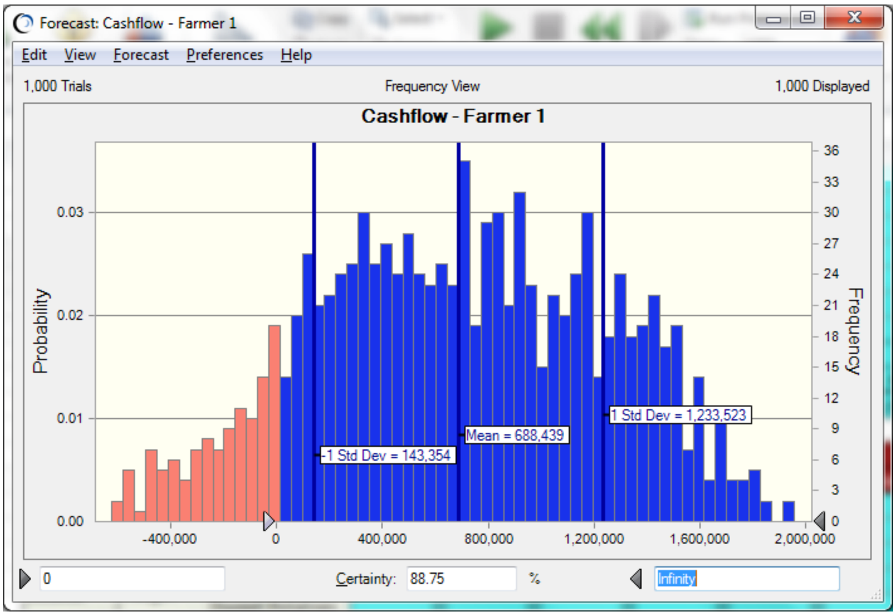

Figure 6: Histogram with simulated consolidated cash flow results, thousand trials

Interesting also is to calculate the average of the thousand simulation rounds: 688,439. However, we are not just interested in the average; we are also concerned with the risk associated with this mean. If the data would be tightly clustered in the centre, there would be little risk, if the data are scattered over a wide range, there is more risk. In the latter case, it is more likely that events occur far away from the mean, with very different cash flows from what we might expect. This spread or degree of scatter is measured in terms of the standard deviation. In our case, the standard deviation, or "σ", turns out to be 545,084. Therefore, if we would make changes to the inputs of our model we would observe the effect on the standard deviation. An increase in the standard deviation would equal higher risk levels, while a decreasing value would indicate risk mitigation. Knowing the standard deviation and knowing the risk are crucial for risk management. We can take this a step further. If we consider the thousand simulated results as the total domain of possible outcomes, we can split the results in percentages below and above a certain threshold. Putting this threshold at a cash flow of zero reveals that 88.75% of the histogram's surface is in blue, or in the positive domain. Conversely, we notice that there is chance of 11.25% that we end up with a cash flow that is negative. This is a strong indicator for the probability of default (PD) rate because it indicates that this farmer, in eleven percent of the cases, does not have a viable business, so there is default probability of 11.25%.

Knowing the risk is no doubt helpful, but equally interesting is discerning the causes of this risk. Monte Carlo simulations allow for producing sensitivity analyses that dissect the entire risk into the various risk drivers.

Sensitivity analysis - Farmer 1			
Factor		Sensitivity	
CPI		19%	Very low
Harvest margin	Maize	77%	High
	Wheat	86%	Very high
	Dry Beans	79%	High
	Potatoes	79%	High
	Sweet Potatoes	78%	High
	Tea	81%	Very high
	Coffee	61%	High
Other income			
Life costs		2%	Very low
Health costs		4%	Very low
Other expenses		19%	Very low

Figure 7: Sensitivity analysis with most important risk drivers

Identifying these risk drivers will be helpful in taking risk mitigation measures. This could be done by, for instance, advising the farmer to switch from high-risk to low-risk crops, to suggest financing for irrigation or crop protection, or to offer insurances to cover health care and life risks.

Portfolio approach

The calculation and analysis above pertained to one individual farmer. The same Monte Carlo approach can equally be applied to a portfolio of agricultural credits. To a certain extent, risks in a portfolio can be mitigated, and it is worthwhile to study the average and standard deviation of the portfolio as well as the main risk drivers. This methodology hands the risk manager proper tooling to indeed manage the risks.

Back testing

For banks and financial institutions there is the option to back test the performance of the ARA model against the historical loan application database. Since, the actual reimbursements are known the validity of the model can be checked.

Conclusion

The Agri-Risk Analyser (ARA) is a crucial tool for calculating the PD, as well as the associated risk per farmer and for a portfolio of agricultural credits. The Monte Carlo simulation also reveals the most important risk drivers. The ARA is a practical instrument for making accurate lending decisions concerning individual farm loan applications and for portfolio risk management. The information provided allows for targeted advice to the clientele, which in turn supports the cross-selling of insurance and other lending products.

André Koch (andre@stachanov.com)

Stachanov Solutions & Services bv, Amsterdam, November 2015