

RESEARCH ARTICLE

Open Access

Stochastic economic analyses of even-aged timber growing



Ian Ferguson

Abstract

Background: StochPV is a computer programme enabling the risks and uncertainties inherent in even-aged timber growing to be recognised through stochastic analysis. Most financial analyses of even-aged plantation or forest stands have used deterministic analyses, such that a single set of volumes, costs and prices yields a single answer for the present value. Risks and uncertainties are ignored, even though well-recognised by practitioners. StochPV provides a probability distribution, not just a single value, for the present value of an even-aged stand, allowing practitioners to better gauge their decisions regarding the impact of risks and uncertainties on silvicultural options and stand valuation. More generally, it will enable analysts to explore whether or when deterministic analyses are a reliable guide for such decisions.

Methods: Means and coefficients of variation, together with maximum and minimum values and correlation coefficients of prices and costs, are read in via a spreadsheet and used to estimate beta distribution parameters for each data cell. The resulting distributions can be reviewed visually and amended, if necessary. Fire frequency and fire salvage data are also read in via spreadsheet.

Univariate or multivariate pseudo-random values are simulated for each variable involved, enabling the present value to be calculated for each iteration. Present value is calculated for the stochastic analogues of Faustmann perpetual rotations and a single rotation, the latter taking account of the initial land cost and revenue at the end of that rotation. The present values of the many iterations are then used to calculate mean, standard deviation and other summary statistics.

Results: The programme integrates all known risks relating to stand management in an analytical framework that is relatively simple and communicable. The use of beta distributions provides flexibility in representing the nature of the empirical distributions involved. Inter-relationships between kindred variables are taken into account using multivariate distributions based on copulas and the marginal distributions involved.

Conclusions: The programme results are suitable for evaluation of first-degree stochastic dominance and could also be applied to second-degree stochastic dominance if the decision-maker had a material aversion to risk, all necessary data being available as outputs of the programme.

Keywords: Stochastic, Monte Carlo simulation, Beta distributions, Multivariate distributions, Copulas, Present value

Background

Deterministic analysis of the financial outcomes of forest investment is based on a single set of prices for inputs and outputs, together with a single set of quantities of those outputs and inputs. They are used to calculate annual net revenues which, in turn, are discounted to their present value and aggregated to a single value estimating

the net present value of the stand or forest or per unit area thereof.

Deterministic analyses of present value currently represent the conventional wisdom of foresters in evaluating options in the financial management of radiata pine (*Pinus radiata* D.Don) stands and similar even-aged forests. They could be correct and useful guides if no natural disasters occur and all the probability distributions attached to the value of quantities and prices were normal distributions. But foresters know they live with the risk of natural and

Correspondence: iansf@unimelb.edu.au
School of Ecosystem and Forest Sciences, The University of Melbourne,
Grattan St, Parkville, Vic 3010, Australia

other disasters, together with the risk of changes in the quantities and prices of outputs and inputs due to changes in climatic conditions and changes in markets, not to mention other sources of volatility.

So the question arises—how reliable are deterministic analyses? We obviously should not expect deterministic and stochastic analyses to produce identical present values for the same set of options. However, will they produce the same rank order for the same set of options? This computer programme enables such questions to be pursued.

Of course, there are many partial ‘fixes’ to assess the impact of individual risks, such as using a risk-weighted value for a particular price due to a natural disaster or other exogenous change (Hildebrandt and Knoke 2011). Wald’s (1945) maximin strategy can be applied to indicate the best strategy to take to avoid the worst-case outcome. Fuzzy sets (Varma et al. 2000) can be used to finesse near-objective probabilities with more subjective possibilities. Hildebrandt and Knoke (2011) note the popularity that Ben-Haim’s (2006) so-called info-gap model appears to have gained in ecosystem management, despite concerns about its objectivity. All these methods and many others may have a place for protecting an owner against the worst-case scenario. But they are specific fixes that take no account of all the interactions and impacts that go on within the greater family of risks.

What of the many options and decisions involved in managing a plantation—planting problematic locations and soil types; later tending, fertilising and thinning choices; and choosing market strategies and protective strategies against fire? What of the valuation for the sale of a particular stand? We have considerable experience from historical and other records concerning these options and sales. Their financial present values are generally assessed using deterministic analysis. Some scepticism is warranted, given that risks are not taken into account, apart from ‘fixes’ that address a single risk.

Whether we term it option analysis or robust optimisation or Bayesian analysis or just stochastic analysis, an assessment is surely needed of the impact of all known systematic risks and uncertainties on a comprehensive, integrated model. One such model is StochPV.m, which was developed by the author and is available from MathWorks <https://au.mathworks.com/matlabcentral/fileexchange/58578-stochpv/content/StochPVv3/StochPV3Functions/StochPV.m>. Aside from its merits in illuminating our understanding of the collective risks involved, it enables us to evaluate whether conventional deterministic analysis works sufficiently well in guiding us. The development of StochPV is one step in that process. Some related results using StochPV have already been useful in examining the Future Rotations Problem (Ferguson 2016) and, hopefully, more comparisons will follow.

Stochastic analyses seek to take account of the variations in the prices and quantities of inputs and outputs resulting from variations in site and location and the consequent changes in costs and yields. Likewise, variation in markets and the resulting changes in prices, costs and demand can affect the prices and quantities of outputs. Risks due to these and exogenous factors such as fire and other destructive natural disturbances need to be assessed to provide an estimate of the expected value (*syn.* mean value) of the net present value.

To estimate the mean value and its probability distribution, probability distributions have to be estimated for all prices and quantities of inputs and outputs and for the exogenous disturbances. Monte Carlo simulations (Hildebrandt and Knoke 2011) are then used to estimate the expected value. In the first run of the software, the Monte Carlo simulation draws a pseudo-random value of each variable concerned from each probability distribution. These are then used to calculate a single value of net present value, just as in the deterministic method. By repeating the simulation many thousands of times, one can estimate the probability distribution of net present value and so calculate the expected value and other relevant statistics of interest (such as the standard deviation and skewness) from it or plot the probability distribution.

The point of undertaking this process is that many of the distributions involved are not normal distributions but are skewed or have non-normal levels of kurtosis. If they were all normal and no natural disasters occur, a simpler process could be used to evaluate the mean net present value as it, with some minor caveats, also be normal. The Monte Carlo process is a relatively painless way (for the user) of taking non-normality into account in estimating the probability distribution of net present value. Fire is an obvious example, as are non-normal prices and quantities of inputs and outputs. There are also other complexities to be considered. Some probability distributions are conditional on the state of the forest or plantation at the time. For example, the proportion of sawlog volumes that can be salvaged after a fire may be conditional on the age of the stand. The cost of early years of tending after planting may be conditional on the cost of planting because it provides a proxy for the roughness or difficulty of the terrain. Conditional probabilities complicate the programming of the software but not the process of simulation, other than the time to compute.

Most textbooks on forest economics contain sections dealing with aspects of risk analysis in forest investment at an elementary level. However, Amacher et al. (2009) have contributed greatly to risk analysis at a theoretical level. Hildebrandt and Knoke (2011) provided a comprehensive review of the literature and methodology of

analysing forest investments under uncertainty. However, comprehensive and well-integrated tools for analysing investment in even-aged plantations or forests are sparse, as the Hildebrandt and Knoke (2011) review shows.

Earlier attempts to apply stochastic analyses to radiata pine plantations include those by Kirby et al. (1993), Knott et al. (1995) and Ferguson (2005). Following research by Cassidy et al. (1970) and Lothner et al. (1986), Kirby et al. (1993) used a stochastic model for the analysis of radiata pine investment options. All probability distributions were triangular and independent, and the risk of fire was not incorporated. Knott et al. (1995) also used a stochastic model to analyse the most profitable options in a thinning experiment in radiata pine. Independent triangular distributions were used for volumes and independent normal distributions for costs and prices. The risk of fire was not incorporated. Ferguson's (2005) work was a precursor to the development of StochPV.

The computing and statistical technology underpinning Monte Carlo processes has improved since those pioneering works and now enables asymmetrical distributions, correlations between variables and the risk of fire to be recognised using different distribution shapes for different inputs or outputs and preserving multivariate relationships. Throughout the financial and banking world, these processes are being applied to analyse and understand investment outcomes involving risks and uncertainties. Hildebrandt and Knoke (2011) have pointed to the need for similar development in analyses of forest investments.

Software—StochPV

The aim of this research was to develop a computer programme, StochPV (see “Availability of data and materials” section), that enables the risks and uncertainties inherent in growing timber in even-aged plantations or forest stands to be taken into account through stochastic analysis.

The programme uses a Microsoft Excel® spreadsheet¹ to set up the input data for a particular stand and sequence of harvests, which may involve up to five thinning before the final clear-felling. Up to six different log products can be accommodated.

The programme is written in Matlab® language.² It can be used to generate readable and portable C code using Matlab algorithms and can then be run on most computers using Windows® or Apple® operating systems.

An Excel® spreadsheet, ‘StochPVPaper.xlsx’, showing a complete example of data, is provided for testing the programme (see “Availability of data and materials” section). This spreadsheet can also be used as a template

for new data. Figure 1 shows the sample worksheet, ‘Volumes’, for data on log volumes from that spreadsheet. The worksheet ‘Prices’ is of identical form with appropriate changes in headings. The upper level of Fig. 2 shows the sample worksheet, ‘Costs’, truncated to the right here but otherwise proceeding in like manner to the rotation length. The lower level of Fig. 2 shows the sample worksheet ‘Disasters’, in this case, based on fires.

Three separate Matlab scripts (‘IN_Stoch’, ‘PNV_Stoch’ and ‘OUT_Stoch’) are involved in the programme and can be used as a single linked script (‘StochPV’) or as a guided user interface programme (‘GUIStochPV’). IN_Stoch deals with the relatively straightforward preparation of data input from the Excel worksheet, while Out_Stoch deals with printing and graphing of the results. PNV_Stoch is a script that computes the various financial results and data for graphing.

A flowchart of PNV_Stoch is shown in Fig. 3 and provides some visual idea of the complexity, noting that there are three user options involved. ‘Test’ provides visual displays of the beta distributions for the prices and costs. ‘Det’ provides a deterministic analysis and ‘Stoch’, a stochastic analysis. The complete programme, StochPV, and the example spreadsheet ‘StochPVPaper.xlsx’ are freely available on the Matlab File Exchange site³ under a BSD licence.

In addition to simulating the probable impact of fire, novel features of the model include the use of multivariate beta probability distributions for simulating volume, price and variable cost, based on correlation data and the use of copulas, discussed later. The iterative basis for calculating a stochastic analogue of the Faustmann (1849) formula for perpetual rotations is also novel, as is providing a choice of either deterministic or stochastic modes of analysis. Input data can be tested visually by simulating a random test value of each input variable (other than fire occurrence) and graphing the distribution of each variable to ensure the shape of the distribution is appropriate.

This paper first describes the data requirements and example data. The definition of the beta distribution parameters and their estimation for each relevant cell of the input data are described next. Monte Carlo simulation of pseudo-random numbers from the resulting distributions is then discussed. Copulas provide the basis for estimation and application of multivariate distributions, with the relevant beta distributions of input values providing the marginal distributions. Finally, the scope and limitations of the programme are discussed.

Methods

Data

This programme is for stand-level analyses—a stand being a relatively homogeneous area of trees of uniform age and silvicultural history, seldom exceeding 200 ha in

nj

A	B	C	D	E	F	G	H	I	
	No of Products	No & Volumes of Log Products					(cu m/ha)		
Products not to exceed 6	3	LProd1	LProd2	LProd3	LProd4	LProd5	LProd6		
	Harv1	90.00	0.00	0.00					
Harv# can be a thinning or a	Harv2	65.00	100.00	0.00					
clearfell but clearfell must be	Harv3	45.00	180.00	120.00					
at the rotation length	Harv4								
and no entries beyond	Harv5								
	Harv6								
		Coefficients of Variation of Log Vols					(%)		
Prod# by Harv#		LProd1	LProd2	LProd3	LProd4	LProd5	LProd6		
	Harv1	35.00	0.00	0.00					
	Harv2	25.00	35.00	0.00					
	Harv3	10.00	35.00	30.00					
	Harv4								
	Harv5								
	Harv6								
		Correlation Coeff Matrices of Log Vols					Max 1.0 diags		
Prod# against Prod# by Harv#		LProd1	LProd2	LProd3	LProd4	LProd5	LProd6		
Harv1	LProd1	1	0	0	0	0	0		
	LProd2		1	0	0	0	0		
	LProd3			1	0	0	0		
	LProd4				1	0	0		
	LProd5					1	0		
	LProd6						1		
Harv2	LProd1	1	0.3	0	0	0	0		
	LProd2		1	0	0	0	0		
	LProd3			1	0	0	0		
	LProd4				1	0	0		
	LProd5					1	0		
	LProd6						1		
Harv3	LProd1	1	0.3	0.1	0	0	0		
	LProd2		1	0.7	0	0	0		
	LProd3			1	0	0	0		
	LProd4				1	0	0		
	LProd5					1	0		
	LProd6						1		
Harv4	LProd1	1							
	LProd2		1						
	LProd3			1					
	LProd4				1				
	LProd5					1			
	LProd6						1		
Harv5	LProd1	1							
	LProd2		1						
	LProd3			1					
	LProd4				1				
	LProd5					1			
	LProd6						1		
Harv6	Prod1	1							
	Prod2		1						
	Prod3			1					
	Prod4				1				
	Prod5					1			
	Prod6						1		

Fig. 1 Sample Excel spreadsheet for input of volume data

A	B	C	D	E	F	G	H	I	J	K	L	
ENTER DATA IN COLOURED CELLS: NO NEED TO ENTER BEYOND ROTATION AGE												
	Start Age	Clearfell age	Last year in which Estab Cost was incurred					No of fallow years after clearfell				
	0	35	4					1				
Maintenance Costs & Coefficients of Variation (\$/ha & %)												
By Yrs after Planting	0	1	2	3	4	5	6	7	8	9	10	
Maintenance Coss/ha	121	121	121	121	121	121	121	121	121	121	121	
Coefficients of variation	5	5	5	5	5	5	5	5	5	5	5	
(\$/ha & %)												
Estab & Other Var Cost/	673	1211	161	215	0	942	1211					
Site prep after fire	200	200	1000	1000	1000	1500	1500	1500	1500	1500	1500	
Coefficients of variation	15	15	15	15	15	15	15	15	15	15	15	
Correlation coefficients of Establishment Costs Diag 1.0												
Years after planting	Yr0	Yr1	Yr2	Yr3	Yr4	Yr5	Yr6	Yr7	Yr8	Yr9	Yr10	
Yr0	1	0.8	0.5	0.2								
Yr1		1	0.7	0.3								
Yr2			1	0.7								
Yr3				1								
Yr4					1							
Yr5						1						
Yr6							1					
Yr7								1				
Yr8									1			

A	B	C	D	E	F	G	H	I	J	K	L	M
ENTER DATA IN COLOURED CELLS:												
	Start age	Rotation age	No of Harvests-Max 6				No of Log Products-Max 6					
	0	35	3				3					
Probabilities of Disasters Max 1.0												
By Yrs after planting	0	1	2	3	4	5	6	7	8	9	10	11
Probabilities of Ignitk	0	0	0	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
Severity class	0	0	0	4	4	4	4	4	4	4	4	4
Salvageable Proportions of Log Volumes after Disaster (%)												
Severity class 1	By Prod# & Harv#		Prod1	Prod2	Prod3	Prod4	Prod5	Prod6				
	Harv1	Harv2	Harv3	Harv4	Harv5	Harv6						
	Harv1	0	0	0			Severity class 1= Low severity					
	Harv2	0	60	60			Severity class 4= Very High severity					
	Harv3	0	100	100								
	Harv4											
	Harv5											
	Harv6											
Severity class 2	By Prod# & Harv#		Prod1	Prod2	Prod3	Prod4	Prod5	Prod6				
	Harv1	Harv2	Harv3	Harv4	Harv5	Harv6						
	Harv1	0	0	0								
	Harv2	0	40	40								
	Harv3	0	70	70								
	Harv4											
	Harv5											
	Harv6											
Severity class 3	By Prod# & Harv#		Prod1	Prod2	Prod3	Prod4	Prod5	Prod6				
	Harv1	Harv2	Harv3	Harv4	Harv5	Harv6						
	Harv1	0	0	0								
	Harv2	0	20	20								
	Harv3	0	40	40								
	Harv4											
	Harv5											
	Harv6											
Severity class 4	By Prod# & Harv#		Prod1	Prod2	Prod3	Prod4	Prod5	Prod6				
	Harv1	Harv2	Harv3	Harv4	Harv5	Harv6						
	Harv1	0	0	0								
	Harv2	0	0	0								
	Harv3	0	0	0								
	Harv4											
	Harv5											
	Harv6											

Fig. 2 Sample Excel spreadsheet for input of cost and disaster data

extent. No spatial data are involved. Any single stand-level analysis implicitly assumes that the inputs, outputs and risks for a particular stand are independent of its neighbours. This is not always true, as discussed later.

Beta distributions are estimated for prices and quantities of inputs and outputs based on the user-supplied mean and coefficient of variation for each variable,

together with the minimum and maximum values involved (i.e. the bounds) in the case of prices. There is a wide range of log products each with different prices and quantities so these are often simplified to just pulpwood and sawlogs. However, the volumes of each of these log types are often correlated so the Monte Carlo simulation based on beta distributions cannot be used

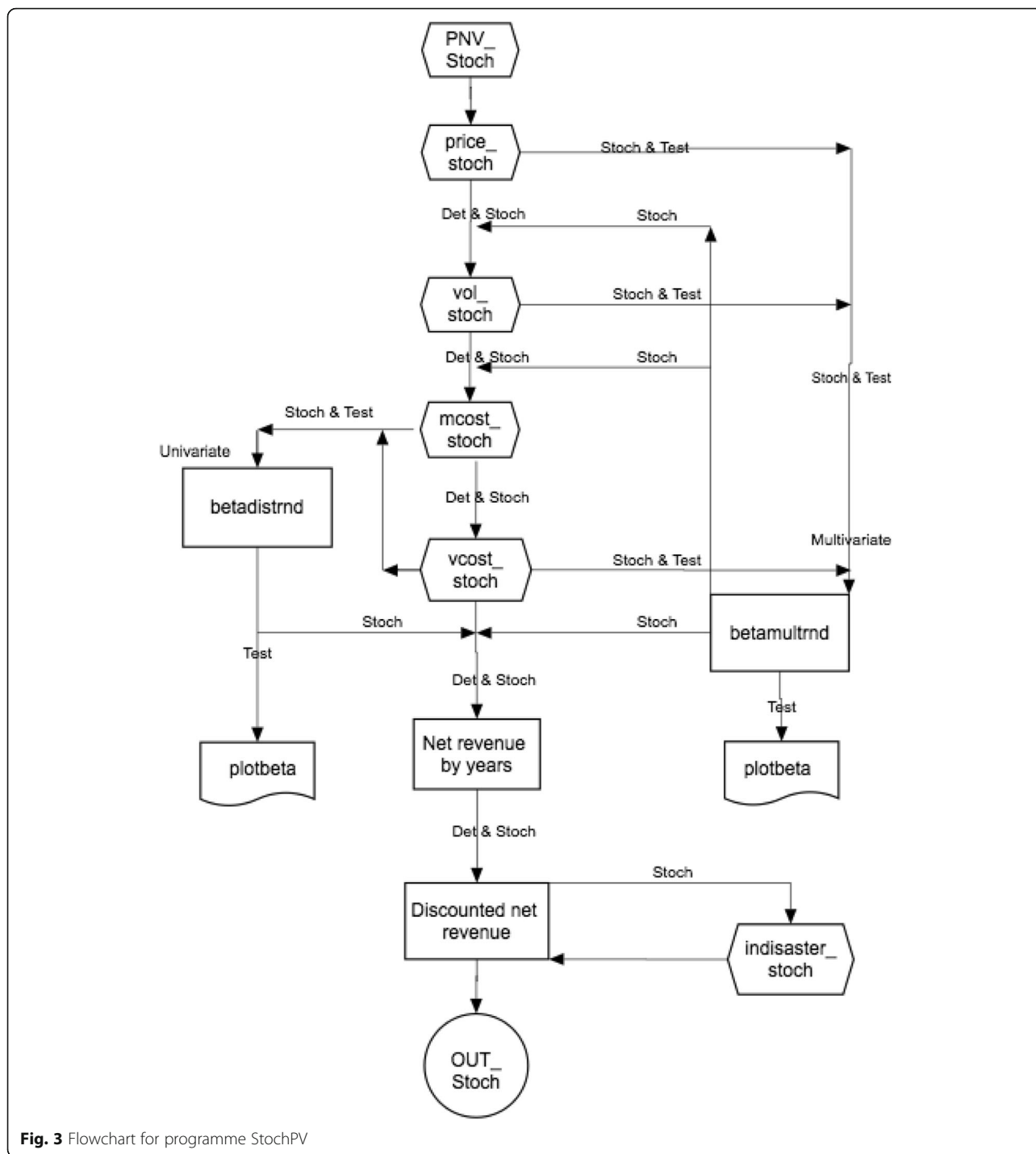


Fig. 3 Flowchart for programme StochPV

independently. Also, the correlation between them needs to be reflected, and this is achieved using copulas (described later) to preserve the multivariate relationships. Similar provisions apply to prices and to the costs of establishment and early tending.

The example data in this study are intended to be illustrative, because of the difficulties posed by commercial confidentiality in reporting actual data. They are

nevertheless realistic for radiata pine plantations in southeastern Australia. They are shown on the example Excel spreadsheet (see “Availability of data and materials” section and Figs. 1 and 2)

Log price and operational cost data reported in Ferguson (2005) were updated to real September 2014 values using the Australian Consumer Price Index. Future real stumpage prices were assumed to be stationary

in this study. However, any forecasts involving non-stationary secular trends could easily be accommodated in the data used in the software programme.

Mean harvest volumes assumed an average mean annual increment of 20 m³ha⁻¹year⁻¹. This value is higher than the average of 17.2 m³ha⁻¹year⁻¹ reported for this region by Ferguson (2014). The stochastic results, however, are very close to that value, after allowing for the impact of fires. The distributions of the log volumes by log product again follow that used in Ferguson (2005).

Some changes were made to coefficients of variation for prices and costs used in Ferguson (2005) in the light of subsequent collection of confidential company data.

Data collected for annual valuations of the estate provided much of the data. Previous estimates proved to be too low, and higher values have been substituted here. However, earlier estimates of the cross-correlations proved to be accurate and remain unchanged.

Estimates of the probability of fire were also increased markedly from those in Ferguson (2005). That study focused on small isolated woodlots in irrigated farmlands, for which the probability of fire was low. The revised data now reflect typical relative frequencies of fire occurrence in the industrial plantation estates in south-eastern Australia and were derived from analysis of confidential company data by region. A uniform distribution was used to simulate the probability of fire occurrence in stand-level analysis. Three levels of fire severity were used, based on age and post-thinning fuel-loads, to model the variation in salvageable volume experienced in past fires (see Fig. 2).

Beta distributions

Equation 1 defines the beta probability distribution. Two parameters (*p* and *q*) determine the shape of the distribution. Two determine ‘location’ *a*, the lower bound, and ‘scale’

(*b* - *a*), where *b* is the upper bound. The so-called standard beta distribution is where *a* = 0 and *b* = 1.

Lee (2004, p. 74) provided a range of graphs illustrating various shapes that the standard beta distribution can take. The probability density of the standard beta distribution (Lee 2004) is given by

$$f(x) = \frac{x^{p-1}(1-x)^{q-1}}{B(p, q)} \tag{1}$$

where 0 ≤ *x* ≤ 1; *p*, *q* > 0 and $B(p, q) = \int_0^1 t^{p-1} (1-t)^{q-1} dt$

Using the method of moments (NIST-SEMATECH 2012), the parameters of a standard beta distribution can be estimated from the estimated mean and standard deviation as shown in Eqs. 2 and 3:

$$p = \mu_x \left[\left(\frac{\mu_x(1-\mu_x)}{\sigma_x^2} \right) - 1 \right] \tag{2}$$

$$q = (1-\mu_x) \left[\left(\frac{\mu_x(1-\mu_x)}{\sigma_x^2} \right) - 1 \right] \tag{3}$$

Estimates of the *actual* mean and population variance (Eqs. 4 and 5) are obtained by using the earlier expressions for the location and scale to transform the standard mean and variance.

$$\mu_x^* = (\mu_x - a) / (b - a) \tag{4}$$

$$\sigma_{x^*} = \sigma_x^2 / (a - b) \tag{5}$$

If substantial amounts of data are available, maximum likelihood can be used to estimate the parameters of the beta distribution but the method of moments works well, subject to testing and adjustment.

The estimation of the parameters is sensitive to the combination of values of the input variables and can result in infeasible solutions. For example, *p* = *q* = 1 constitutes a uniform distribution over the [0,1] bounds. If both *p* and *q* < 1, the distribution is U-shaped, which is unacceptable for these data. The distribution is unimodal only if both *p* and *q* > 1.

If the estimated standard deviation derived from the mean and coefficient of variation lies between 0 and 1, the estimated value of the parameters, *p* and *q*, will be negative, which is infeasible. Hence, the programme tests whether the estimated standard deviation lies between 0 and 1, characterising an unacceptable bimodal distribution. If so, the coefficient of variation is increased to avoid this infeasibility. If this does not overcome the problem, generally because the mean value is small, both beta parameters of the cell are set to one, thus specifying a uniform distribution. This distribution is then used to generate the pseudo-random numbers within the prescribed bounds. Any such changes are flagged in the programme output to enable checking of transcription errors in the data.⁴

Equations 6, 7 and 8 (Bernardo and Smith 1994) give the expected value, variance and mode of the standard beta probability distribution (*π*):

$$E(\pi) = p / (p + q) \tag{6}$$

$$\text{Var}(\pi) = pq / [(p + q)^2 (p + q + 1)] \tag{7}$$

$$\text{Mode}(\pi) = p - 1 / (p + q - 2) \tag{8}$$

These are useful in checking estimates, as discussed later. The values of *p* and *q* determine the shape of the standard beta distribution. For the mode to be in the middle of the range of *θ*, *p* and *q* must be equal. Values of *q* much smaller or larger than *p* (provided both >1)

skew the distribution to the right or left. If the lower and upper bounds of the actual distribution are $[a, b]$, respectively, rather than the $[0,1]$ values of the standard distribution, the scale value $(b - a)$ stretches (or contracts) the distribution and the location value (a) shifts the starting point of the distribution. The transformed value of the random variable, y , is given by $y = a + \theta(b - a)$.

Beta probability density functions

As noted earlier, the user can choose to make a test run to check the values and the nature of the estimated distributions of input variables, other than fire occurrence. In the figures that follow, M denotes the mean value of the distribution and CV the coefficient of variation.

In Fig. 4, the upper level shows the graphs of distributions for log prices and lower level shows those for log volumes, based on the example data worksheet. The values of coefficients of variation for pulpwood prices in Fig. 4 were deliberately chosen to illustrate the sensitivity of the distributions to small changes when means are relatively small. Warning messages flag an increase in the coefficient of variation to avoid infeasibility for the top left case and changes to uniform distributions in the middle and lower left cases.

Figure 5 shows the probability density functions for the first 3 years of variable costs (upper level) and the first 3 years of maintenance costs (lower level).

Figures 4 and 5 display a variety of unimodal shapes. Some price data are skew positive reflecting the general observation that prices seldom rise much above the mean and more often fall below. For variable costs, the reverse holds, and a negative skew is common. For volumes and maintenance costs, the distributions tend to be symmetric but bounded, not with the asymptotic tails that characterise normal distributions.

These graphs of the estimated beta distributions are intended for testing to identify gross errors and shapes that are inconsistent with practitioner knowledge, enabling correction before final analyses are run. They are not intended for reproduction.

Using normal or log-normal distributions instead of beta distributions for stand-level analysis is inappropriate because the asymptotic tails of these distributions inevitably involve the generation of pseudo-random values that are well beyond the bounds of experience, raising issues of credibility. The impact of those extreme values can be muted by simulating large numbers of observations so that the mean, standard deviation, etc. of the final probability distribution are little affected. Alternatively, large values may be avoided by truncating the summary distributions, but that violates the law of probabilities that must sum to unity. Neither method satisfactorily addresses the credibility issues and hence the preference for beta distributions.

Multivariate distributions

Copulas are functions developed by Sklar (1959, 1996) that describe dependencies among random variables by linking a joint multidimensional distribution to its one-dimensional margins. A bivariate normal distribution is an example of a copula, but the concept is not restricted to marginal distributions that are normal.

Let $F(x_1, x_2)$ be a two-dimensional joint distribution with continuous marginal distributions $F_1(x_1)$, $F_2(x_2)$. It has a unique copula representation, known as Sklar's theorem:

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)) \quad (9)$$

For any $(u_1, u_2) \in [0, 1]^2$, the copula representation can alternatively be represented as

$$F(F_1^{-1}(u_1), F_2^{-1}(u_2)) = C(u_1, u_2) \quad (10)$$

If one can generate a joint uniform distribution that exhibits the same association between variables as that of the multivariate distribution sought, then the marginal distributions can be transformed to those with the desired properties (e.g. beta) using the inverse distribution function. The transformed joint beta distributions will retain the multivariate association specified. The inverse functions used need not be the same for each of the variables of interest in the copula, unlike the case above. For example, a bivariate distribution might use beta and gamma inverses. Nor is the theorem and process restricted to two dimensions.

The linear correlation coefficient, ρ , might seem a logical choice to estimate and then use to simulate the copula association for each case and iteration. But the choice of measure of association is not that simple because these transformations are not invariant to a linear correlation coefficient. However, they are invariant to Kendall's rank correlation τ . Fortunately, an exact relationship exists between the two, enabling estimation via the rank correlation coefficient and simulation via the transformed linear correlation coefficient:

$$\tau = 2 \sin^{-1}(\rho) / \pi \quad (11)$$

Instead of simulating the association via the linear correlation coefficient, others may be used, such as a multivariate 't' distribution or Gumbel's multivariate extreme value distribution. Kendall's rank correlation remains appropriate for the initial measure of association, but the relationship in Eq. 11 changes (Perkins and Lane 2003). These alternative distributions have obvious applications in simulation studies using small samples or extreme values but are seldom appropriate in the context of this study.

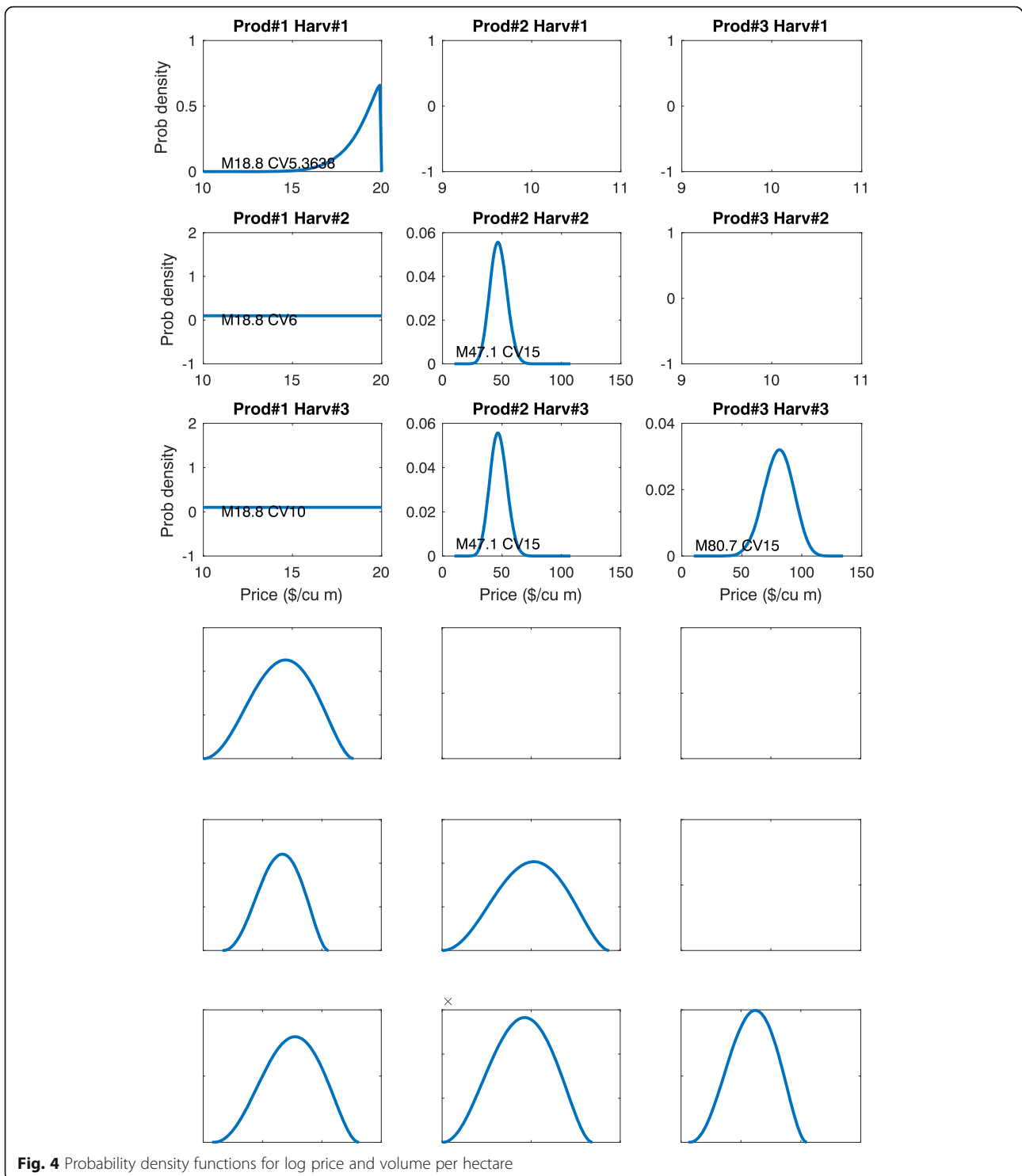


Fig. 4 Probability density functions for log price and volume per hectare

Monte Carlo simulation

Once a satisfactory set of parameter values for the distributions is obtained, a Monte Carlo approach is used to simulate pseudo-random values of the variable concerned.

For maintenance costs and those variable costs beyond an establishment phase nominated by the user, a uniform

random number generator simulates pseudo-random values using the relevant mean and bounds. Bounds for these variables are automatically set to zero and mean plus three times the standard deviation, respectively.

The inter-relationships within the respective price and volume data are estimated using the correlation

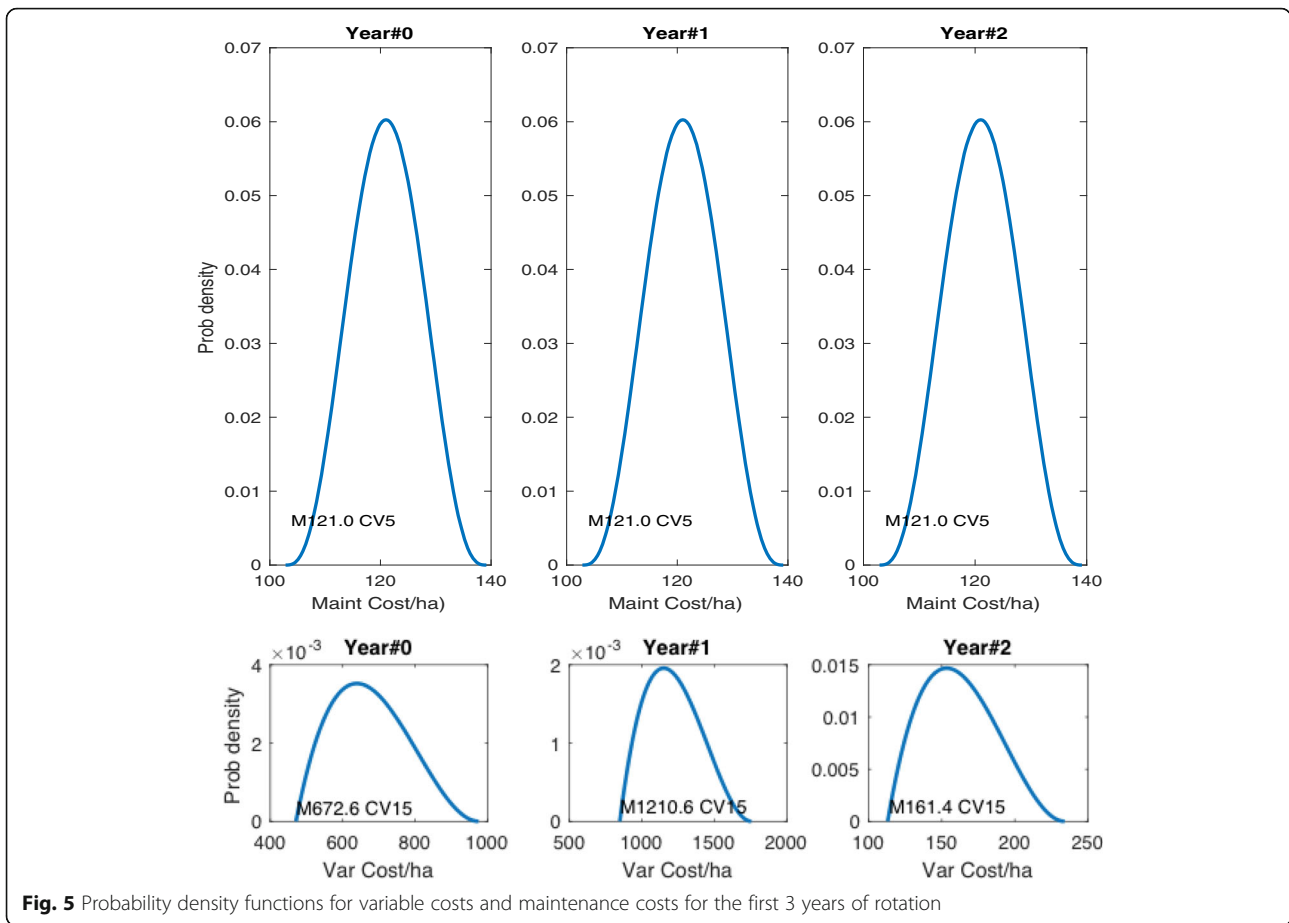


Fig. 5 Probability density functions for variable costs and maintenance costs for the first 3 years of rotation

coefficients provided in the input data. For an individual stand, prices for different log products will commonly be moderately closely correlated, if only as a result of common location relative to processing plants. Similarly, high volumes of sawlogs often reflect high volumes of pulpwood. For those variable costs in the establishment phase, cross-correlations between operations in successive years are supplied in the example data. If the terrain and ground conditions are difficult at the time of site preparation, they are likely to be difficult in later operations of planting and tending. Hence, the duration of the establishment phase, in which correlations between successive years are likely to be important, is included as an input value and, together with the cross-correlation estimate, is used to model the multivariate distribution of variable costs during that period. All correlations are empirical matters that hinge on the field data and estimates derived from that data.

For variable costs during the establishment phase and harvest yields and prices, a multivariate random number generator simulates the pseudo-random values, using a copula function to reflect the inter-relationships involved. The estimated beta distributions for each

variable provide the marginal distributions for the copula. Where a uniform form of a beta distribution occurs, multivariate correlations are still preserved because the uniform is a special case of the beta. The resulting parameter values for a standard beta distribution are then transformed to enable pseudo-random values to be simulated using the specified range centred on the specified mean. This approach ensures that large-sample means of all simulated variables are equal to the means in the input data.

Costs, prices and volumes normally have a zero minimum. However, the user can set the minimum bounds for prices at a higher or lower value than zero if desired. The maxima for prices are ideally based on historical data. Where these are not available, Martz and Waller (1982, p. 270) provided a formula for calculating the mean and variance of a truncated beta distribution at a chosen probability level (e.g. 99%). It involves somewhat laborious computation, however, and does not seem warranted given our experience of the rapid approach of beta distributions to the bounds, as shown in the preceding graphs. Where price maxima cannot be identified from historical data or market experience, the programme automatically sets the maximum at a value

of the mean plus three times the standard deviation, approximating the 99% probability level.

Fire is the sole natural disaster considered explicitly in the example used here. Fire occurrence is determined using an annual comparison of a simulated pseudo-random proportion (via a uniform distribution) with the user-supplied frequency of fire occurrence. The user also supplies an estimate of the proportions of the log classes that are salvageable after a fire for a user-specified fire severity class, as shown in the example data for disasters. Up to four levels of salvageability can be specified by the user according to their estimate of the severity of a fire at that stand age and fuel condition, preferably based on historical data. If a fire occurs, the input data on fire severity at that age then determines the proportions of harvest yield salvageable. Additional site preparation costs, depending on the age of the stand, are added for fire-affected stands.

Where a fire occurs in a stand that is salvageable following the fire, the standing volume at the time of the fire is estimated by adjustment of the volumes expected at clear-felling, plus any thinning volumes not harvested in the period between the fire and clear-felling, based on the proportion of years expired before clear-felling. Thus, a stand killed by fire at age 20 years, with an intended rotation length of 30 years, is estimated to have two thirds of the accumulated volume at 30 years. At young stand ages, this linear adjustment overestimates the volumes in larger size classes somewhat if the growth function follows the normal sigmoidal form.

Where a fire occurs in a stand that is too young to be salvageable, additional site preparation costs are added following a fire.

Maintenance costs include the cost of insuring for replanting after fire. The direct replanting costs over the first 5 years are summed and recouped by insurance payments at the time of clear-felling, with no allowance for interest costs.

Results

Deterministic

In deterministic analyses, Faustmann's (1849) land expectation value (LEV) is the present value of a forest over a given rotation length, repeated to infinity. The LEV refers to the amount a would-be forest grower could afford to pay and just break even at a given discount rate or, equivalently, the value of bare land for growing an even-aged forest on it. The word 'Expectation' has nothing to do with the statistical term 'expected value,' which is the mean of a sample or population. It is simply a term that resulted from the widely used translations of the German 'Bodenerwartungswerte' as land expectation value or soil expectation value.

An LEV is calculated using the Faustmann formula to estimate the present value of bare land. It involves a rotation of length t years repeated in perpetuity. If PV denotes the present value of a single rotation of length t years, using a discount rate of r and excluding any consideration of land costs,

$$LEV = PV / (1 - (1 + r)^{-t}) \quad (12)$$

StochPV also calculates net present value (NPV), the present value of a single rotation, including land costs at the start and land revenues at the end, an alternative statistic that is also often of use:

$$NPV = PV - L + L(1 + r)^{-t} \quad (13)$$

Stochastic

The economic criteria for stochastic analyses are analogous to those for deterministic but now account for risk and uncertainty based on the probability distributions and prices of the inputs and outputs involved.

The Faustmann formula for deterministic evaluation of LEV is based on an infinite series of fixed rotations. It is no longer appropriate for stochastic analyses because the rotations may vary in length for each iteration of the Monte Carlo simulation due to natural disasters. The stochastic analogue is the 'expected value of the present value of bare land' (EVL) over an infinite series of stochastically determined rotations.

Amacher et al. (2009) provide an elegant theoretical analysis of the risk of catastrophic events such as fire, based on a homogeneous Poisson process. In principle, their model is a stochastic analogue of the deterministic Faustmann model, based on an infinite series of rotations of stochastically determined rotation lengths, rather than fixed length.

Amacher et al. (2009, p273) derive an equation for calculating the expected value of the present value of bare land (EVL). While correct, given the underlying assumptions, their equation is inappropriate for calculating EVL for even-aged plantations of radiata pine or similar forests because of three assumptions that do not hold:

- (1) The probability of fire is not constant over the life of a radiata pine plantation.
- (2) Timber is salvageable from radiata pine plantations of 15 years of age or more following fires.
- (3) Adding a risk premium to the discount rate to recognise the risks posed by fires is inappropriate because StochPV already accounts for the probability of fire and fire damage directly.

Amacher et al.'s addition of a risk premium may be useful as a theoretical device to explore the impact of

fire on a deterministic analysis. However, their model ignores all the other systematic risks. The addition of a risk premium to the discount rate double-counts the relevant risks in StochPV and related models that incorporate all systematic risks. A risk premium should not be used (see Ferguson 2015).

The stochastic simulation model described in this paper explicitly incorporates varying probabilities of fire occurrence and severities of fire damage, according to the age and condition of the plantation. The revenues from salvageable timber vary accordingly.

Amacher et al. (2009) derived and used the expected value of the length of successive rotation infinity to calculate EVL. This approach is impractical in the empirical model described here as it would imply an infinite simulation of stochastically determined rotations. Instead, we simulate a finite series of stochastically determined rotations, terminating the simulation once the marginal present value of the next rotation is negligible.

The software is presently set to terminate once the last rotation results in an accumulated period across all past rotations above 100 years. Using this limit, the marginal present value of an additional \$1.00 of discounted net revenue from the last rotation is less than \$0.003.

Operationally, two stages of calculation are involved in simulation:

- (1) For any single run of the simulation, simulations of successive rotations are repeated, accumulating the appropriately discounted net revenues. They continue until the accumulated time of successive rotations reaches a point at which the *marginal* present value is negligible.
- (2) Stage 1 is repeated for each of the iterations (e.g. 20,000), and the resulting series of EVL values are used to calculate summary statistics (mean and standard deviation, etc.) and probability distributions.

For estimates of EVL, the number of observations for EVL statistics and distributions will approximately equal the number of iterations (e.g. 20,000) divided by the expected value of the length of rotations. In this study, the expected value of the length of rotations is 25.86 years for a target rotation of 29 years, including 1 year of fallow.

The other economic criterion for stochastic analyses is the direct stochastic analogue of NPV; the expected value of the present value of a single rotation, EPV, includes the initial cost of bare land and the revenue from the sale thereof at the end of a single rotation.

User options

The number of iterations is set by the user. Two alternative discount rates and two sets of initial and final (i.e. at

clear-felling) land prices can be chosen or user values substituted.

The Excel worksheets must conform to a standard Excel template, as used on the example spreadsheet (see Figs. 1 and 2 and “Availability of data and materials” section). The nominal length of rotation can be extended beyond that used here for fast-growing plantations by extending the relevant sections of all worksheets to the right.

The number of log products is nominally limited to six but can be expanded by extension of the relevant cells of worksheets. However, extension requires more data for correlations, etc. and more computing time. Pooling of log products with similar means and coefficients of variation may, therefore, be preferable.

Successive harvests can number up to six, but more than four (including clear-felling) is unusual.

Outputs

In deterministic mode, the analysis ignores disasters and other risks and simply uses the means of prices, costs and yields in the conventional manner for calculating LEV and NPV. Deterministic output is shown in Fig. 6.

In stochastic mode, the final set of iterations is summarised to provide mean EVL and EPV values and their probability (relative frequency) distributions. In stochastic analyses, means, standard deviations, maxima and minima of EVL and EPV are listed for each of the two discount rates and two land prices specified. They are also written to a file whose name based on that of the input data, and EVL and EPV components are summarised in histograms as the probability mass functions, as shown in the upper level of Fig. 7. In stochastic mode, fire occurrence can have a marked impact on harvest yields and thus on the average mean annual increment. Mean annual increment and its standard deviation of the mean annual increment are therefore calculated and reported for all stochastic runs, as well as the number of fire disasters and the mean age at which the disasters occurred. The written output (upper level, Fig. 7) also flags chosen options together with any changes made automatically in estimating beta distributions and a numerical summary of output.

A stochastic run of 20,000 iterations for four harvests and six log products takes about 30 s using a 2.9 GHz Intel Core i9 iMac computer.

Testing

The deterministic results for the example spreadsheet (see Fig. 6) in this study were checked against independent calculations done directly on Excel® spreadsheets. The output of a deterministic analysis also facilitates checking of all input values to ensure transcription onto the spreadsheet is correct.

1. Italicized text shows user response to Program requests for input data
2. Text between ‘Normal end of input’ and ‘StochPVPaper.xlsx’ is for quick checking of input data.
3. Text after ‘StochPVPaper.xlsx’ is summary of run specifications and the results that are written to and saved in the Excel spreadsheet of that name, the ‘Normal end etc’ verifying that no problems arose in the computations.

```
>> StochPV
Enter Excel file name with suffix containing data StochPVPaper.xlsx
Accept default alternative discount rates [6.00,8.00] - Y/N y
Enter no. of iterations (zero for determ,1 for pdfs, number for stoch)) 0
Accept default Land prices (Start age $1500/ha,Clearfell $1500/ha) Y/N y
Normal end of input
Log prices for successive harvest ops
Deterministic values for its==0
Log volumes for successive harvest ops
Deterministic values for its==0
Values (stochastic if its==1) of first 6 maintenance costs
121.00 121.00 121.00 121.00 0.00 121.00
Values (deterministic for its ==0)of first 6 variable costs
672.57 1210.62 161.42 215.22 0.00 941.59
StochPVPaper.xlsx
its 0 Mean mai 17.1429 STDmai 0 Disasters 0 Mean disaster age 0
Normal end - proceed to OUT_Stoch to obtain results
Disc Deterministic: Present value before tax ($/ha)
rate Mean SD Min Max Obs
6.00 LEV -173 0 0 0 0
8.00 LEV -1606 0 0 0 0
6.00 N1R -1467 0 0 0 0
8.00 N1R -2911 0 0 0 0
LEV: Deterministic Faustmann Land Expectation Value - perpetual rotations
N1R: Deterministic NPV one rotation incl Land Prices at start & clearfell ages
```

Fig. 6 Example of user input and responses: deterministic run

In testing mode, a single set of simulated values is calculated for all input variables, together with graphs (see Figs. 1 and 2) of the individual probability density distributions to facilitate checking. The printed output for this option is very similar to that in Fig. 6 and is not reproduced here.

In stochastic mode, the frequencies of fire occurrence can be set to zero to enable the impact of fire to be evaluated or to check the stochastic processing of the inputs and outputs, other than fire. Comparison of the mean annual increments from such a stochastic run with that of the corresponding deterministic run provided a check that the stochastic version is calculating correctly.

Conclusions

Although the programme had its origins in the stochastic analysis of radiata pine plantations, it is widely applicable to other species of even-aged plantations and forests. For tropical plantations, for example, the disasters might be framed in terms of the frequency of occurrence of cyclones, rather than fires. In less fire-prone temperate regions, windblow or snow or hail damage may be the relevant natural disaster. The programme is capable of analysing any even-aged forest for which appropriate models are available to predict future log volumes by grades and their prices.

The programme integrates the systematic risks relating to stand management in an analytical framework that is

1. Italicized text is user response to Program requests for input data
2. Text between 'Normal end of input' and 'StochPVPaper.xlsx' shows changes to specification of input data made by program because Price data was infeasible.
3. Text after 'StochPVPaper.xlsx' is summary of run specifications and numerical results written to and saved in the Excel spreadsheet of that name, the 'Normal end etc' verifying that no further problems arose in the computations, and the graphical results.

Enter Excel file name with suffix containing data *StochPVPaper.xlsx*
 Accept default alternative discount rates [6.00,8.00] - Y/N y
 Enter no. of iterations (zero for determ,1 for pdfs, number for stoch)) 10000
 Accept default Land prices (Start age \$1500/ha,Clearfell \$1500/ha) Y/N y
 Normal end of input
 PRICE cv(1,1) < 1, increased to :5.3638
 PRICE p or q(2,1) < 1 or Inf, switch to uniform :
 PRICE p or q(3,1) < 1 or Inf, switch to uniform :
 StochPVPaper.xlsx
 its 10000 Mean mai 17.1648 STDmai 3.0713 Disasters 892 Mean disaster age 18.8722
 Normal end - proceed to OUT_Stoch to obtain results

Disc rate	Stochastic:	Mean	SD	Min	Max	Obs
6.00	EVL	1385	1102	-3272	5252	3134
8.00	EVL	-253	679	-2982	2128	3134
6.00	E1R	-186	1456	-4813	4703	3134
8.00	E1R	-1688	934	-4695	2510	3134

EVL: Stochastic Expected Value of Land - approx perpetual rotations
 E1R: Stochastic Expected Value of NPV of one rotation incl Land Prices at start & clearfell ages

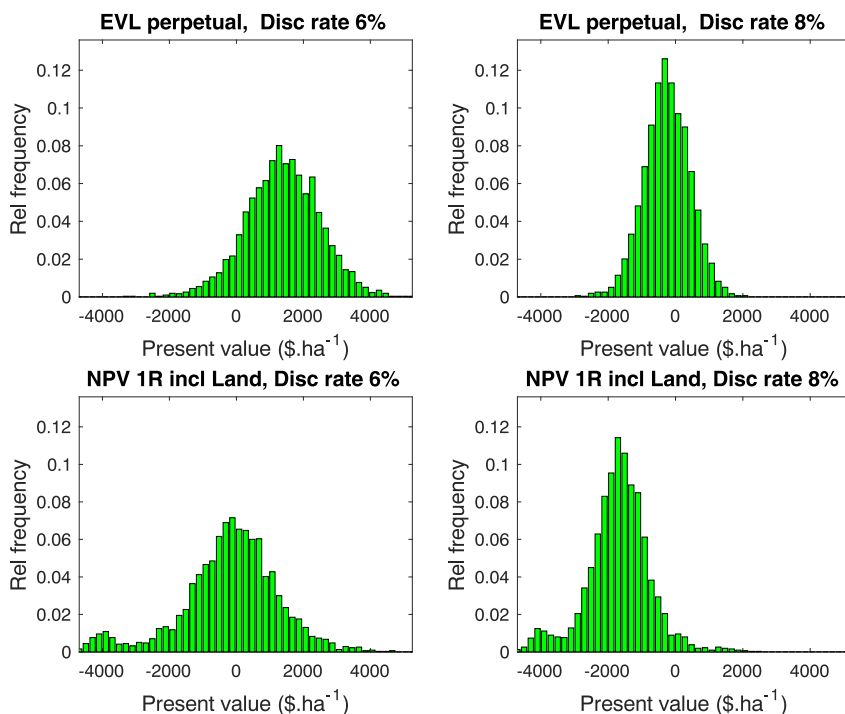


Fig. 7 Example of user input and responses: stochastic run

relatively simple and communicable (Hildebrandt and Knoke 2011) using spreadsheet data and graphical as well as numerical outputs. Unlike almost all past models of this type, the inter-relationships between kindred variables are taken into account using multivariate distributions based on copulas and the marginal distributions involved. The use of beta distributions provides flexibility in representing the nature of the empirical distributions of the quantities and prices of inputs and outputs.

The illustrative results highlight the importance of disasters. In the unlikely event that the probability of fire occurrence could be reduced to zero, the expected value of bare land in the illustrative data used in this study would rise by \$300 to \$500 ha⁻¹. The results of an estate-wide stochastic analysis (Ferguson 2015) yielded even larger margins. These larger margins underscore a limitation of stand-level analyses that implicitly assume inputs, outputs and risks for a particular stand are independent of their neighbours. Fire is a contagious phenomenon, especially under extreme weather conditions, so the assumption of independence does not always hold. Additional safety margins, therefore, need to be applied to the results of stand-level analyses of this type. Or they need to be checked against stochastic analyses of the entire estate, similar to those in Ferguson (2015).

There may be some situations where obtaining data on the quantities and prices of inputs and outputs, and for fire, could be too costly or difficult to collect. For the owners of medium and large plantations or forests, most of the raw data are already collected in annual valuations of the estate and financial management. Collecting these data might be too costly or difficult for the owners of small plantations or forests, but this could be overcome with assistance through collectives or from public research bodies.

The example data and outcomes of analyses using it are only illustrative of those for radiata pine plantations. In the terminology of Hildebrandt and Knoke (2011), the StochPV results are suitable for evaluation of first-degree stochastic dominance (see also Caulfield 1988 and Lien et al. 2007). The results could also be applied to second-degree stochastic dominance if the decision-maker had a material aversion to risk, all the data needed being available as outputs of the programme. However, assessment of risk aversion in a practical setting is not easy, especially where potential reputational and personal financial risks are involved. My preference, therefore, is to use the risk-neutral stochastic stand-level results and, if in doubt, apply a safety margin (Damodaran 2014, p. 915).

Endnotes

¹Excel is used purely as a vehicle for input of data and output of numerical results. It plays no part in computations.

²Matlab is a commercial software package available from www.mathworks.com.

³<http://www.mathworks.com/matlabcentral/fileexchange/?utf8=✓&term=StochPV>

⁴See Fig. 7 for an example.

Acknowledgements

IF would like to thank the colleagues in the Australian and New Zealand forest industries for their assistance and insights over a number of years and to acknowledge the helpful comments of two anonymous reviewers. Special thanks go to Ron Harper (HVP Plantations Pty Ltd) for providing the data on fire salvage.

Availability of data and materials

A complete copy of the programme StochPV, together with a spreadsheet 'StochPVPaper.xlsx' showing the example data used in this paper and a brief User Manual, is available in the Mathworks' File Exchange repository (<https://www.mathworks.com/matlabcentral/fileexchange/>). These are open-access materials freely available under BSD licence.

Author's information

At the time of programming and much of the writing, IF worked for the University of Melbourne and was a Director of HVP Plantations Pty Ltd and Taumata Plantations Inc. He retired from the University in 2003 and from other positions at the end of 2015.

Competing interests

The author declares that he has no competing interests.

Received: 25 March 2016 Accepted: 13 February 2017

Published online: 28 April 2017

References

- Amacher, G. S., Ollikainen, M., & Koskella, E. (2009). *Economics of forest resources*. Cambridge, MA: The MIT Press.
- Ben-Haim, Y. (2006). *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*, 2nd ed. London: Academic Press.
- Bernardo, J. M., & Smith, A. F. M. (1994). *Bayesian theory*. Chichester: John Wiley & Sons.
- Cassidy, P. A., Rodgers, J. L., & McCarthy, W. O. (1970). A simulation approach to risk assessment in investment analysis. *Review of Marketing and Agricultural Economics*, 38, 3–24.
- Caulfield, J. P. (1988). A stochastic efficiency approach for determining the economic rotation of a forest stand. *Forest Science*, 34, 441–457.
- Damodaran, A. (2014). *Investment valuation* (3rd ed.). Chichester: John Wiley & Sons.
- Faustmann, M. (1849). On the determination of the value which forest land and immature stands pose for forestry (English translation). In M. Gane (Ed.), (1968), *Martin Faustmann and the Evolution of Discounted Cash Flow*. Oxford Forestry Institute Paper 42. Oxford, UK: Oxford University/Commonwealth Forestry Institute.
- Ferguson, I. (2005). Chapter 10, economic and policy implications. In S. Nambiar & I. Ferguson (Eds.), *New forests: wood production and environmental services*. Melbourne: CSIRO Publishing.
- Ferguson, I. (2014). Australian plantation inventory: ownership changes, availability and policy. *Australian Forestry*, 77(1), 25–38.
- Ferguson, I. (2015). Forest valuation and risk. *Australian Forestry*, 79(1), 32–42.
- Ferguson, I. (2016). Future rotations problem. *Australian Forestry*, 79(2), 122–125.
- Hildebrandt, P., & Knoke, T. (2011). Investment decisions under uncertainty—a methodological review on forest science studies. *Forest Policy and Economics*, 13(1), 1–15.
- Kirby, M. C., Sinden, J. A., & Kaine, G. W. (1993). Appraisal of agroforestry investment under uncertainty: a South Australian case study. *Australian Forestry*, 56(2), 109–119.
- Knott, J. L., Turner, J., & Lambert, M. J. (1995). Fertilization of *Pinus radiata* plantations after thinning. II. Investment potential. *Australian Forestry*, 59(1), 22–31.
- Lee, P. M. (2004). *Bayesian statistics: an introduction* (3rd ed.). London: Hodder Arnold.

- Lien, G., Størdal, S., Hardaker, J. B., & Asheim, L. J. (2007). Risk aversion and optimal forest replanting: a stochastic efficiency study. *European Journal of Operational Research*, *181*, 1584–1592.
- Lothner, D.C., Hoganson, H.M., & Rubin, P.A. (1986). Examining short-rotation hybrid poplar investments by using stochastic simulation. *Canadian Journal of Forest Research*, *16*, 1207–1213.
- Martz, H. F., & Waller, R. A. (1982). *Bayesian reliability analysis*. New York: John Wiley & Sons.
- NIST-SEMATECH (2012). *e-Handbook of Statistical Methods*, Available on <http://www.itl.nist.gov/div898/handbook/>, 1/8/2016.
- Perkins, P & Lane, T (2003). Monte Carlo simulation in Matlab using copulas. *Matlab News & Notes*, 4-5.
- Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, *8*, 229–231.
- Sklar, A. (1996). Random variables, distribution functions, and copulas—a personal look backward and forward. In L. Rüschendorf, B. Schweizer, & M. Taylor (Eds.), *Distributions with fixed marginals and related topics* (pp. 1–14). Hayward, CA: Institute of Mathematical Statistics.
- Varma, V., Ferguson, I., & Wild, I. (2000). Decision support system for the sustainable forest management. *Forest Ecology and Management*, *128*(1-2), 49–55.
- Wald, A. (1945). Statistical decision functions which minimize the maximum risk. *The Annals of Mathematics*, *46*(2), 265–280.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
