United States Department of Agriculture

Agroforestry Land-use Economic Yield and Risk (ALLEY) Model 2.0:

A Computer Suite to Simulate and Compare Stochastic Yield and Returns of Alley Crop, Monocrop, and Pine Plantation Systems in the U.S. South

Gregory E. Frey, Michael A. Cary, Barry K. Goodwin, and D. Evan Mercer

Forest Service
Southern Research Station
e-General Technical Report SRS-235
November 2018

Authors

Gregory E. Frey: Research Forester; U.S. Department of Agriculture Forest Service, Southern Research Station, Forest Science and Assessment Center; 3041 E. Cornwallis Rd., Research Triangle Park, NC 27709.

Michael A. Cary: PhD Student; West Virginia University, Eberly College of Arts and Sciences; Morgantown, WV 26506.

Barry K. Goodwin: Professor; North Carolina State University, Department of Agricultural and Resource Economics; Raleigh, NC 27695.

D. Evan Mercer: Research Economist (retired); U.S. Department of Agriculture Forest Service, Southern Research Station; Research Triangle Park, NC 27709.

Acknowledgments

Initial work on this model was funded by Joint Venture Agreement 12-JV-11330143-104 between Virginia State University and U.S. Department of Agriculture Forest Service, Southern Research Station.

The authors thank **Drs. Carola Paul** and **Hernan Tejeda** for their valuable comments on an early draft of this manuscript.

Abstract

Alley crop systems, which combine field crops in wide alleys between rows of trees, have been presented as a sustainable land use that can generate economic benefits and ecosystem services. We created models of stochastic processes that simulate yields, prices, and costs for an alley crop system and two competing land uses: monocrop and pine plantation. The suite of models, called Agroforestry Land-use Economic Yield and Risk (ALLEY) Model 2.0, uses a Monte Carlo approach, iterating numerous times over a given time horizon, to estimate and compare expected values and distributions of potential financial returns from the three land uses with selected management variables. ALLEY Model 2.0 includes sample historical data on crop yields in Halifax County, NC, crop and timber prices in North Carolina, and costs from the U.S. Southeast. Timber growth and yield were simulated using biometric equations from past research on loblolly pine in the U.S. Southeast. Other parameters with little historical data, such as competition between trees and annual row crops in alleys, were based on literature review and authors' estimates. ALLEY Model 2.0 may be used for research on agroforestry decision making, testing of crop characteristics that improve returns in agroforestry settings, evaluation of current policies and future policy options, or assessment of potential changes in returns due to the effects of a changing climate. An example application is provided. The ALLEY Model 2.0 software suite and sample data file can be accessed at https://www.srs.fs.usda.gov/pubs/gtr/gtr_srs235/.

Keywords: Agroforestry, Gaussian copula, Monte Carlo method, risk and uncertainty, stochastic modeling.

Cover Photograph
Alley crop system by Reza Rafie of Virginia State University

Product Disclaimer
The use of trade or firm names in this publication is for reader information and does not imply endorsement by the U.S. Department of Agriculture of any product or service.

November 2018

Forest Service
Research & Development
Southern Research Station
e-General Technical Report SRS-235

Southern Research Station
200 W.T. Weaver Blvd.
Asheville, NC 28804
www.srs.fs.usda.gov

Agroforestry Land-use Economic Yield and Risk (ALLEY) Model 2.0:

A Computer Suite to Simulate and Compare Stochastic Yield and Returns of Alley Crop, Monocrop, and Pine Plantation Systems in the U.S. South

Gregory E. Frey

Michael A. Cary

Barry K. Goodwin

D. Evan Mercer

Contents

INTRODUCTION 1
Objectives, Principal Outputs, and Target Users 2

MODEL OVERVIEW 3
Model Input Parameters 5
Model Outputs 6
Model Options 6

OBTAINING AND UTILIZING THE MODEL 9
Data Sources 10
Data Format 10

DETAILED MODEL CONSTRUCTION 12
Parameter Estimation for Returns Elements 12
Input of Other Parameters 16
Monte Carlo Simulation 16
Calculation and Output of the Financial Returns Indicators and Management Variables 25
Detailed Model Options 25

EXAMPLE APPLICATION 29
Data and Method 29
Results 29

AREAS FOR FUTURE IMPROVEMENTS TO THE MODEL 32

CONCLUSION 33

LITERATURE CITED 34

APPENDIX: ADDITIONAL PARAMETER LISTS 35

Introduction

Agroforestry systems have been presented as potentially sustainable land uses that can generate economic benefits and ecosystem services. However, they have not been widely adopted in the U.S. Southeast or mid-Atlantic States, possibly due to a lack of information or the perception that they are not financially competitive with alternative land uses (Workman and others 2003).

Alley crop systems are agroforestry systems in which field crops are cultivated between rows of trees. A noteworthy economic advantage of alley crop systems is that trees mitigate risk, providing an opportunity to use timber sales to offset losses from agricultural products. Alley crop systems also provide opportunities to enhance annual returns by changing the crops (including niche products) that are produced from year to year depending on market conditions and tree maturity. This versatility may be crucial to achieving significant long-run adoption and diffusion of the system.

The Agroforestry Land-use Economic Yield and Risk (ALLEY) Model 2.0 is a suite of computer functions and scripts, in the MATLAB[1] computer programming language, for simulating financial returns from an alley crop system and two competing land uses: monocrop system and pine plantation. ALLEY Model 2.0 provides a method to generate information about financial aspects of alley cropping specifically, and agroforestry-based land management options in general, via simulations based on trends extracted from historical data. Researchers and extension professionals can use the model to answer questions that will help landowners make better-informed decisions. An early version of this model was described in Cary and others (2014);[2] the underlying models have been updated and improved, including new parameters and data.

ALLEY Model 2.0 simulates stochastic yields, prices, and costs for the three land uses. The simulation repeats the stochastic process a pre-defined number of times to generate a distribution of financial returns. It allows for comparisons among long-run distributions of economic returns of the land uses, e.g., testing for stochastic dominance or other features that might indicate financial preference for long-run risk-neutral or risk-averse individuals. ALLEY Model 2.0 also allows the user to model and observe the effects of changes to various factors that create uncertainty, such as changes in policy, incorporation of new alternative crops with different risk profiles, increased probability of catastrophic events (such as weather events that create a catastrophic reduction in crop yields), and different assumptions about the level of competition between agroforestry system components.

[1] MATLAB is a product of the Mathworks® Company. See https://www.mathworks.com/products/matlab/.

[2] While the Cary and others (2014) text did not use the "ALLEY" naming convention, we consider the model described there to be the 1.0 version of this suite. Hence, the version described here is called "ALLEY Model 2.0."

Objectives, Principal Outputs, and Target Users

The objective of ALLEY Model 2.0 is to simulate long-run financial returns and risk of alternative land uses—alley crop, monocrop, and pine plantation—that offer differing forms of versatility in decision making, using techniques that are flexible enough to allow for changing various assumptions that could alter the resulting simulated returns and risk. The principal model outputs are the simulated long-run distributions of indicators of financial returns, or profitability, which are net present value (NPV), soil expectation value (SEV), and annual equivalent income (AEI). The distributions of these long-run financial and management indicators can then be graphed visually as probability or cumulative distribution functions, and statistics such as mean and standard deviation can be calculated.

ALLEY Model 2.0 is intended for use by research and extension organizations interested in exploring the conditions under which alley crop systems may or may not be feasible and beneficial for various types of client landowners in their region of interest. Users of the model will need at least a basic understanding of computer modeling and statistics. ALLEY Model 2.0 is not intended as a decision-support system for individual landowners.

Model Overview

ALLEY Model 2.0 was created and run in MATLAB version 2015b with Statistical Toolbox but should be compatible with other versions of MATLAB as well as open-source software such as GNU Octave.[3] The model uses a stochastic Gaussian copula process in a Monte Carlo framework to simulate financial returns from an alley crop system with annual crops in the alleys and loblolly pine (*Pinus taeda*) in the tree rows, a monocrop system with annual crops, and a conventional loblolly pine plantation.

The model overview is shown in figure 1.[4] The programming is designed to be flexible and modular, so users can choose to run scripts independently, depending on their needs. The model inputs various parameters related to management choices and uses those to generate stochastic "returns elements" of each annual crop or timber product. **The stochastic "returns elements" are defined as three factors that drive financial returns for crop or timber products: input cost, output price, and output yield**. The model simulates these returns elements over a pre-defined time horizon, and combines them to calculate financial returns indicators NPV, SEV, and AEI. That time horizon is then repeated a large number of times in a Monte Carlo framework to find a distribution of long-run returns for a particular land use.

Random annual shocks to the three returns elements for all crops (input cost, output price, and output yield) and two pine returns elements (timber product prices and timber input cost index) are generated through a stochastic Gaussian copula approach (Frees and Valdez 1998). Random annual shocks for all returns elements are generated from a multivariate normal distribution function, creating a joint distribution of the annual shocks, which are correlated.[5] Since returns elements may not be normally distributed (Goodwin and Ker 2002), the random normal shocks may be translated to alternative parametric distributions including beta or lognormal, as set by the user. The parameterized shock is then translated to an actual value of the returns element using one of several possible "autoregression-trend functions" that can include one of various time trends, and/or autoregression, as set by the user. Timber yield is modeled independently of timber price and timber input cost, as described in more detail below. Details of the returns elements generation process are given in the Detailed Model Construction section below under Generation of the Simulated Returns Elements for the *m*-year Time Horizon.

ALLEY Model 2.0 includes a decision rule for selecting an annual crop in each year, which is related to the expected returns of each annual crop alternative. Details of the crop selection rule are given in the Detailed Model Construction section under *Crop rotation and selection*. The model also allows the user to activate/deactivate several options, which are summarized below under Model Options and later described in the Detailed Model Construction section under Detailed Model Options.

The model is both data- and parameter-intensive. This allows users the flexibility to generate parameters from historical data and/or alter assumptions about various parameters to gauge their

[3] GNU Octave is freely redistributable software written by John W. Eaton and many others. GNU Octave is mostly compatible with MATLAB. We have not tested our code on GNU Octave. See https://www.gnu.org/software/octave/about.html.

[4] Definitions of the input parameters and outputs are given below in tables 1 and 2.

[5] It is well known, for example, that crop yields and prices are negatively correlated (Goodwin and Ker 2002), because of supply and demand effects.

effect. The model can be run either by (1) importing historical data and using these data to estimate parameters regarding the returns elements, and directly entering parameters regarding particular management choices; or (2) directly entering assumed known parameters. The model can accept traditional row crops with publicly available historical data series such as corn, soybeans, and cotton, or the user can incorporate other annual crops as desired by either providing the relevant historical data or directly entering the relevant parameters. Loblolly pine was chosen because it is one of the main trees planted for timber worldwide, and thus is a commodity crop that has well-documented growth and yield equations and historical timber prices. At the same time, this choice limits the direct application of the model to areas where loblolly is a common crop tree, such as the U.S. South and, possibly, Latin America.

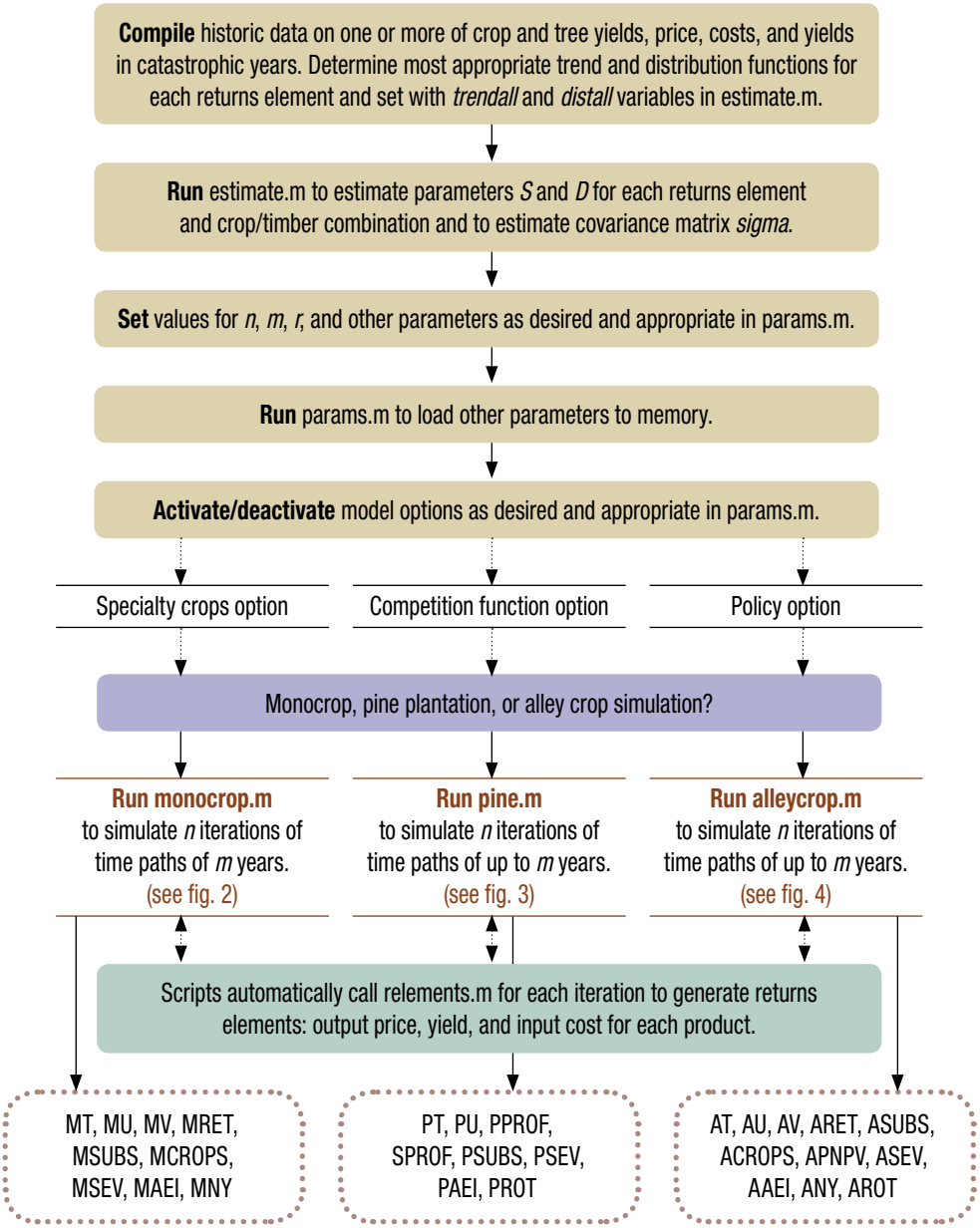

Figure 1—Overview of ALLEY Model 2.0. See table 1 for definition of inputs and table 2 for outputs.

Model Input Parameters

ALLEY Model 2.0 includes numerous parameters that define alley crop, monocrop, and pine plantation management and how financial returns are generated. The authors have provided default values for parameters based on the relevant literature, historical data, and, in some cases, the authors' informed assumptions. Very little is known about the biophysical or socioeconomic aspects of alley crop systems, such as potential crop yields or market opportunities, in the U.S. South with a high degree of certainty. Therefore, our computer code is open source and allows the user to adjust parameters as desired, to allow for different geographic regions, to utilize the user's own prior knowledge, to compare various management options, and in response to new information that may become available in the future.

Table 1 lists the key input parameters that are of main importance to the user. Other parameters are listed in the appendix. The user can easily change the default values by directly modifying the code of the scripts estimate.m and params.m.

Table 1—Key input parameters of ALLEY Model 2.0 (other parameters given in the appendix)

Parameter	Format	Description	Units	Default value	Script or function
nec	Scalar	Number of returns elements to be simulated for each crop	Unitless	3	Coded into estimate.m
nep	Scalar	Number of returns elements to be simulated for pine	Unitless	2	Coded into estimate.m
ne	Scalar	*nec* + *nep*	Unitless	5	Calculated in estimate.m
q	1 x *ne* matrix	Number of values to be simulated for each returns element (e.g., number of crops) (The first *nec* elements represent number of crops; remaining *nep* elements represent number of pine price and cost categories.)	Unitless	[3 3 3 2 1]	Coded into estimate.m
data	*T* x sum(*q*) matrix	Historical data matrix for estimating parameters *S*, *D*, and *sigma*	Variable units	No default value	Imported to estimate.m from Excel file
trendall	1 x *ne* matrix	Selection of autoregression-trend functional form for each returns element (See section on Parameter Estimation for Returns Elements under Detailed Model Construction for definition of values.)	Unitless	[6 5 2 2 0]	Coded into estimate.m
distall	1 x *ne* matrix	Selection of distribution function for shock values for each returns element (See section on Parameter Estimation for Returns Elements under Detailed Model Construction for definition of values.)	Unitless	[1 2 0 1 0]	Coded into estimate.m
perc	Scalar	Percent deviation value from minimum and maximum residuals, used for fitting distribution function	Unitless	0.95	Coded into estimate.m
S	4 x *q*(1) x *ne* matrix	Parameters for the autoregression-trend function for all returns elements for all crop/timber products	Variable units	No default value	Output of estimate.m or may be input directly to memory
D	4 x *q*(1) x *ne* matrix	Parameters of the distribution functions for the shock values for all returns elements for all crop/timber products	Variable units	No default value	Output of estimate.m or may be input directly to memory
sigma	sum(*q*) x sum(*q*) matrix	Covariance matrix of normalized shocks for all returns elements for all crop/timber products	Variable units	No default value	Output of estimate.m or may be input directly to memory
seedall	*q*(1) x *ne* matrix	Seed of historical values from most recent year of all returns elements for all crops	Variable units	No default value	Output of estimate.m or may be input directly to memory
rh	*q*(1) x 5 matrix	Past 5 years (years *t* = -5 through -1) of revenue from each crop, to be used for the Farm Bill payment scenario	$/ha	No default value	Output of estimate.m or may be input directly to memory
n	Scalar	Number of iterations of the model	Unitless	10,000	Input in params.m
m	Scalar	Time horizon of the model	Years	40	Input in params.m
r	Scalar	Discount rate	Unitless	0.05	Input in params.m

Model Outputs

The model generates distributions for three financial returns indicators. The primary indicator generated is net present value (NPV), which is the sum of discounted net returns over the given time horizon. This NPV is also transformed in capital terms to equivalent values—over an infinite time horizon (soil expectation value, SEV) and over 1 year (annual equivalent income, AEI). The simulation is iterated *n* times, using a Monte Carlo framework, generating a distribution of each of the indicators above. That distribution can be plotted on a graph, and indicators of risk such as standard deviation can be estimated. These indicators are output in the matrices MT, PT, and AT (see table 2). The financial returns indicators are defined as follows (Cubbage and others 2015):

$$rev_{crop_t,t} = price_{crop_t,t} \cdot yield_{crop_t,t} \tag{1}$$

$$ret_{crop_t,t} = rev_{crop_t,t} - cost_{crop_t,t} \tag{2}$$

$$NPV = \sum_{t=1}^{m} \frac{\left(ret_{crop_t,t} + subs_{crop_t,t}\right) + \left(ret_{pine,t} + subs_{pine,t}\right)}{(1+r)^t} \tag{3}$$

$$SEV = NPV + \frac{NPV}{(1+r)^m - 1} \tag{4}$$

$$AEI = SEV \cdot r \tag{5}$$

where

$rev_{crop,t}$ and $ret_{crop,t}$ = revenue and net returns, respectively, in a specific year, *t,* from a specific crop selected in year *t*
NPV, *SEV*, and *AEI* = financial returns indicators of net present value, soil expectation value, and annual equivalent income
price, *yield*, *cost* = annual returns elements of product price, product yield, and input cost, respectively, for an individual crop or timber product
t = year
m = maximum time horizon or the rotation length of the pine plantation in years
r = discount rate
$crop_t$ = the specific annual crop planted in year *t* (chosen from a set of alternative crops using a selection method described below)
pine = pine timber product
subs = annual government payments/ha related to a particular crop or pine (relevant only when the policy option is activated) and may be a function of revenues or costs.

Table 2 describes the model outputs, which include averages and distributions of NPV, SEV, and AEI (in matrices MT, PT, and AT), management variables such as timber rotation age (PROT and AROT), and number of years each annual crop is selected (MNY and ANY). The user can also view other, more detailed outputs, such as average and individual price and yield paths over time.

Model Options

The user can activate or deactivate several pre-programmed options in the params.m script. These options describe the incorporation of new alternative crops with different risk profiles, different assumptions about the level of interspecific competition between agroforestry system components, and changes in policy (table 3). To activate an option, the user inputs 1 or 0 at the prompt after running params.m. If selected, params.m will also ask the user to input various necessary parameters for each.

Table 2—Outputs of ALLEY Model 2.0

Outputs of monocrop.m

Name of output	Format	Description
MT	*n* x (*q*(1)+3) matrix	Financial returns indicators for *n* iterations, based on full *m*-year period. Columns represent: 1: NPV; 2: SEV; 3: AEI; 4 through *q*(1)+3: number of years each crop is selected for cultivation.
MU	*n* x *m* x (*q*(1)+2) matrix	Yearly indicators for all *m* years, over all *n* iterations. Pages (third dimension) represent: 1 through *q*(1): returns from each crop; *q*(1)+1: profit (including returns and policy payments) from selected crop; *q*(1)+2: policy payments for selected crop.
MV	*n* x *m* x (sum(*q*)+2**q*(1)) matrix	Yearly indicators for all *m* years, over all *n* iterations. Pages represent: 1 through *q*(1): prices for all crops; *q*(1)+1 through sum(*q*(1:2)): yield for all crops; sum(*q*(1:2))+1 through sum(*q*(1:3)): cost for all crops; sum(*q*(1:3))+1 through sum(*q*(1:3))+*q*(1): revenues for all crops; sum(*q*(1:3))+*q*(1)+1 through sum(*q*(1:3))+2**q*(1)): returns for all crops.
MRET	*m* x 2*(*q*(1)+2) matrix	Mean and standard deviations of yearly indicators given in MU, over *n* iterations.
MSUBS	scalar	Overall mean annual policy payments.
MCROPS	*m* x (sum(*q*)+2**q*(1)) matrix	Mean of yearly indicators given in MV, over *n* iterations.
MSEV	1 x 2 vector	Mean and standard deviations of SEV from MT column 2.
MAEI	1 x 2 vector	Mean and standard deviations of AEI from MT column 3.
MNY	1 x *q*(1) vector	Mean number of years during the *m*-year time horizon that each crop is selected for planting.

Outputs of pine.m

Name of Output	Format	Description
PT	*n* x 3 matrix	Financial returns and other outputs for *n* iterations, based on full period of the length of the pine rotation. Columns represent: 1: NPV; 2: SEV; 3: AEI; 4: pine rotation length.
PU	*n* x *m* x 4 matrix	Yearly indicators for all *n* iterations. Pages represent: 1: costs; 2: revenues; 3: policy payments; 4: estimated SEV at that age.
PPROF	*m* x 4 matrix	Average values for each indicator in PU, at each possible rotation age.
SPROF	*m* x 4 matrix	Standard deviations for each indicator in PU, at each possible rotation age.
PSUBS	Scalar	Mean annual policy payments.
PSEV	1 x 2 vector	Mean and standard deviations of SEV from PT column 2.
PAEI	1 x 2 vector	Mean and standard deviations of AEI from PT column 3.
PROT	1 x 2 vector	Mean and standard deviations of optimal rotation length from PT column 1.
catyear	vector	List of years of catastrophic destruction of timber.

(continued)

Table 2 (*continued*)—Outputs of ALLEY Model 2.0

Outputs of alleycrop.m

Name of Output	Format	Description
AT	*n* x (*q*(1)+4) matrix	Financial returns and other outputs for *n* iterations, based on full period of the length of the pine rotation. Columns represent: 1: alley crop pine rotation length; 2: SEV; 3: AEI; 4: alley crop pine rotation length; 5 through *q*(1)+4: number of years each crop is selected for cultivation.
AU	*n* x *m* x (2**q*(1)+6) matrix	Yearly indicators for all *m* years, over all *n* iterations. Pages represent: 1 through *q*(1): returns from each crop; *q*(1)+1: profit (including returns and policy payments) from selected crop; (*q*(1)+2) through (2**q*(1)+1): number of years each crop is selected for cultivation up to that year; (2**q*(1)+2): pine costs; (2**q*(1)+3): revenues; (2**q*(1)+4): pine policy payments; (2**q*(1)+5): policy payments for selected crop; (2**q*(1)+6): estimated SEV at that age.
AV	*n* x *m* x sum(*q*)+2**q*(1)+1 matrix	Yearly indicators for all *m* years, over all *n* iterations. Pages represent: 1 through sum(*q*): values for all returns elements; (sum(*q*)+1) though (sum(*q*)+*q*(1)): revenues for all crops; sum(*q*)+*q*(1)+1 through sum(*q*)+2**q*(1)): returns for all crops; sum(*q*)+2**q*(1)+1: pine NPV at that age.
ARET	*m* x (2**q*(1)+6) matrix	Mean and standard deviations of yearly indicators given in MU, over *n* iterations.
ASUBS	Scalar	Mean annual policy payments.
ACROPS	*m* x sum(*q*)+2**q*(1)+1 matrix	Mean of yearly indicators given in AV, over *n* iterations.
APNPV	*m* x 1 vector	Average NPV at each year.
ASEV	1 x 2 vector	Mean and standard deviations of SEV from AT column 2.
AAEI	1 x 2 vector	Mean and standard deviations of AEI from AT column 3.
ANY	1 x *q*(1) vector	Mean number of years during the time horizon of the alley crop pine rotation that each crop is selected for planting.
AROT	1 x 2 vector	Mean and standard deviations of optimal rotation length from PT column 1.
catyear	vector	List of years of catastrophic destruction of timber.

Table 3—Model options, activated by inputting 1 at prompt in the params.m script

Model option	Description
Specialty crops	Includes the possibility to include hypothetical or real specialty crops, alongside commodity crops for which, presumably, parameters have been estimated from data. Unlike commodity crops, specialty crops are unlikely to have long time series historical data on price and yield in a locality.
Competition function	Includes a competition function that alters the yield of annual crops in an alley crop setting due to interspecific interactions between trees and annual crops (0 = no competition function, 1 = competition function).
Policy	Includes Farm Bill ARC-IC program payments for annual crops and 50 percent cost-share of plantation establishment for pine (0 = no policy payments, 1 = policy payments).

Obtaining and Utilizing the Model

ALLEY Model 2.0 is open source and public domain. When referencing ALLEY Model 2.0, please provide appropriate citation to this document. The authors ask that users please provide input and feedback on the model in order to improve and understand uses. Please send an email to gregory.e.frey@usda.gov telling us about the purpose for using the model, the model's usefulness, and suggestions for improving the model.

The full ALLEY Model 2.0 suite contains 12 total MATLAB files, which are listed in table 4. The model is accompanied by a sample data file in Microsoft® Excel format, entitled data.xlsx. These sample data are the data described in the Example Application section below. The ALLEY Model 2.0 software suite and sample data file can be accessed at https://www.srs.fs.usda.gov/pubs/gtr/gtr_srs235/.

Table 4—Scripts and functions in the ALLEY Model 2.0 suite

Scripts (called by user)	Description	Typical order
estimate.m	Estimates parameters *S*, *D*, *sigmaalley*, *sigmamono*, and *sigmapine*.	1
params.m	Loads the basic parameters for all subsequent scripts and functions into memory.	2
alleycrop.m	Simulation of alley crop system with pine and an annual crop, for up to *m* years	3–5
monocrop.m	Simulation of monocrop system over *m* years, with *m* set to 40 for the results given here.	3–5
pine.m	Simulation of pine plantation, for up to *m* years.	3–5
Functions (called by scripts or other functions)	**Description**	**Called by**
arc.m	Calculates payment due from ARC-IC program as described in FSA (2014).	monocrop.m and alleycrop.m
compfunc.m	Defines the competition function that reduces yield on annual crop based on size of trees in alley crop system.	alleycrop.m
harvest.m	Determines if a thinning or clearcut operation will occur, the number and which trees will be cut, and the harvested timber volume. Based on stand parameters; programmed decision parameters *maxage*, *BAdecSI*, *thinlimit*, and *thinparams*; and programmed harvest parameters *sawmin*, *sawtop*, *chipmin*, *chiptop*, *pulpmin*, and *pulptop*.	alleycrop.m and pine.m
lagmat.m	Creates a matrix of lagged time series.	estimate.m
pinegrowth.m	Generates pine growth and yield for pines as described in Burkhart and others (2008).	pine.m and alleycrop.m
relements.m	Generates returns elements (price, yield, input cost) based on a copula approach, starting with randomly generated shock and converting to appropriate values using selected autoregression-trend function and distribution.	alleycrop.m, monocrop.m, and pine.m
youngpinegrowth.m	Generates pine growth and yield for young pines (before onset of intraspecific competition) as described in Westfall and others (2004).	pine.m and alleycrop.m

Data Sources

It is recommended that, for the most realistic models, the input parameters be based on historical data. When adequate historical data are not available, it may be necessary to make certain assumptions about parameter values or utilize default values. Some sources of data are freely available on the Internet and appropriate for the purposes of this model. For example, historical crop data can be found at the U.S. Department of Agriculture (USDA) Economic Research Service's "Commodity Costs and Returns" website, based on the Agricultural Resource Management Survey (ERS 2016), or from USDA National Agricultural Statistics Service's "Quick Stats" website (NASS 2016). Pine growth and yield parameters and equations from Westfall and others (2004) and Burkhart and others (2008) are pre-programmed into the model. Historical pine sawtimber and pulpwood price data can be obtained from subscription services, such as TimberMart-South, or online from such services as North Carolina Cooperative Extension's "Price Report" (NCCE 2014) or Louisiana Department of Agriculture and Forestry's "Quarterly Report of Forest Products" (LADAF 2016). Pine plantation historical cost data in the U.S. South can be found in a series of publications from Alabama Cooperative Extension Service (e.g., Dooley and Barlow 2013) or other local sources.

Data Format

Units—The model is generally conceived in terms of metric tons for volume/mass, ha for land area, and US dollars (US$) for monetary values. This translates into metric tons/ha for yield, US$/metric ton for price, and US$/ha for costs. However, the model does not place stringent requirements on units, with the exception of land area. All land area units should be in ha, and correspondingly, yield and costs should be in volume or mass/ha and monetary units/ha. Otherwise, the model can accept the units desired by the user.

It is generally recommended that the user provide standard volume/mass units that apply to all crops, such as metric tons rather than bushels for corn and hundredweight for cotton; however, as long as the same units are used for price and yield for each individual crop, the model will function.

Adjustment for Inflation—When using historical data, it is important to realize that monetary values are affected by inflation over time. Most historical price and cost datasets are given in nominal terms; that is, they do not adjust for inflation. This model assumes real (inflation-adjusted) prices and costs; therefore, historical data must be adjusted before being imported into the model. An inflation index such as the producer or consumer price index (PPI or CPI) is appropriate for this task.

Crop Costs—Yield in units such as metric tons/ha and price in dollars per metric ton is relatively straightforward. However, "costs of production" in various contexts may or may not include a wide variety of fixed and variable costs, opportunity costs, overhead costs, etc. These decisions are left to the discretion of the user, and we recommend detailed documentation of the assumptions made. However, the data provided utilize one set of assumptions, and where possible and appropriate, the authors recommend following the convention we use for our data files and default values: include those costs related to direct expenses, allocated overhead, depreciation, and contract services.[6] The costs we do not include are opportunity costs of land or unpaid labor.[7]

[6] These categories in the ERS (2016) Commodity Costs and Returns spreadsheets were used: seed; fertilizer, lime, and gypsum; chemicals; custom operations; fuel, lube, and electricity; repairs; hired labor; purchased irrigation water; interest; taxes and insurance; general farm overhead; capital recovery of machinery and equipment; ginning.

[7] These categories in the ERS (2016) Commodity Costs and Returns spreadsheets were not used: unpaid labor; land.

Pine Costs—Pine costs include (1) establishment costs in year 1, (2) hardwood release/competition control in year 8, and (3) annual management costs in every year. Since very different types of costs, with different relative values, are undertaken in any given year, it is difficult to model this variability in costs as a single variable. However, we assume these costs to be strongly positively correlated. Therefore, we considered it feasible to model a single stochastic pine input cost index variable. The historical cost index was derived by summing the historical relevant real costs in a given year and then normalized by dividing by the average of those costs for all historical years available. This creates a historical cost index of average 1. This variable is then multiplied by a fixed average value for the activity that is scheduled to occur in a given year.

Crops with Multiple Products—Some annual crops produce two or more products that are measured or marketed separately. An example of this is cotton, which produces cotton lint and cottonseed. Historical data on yield may give separate values for the yield and price of the cotton lint and cottonseed. Each product contributes a significant portion of the total revenue provided from cotton crops, although cotton lint is the primary component of revenue.

Rather than further complicate the model coding by adding different components to model the multiple products separately, each with stochastic yields and price, the user must address this issue directly in the data. This was a decision made in the development of ALLEY Model 2.0, in order to reduce the complexity of the model, at the expense of some nuance.

The method for addressing this in the data is best explained through an example. Our historical data provided historical annual yields of both cotton lint and cottonseed, and historical annual prices for both products. We first calculated a historical annual cotton revenue by summing the revenues from both products. We then found an "imputed price of cotton lint" by dividing the total revenue by the yield of cotton lint. Thus in ALLEY Model 2.0 we can model cotton lint yield as our product and this "imputed price" as our price. The imputed price includes the additional value of the cottonseed.

Data File Organization—If historical data in an Excel file are used to estimate the parameters, the data should be organized with years in the rows and returns elements in columns. The first column states the years, in descending order (most recent year at the top). Following that, the next columns should include the data for annual crop output prices, in real (inflation-adjusted) dollars per unit. There is one column for output price for each crop. The next columns are annual crop yield in units/ha, one column for each crop, in the same order as for the output prices. The next columns are input costs in real US$/ha, one column for crop, in the same order as prices and yields.

Following the crop data, the next columns are timber prices, in real dollars per ton. The prices should include pine sawtimber and pulpwood historical data. The units should be the same for both timber products, which is metric tons. This is because the timber yield model (independent from other returns elements) outputs tons/ha. After the timber price, the final row should be an index of real pine plantation input costs (described above in the section on Pine Costs).

The data file should be in the same folder with the MATLAB scripts and functions, and named data.xlsx. The data should be listed in a worksheet titled 'Data.'

Detailed Model Construction

The ALLEY Model 2.0 suite includes a set of scripts and functions that model stochastic price, yield, and cost processes, in a Monte Carlo framework. The models are for multiple alternative annual crops, and for loblolly pine (*Pinus taeda*). The MATLAB programs we created are available from the authors and include various scripts, which are called by the user in MATLAB, and functions, which are called as needed by the scripts and other functions (table 4).

The basic model structure follows the following outline:

- Parameter estimation for returns elements from historical data using the script estimate.m
- Input of other parameters and model options using the script params.m
- Monte Carlo simulation to generate returns elements time paths and modeling of decision rules using scripts monocrop.m, pine.m, and alleycrop.m (each of these three scripts calls the function relements.m for returns elements time paths)
- Calculation and output of the financial returns indicators and management variables, over the *m*-year time horizon, over *n* iterations

Parameter Estimation for Returns Elements

The first step in the ALLEY Model 2.0 suite is to load values for the vectors *trendall* and *distall* and estimate or load values for the matrices *S*, *D*, and *sigma*. *trendall* and *distall* are matrices that store the selected functional forms for the autoregression-trend function and distribution function. The matrix *S* represents the parameters for the autoregression-trend function for all returns elements for all crop/timber products, and *D* represents the parameters of the distribution functions for the shock values for all returns elements for all crop/timber products (table 1). Both matrices are of size 4 x *q*(1) x *ne*, where *q* is the matrix of the number of crops and pine products for each returns element and *ne* is the number of returns elements to be simulated, with each returns element for each crop/pine occupying a column (second dimension) on each page (third dimension) (table 1). The matrix *sigma* is the covariance matrix of normalized shocks for all returns elements for all crop/timber products (table 1). The script estimate.m imports historical data on returns elements from an Excel file[8] to estimate *S*, *D*, and *sigma*, based on function and distribution types set by the user in *trendall* and *distall*. The authors recommend using this approach. It is, however, possible to load values for *S*, *D*, and *sigma* directly into memory.

In either case, the first step here is to set values for the row vectors *trendall* and *distall*. *trendall* tells the scripts (estimate.m and others that call the function relements.m) which of seven types of autoregression-trend functions describes each returns element. *distall* tells the scripts which of three types of distribution functions describes the annual shocks for each returns element. Each returns element occupies one value in the vector, in the same order as the historic data in the Excel file. All crops use the same value of *trendall* and *distall* for a particular risk element.

[8] The location and name of the Excel file should be typed into the data import line with the command function xlsread in the script estimate.m. Alternatively, if the data are not in an Excel format, the user can either replace the import command line in estimate.m with a matrix of data directly, or change the import command using one of numerous data import tools available. See https://www.mathworks.com/help/matlab/import_export/supported-file-formats.html.

Possibilities for each value in *trendall* include various possible trends and autoregression types. The coefficients from fitting the model are used to populate the matrix *S*. In the following, *y* represents the returns element to be estimated (price, yield, etc.) for each crop/pine.

0. $trendall_k == 0$: full mean reversion plus shock

$$y_t = \bar{y} + \varepsilon_t \quad \text{(E.T.0.1)}$$

$$\hat{y}_t = \bar{y} \quad \text{(E.T.0.2)}$$

$$\hat{\varepsilon}_t = y_t - \hat{y}_t \quad \text{(E.T.0.3)}$$

$$S = \bar{y} \quad \text{(E.T.0.4)}$$

1. $trendall_k == 1$: random walk - no mean reversion, no time trend

$$y_t = y_{t-1} + \varepsilon_t \quad \text{(E.T.1.1)}$$

$$\hat{y}_t = y_{t-1} \quad \text{(E.T.1.2)}$$

$$\hat{\varepsilon}_t = y_t - \hat{y}_t \quad \text{(E.T.1.3)}$$

$$S = 0 \quad \text{(E.T.1.4)}$$

2. $trendall_k == 2$: no time trend, AR1 autoregression

$$y_t = \beta_1 + \beta_2 \cdot y_{t-1} + \varepsilon_t \quad \text{(E.T.2.1)}$$

$$\hat{y}_t = \hat{\beta}_1 + \hat{\beta}_2 \cdot y_{t-1} \quad \text{(E.T.2.2)}$$

$$\hat{\varepsilon}_t = y_t - \hat{y}_t \quad \text{(E.T.2.3)}$$

$$S = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} \quad \text{(E.T.2.4)}$$

3. $trendallk == 3$: linear time trend, AR1 autoregression

$$y_t = \beta_1 + \beta_2 \cdot t + \beta_3 \cdot y_{t-1} + \varepsilon_t \quad \text{(E.T.3.1)}$$

$$\hat{y}_t = \hat{\beta}_1 + \hat{\beta}_2 \cdot t + \hat{\beta}_3 \cdot y_{t-1} \quad \text{(E.T.3.2)}$$

$$\hat{\varepsilon}_t = y_t - \hat{y}_t \quad \text{(E.T.3.3)}$$

$$S = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} \quad \text{(E.T.3.4)}$$

4. $trendall_k == 4$: logarithmic time trend, AR1 autoregression

$$y_t = \beta_1 - \frac{1}{ln(1-\beta_2)} \cdot \beta_2{}^t + \beta_3 \cdot y_{t-1} + \varepsilon_t \quad \text{(E.T.4.1)}$$

$$\hat{y}_t = \hat{\beta}_1 - \frac{1}{ln(1-\hat{\beta}_2)} \cdot \hat{\beta}_2{}^t + \hat{\beta}_3 \cdot y_{t-1} \quad \text{(E.T.4.2)}$$

$$\hat{\varepsilon}_t = y_t - \hat{y}_t \quad \text{(E.T.4.3)}$$

$$S = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} \quad \text{(E.T.4.4)}$$

Autoregression-Trend and Distribution Functions

The following *trendall* and *distall* functions [e.g., "(**E.T.0.1**)"] are enumerated as follows:

The first letter is either **E** for "estimate" or **S** for "simulate."

The second letter is **T** for "trend" or **D** for "distribution."

The first number is for the autoregression-trend or distribution type number, which corresponds to the value in the *trendall* vector.

The final number is the equation number.

5. $trendall_k$ == 5 : increasing exponential time trend, AR1 autoregression

$$y_t = \beta_1 - e^{\beta_2 \cdot t} + \beta_3 \cdot y_{t-1} + \varepsilon_t \quad \text{(E.T.5.1)}$$

$$\hat{y}_t = \hat{\beta}_1 - e^{\hat{\beta}_2 \cdot t} + \hat{\beta}_3 \cdot y_{t-1} \quad \text{(E.T.5.2)}$$

$$\hat{\varepsilon}_t = y_t - \hat{y}_t \quad \text{(E.T.5.3)}$$

$$S = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} \quad \text{(E.T.5.4)}$$

6. $trendall_k$ == 6 : decreasing exponential time trend, AR1 autoregression

$$y_t = \beta_1 + \beta_2 \cdot e^{\beta_3 \cdot t} + \beta_4 \cdot y_{t-1} + \varepsilon_t \quad \text{(E.T.6.1)}$$

$$\hat{y}_t = \hat{\beta}_1 - \hat{\beta}_2 \cdot e^{\hat{\beta}_3 \cdot t} + \hat{\beta}_4 \cdot y_{t-1} \quad \text{(E.T.6.2)}$$

$$\hat{\varepsilon}_t = y_t - \hat{y}_t \quad \text{(E.T.6.3)}$$

$$S = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{bmatrix} \quad \text{(E.T.6.4)}$$

where

trendall = vector for selection of autoregression-trend functional form for each returns element
k = index of the returns element
y_t = value of returns element in year t
$\hat{y}$ = estimated value of y
$\bar{y}$ = historical mean of y
ε = unobserved shock term
$\hat{\varepsilon}$ = residual from regression
S = parameters for the autoregression-trend function for all returns elements for all crop/timber products
β, $\hat{\beta}$ = true and estimated coefficients of the autoregression-trend function

Possibilities for each value in *distall* include normal- (*N*), lognormal- (*LN*), and beta- (*Beta*) distributed residuals. Coefficients from fitting these distributions to the data are used to populate the matrix *D*. Since the residuals from the autoregression-trend functions above will be centered around 0, whereas the lognormal distribution is strictly positive and the beta distribution is on the interval [0, 1], it is first necessary to shift the values of the residuals to those intervals before fitting the distribution to them. The residuals, ε are adjusted to new values, $\hat{g}$, using adjustment factors *minadj* and *maxadj* based on the minimum and maximum values in ε and a pre-defined percent deviation value *perc*. Once a distribution is fit to the new values in $\hat{g}$, they are converted to a standard normal distribution defined by :

0. $distall_k$ == 0 : normal distribution

$$\hat{\varepsilon} \sim N(0, \hat{\sigma}) \rightarrow \hat{t} \sim N(0,1) \quad \text{(E.D.0.1)}$$

$$D = \begin{bmatrix} \hat{\sigma} \\ 0 \end{bmatrix} \quad \text{(E.D.0.4)}$$

1. $distall_k$ == 1 : lognormal distribution

$$minadj := \frac{\min(\hat{\varepsilon})}{perc} \tag{E.D.1.1}$$

$$\hat{g}_t = \hat{\varepsilon}_t - minadj \tag{E.D.1.2}$$

$$\hat{g} \sim LN(\hat{d}_1, \hat{d}_2) \rightarrow \hat{\tau} \sim N(0,1) \tag{E.D.1.3}$$

$$D = \begin{bmatrix} minadj \\ \hat{d}_1 \\ \hat{d}_2 \end{bmatrix} \tag{E.D.1.4}$$

2. $distall_k$ == 2 : beta distribution

$$minadj := \frac{\min(\hat{\varepsilon})}{perc} \tag{E.D.2.1}$$

$$maxadj := \frac{\max(\hat{\varepsilon})}{perc} \tag{E.D.2.2}$$

$$\hat{g}_t = \frac{\hat{\varepsilon}_t - minadj}{maxadj - minadj} \tag{E.D.2.3}$$

$$\hat{g} \sim Beta(\hat{d}_1, \hat{d}_2) \rightarrow \hat{\tau} \sim N(0,1) \tag{E.D.2.4}$$

$$D = \begin{bmatrix} minadj \\ maxadj \\ \hat{d}_1 \\ \hat{d}_2 \end{bmatrix} \tag{E.D.2.5}$$

where

distall = vector for selection of distribution functional form for the shock values for each returns element
minadj, *maxadj* = adjustment factors
perc = percent deviation value
$\hat{g}$ = adjusted residual
$\hat{\tau}$ = normalized residual
D = parameters for the distribution functions for the shock values for all returns elements for all crop/timber products
$\hat{\sigma}$ = estimated standard deviation of the distribution
$\hat{d}_1$, $\hat{d}_2$ = other estimated coefficients of the distribution, based on fit

A single covariance matrix *sigma* is then estimated for all the standardized residuals, $\hat{\tau}$, for all returns elements and crops/pine:

$$sigma = \begin{bmatrix} 1 & cov(\hat{\tau}_{1,1}, \hat{\tau}_{1,2}) & \cdots & cov(\hat{\tau}_{ne,1}, \hat{\tau}_{ne,c}) \\ cov(\hat{\tau}_{1,2}, \hat{\tau}_{1,1}) & 1 & \cdots & cov(\hat{\tau}_{ne,2}, \hat{\tau}_{ne,c}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(\hat{\tau}_{ne,c}, \hat{\tau}_{ne,1}) & cov(\hat{\tau}_{ne,c}, \hat{\tau}_{ne,2}) & \cdots & 1 \end{bmatrix} \tag{6}$$

where

c = total number of crops and pine products

The script estimate.m also outputs matrices for *seedall* and *rh*, based on the imported data. *seedall* is a matrix of the most recent historical values of each returns element for each crop. This information is stored and used for returns elements that have autoregressive characteristics in the autoregression-trend function, to provide a value for y_{t-1} when $t = 0$ in the Monte Carlo simulation. *rh* is a matrix of the most recent 5 years of historical revenue from each annual crop. This assumes that price and yield are the first two returns elements in the data matrix and multiplies them

together. The script simply loads the appropriate values from the imported data matrix. If the cells for the appropriate values for *seedall* and *rh* are empty in the data matrix (i.e., missing value, denoted in MATLAB as "NaN"), the script estimate.m will prompt the user to enter a value.

Quality Control—Within the framework described above, there is significant freedom to vary the types of trends and residual distributions of the returns elements. Users should first select autoregression-trend functions and residual distributions based on their own understanding of returns elements. After estimating the parameters, it is good practice to run the model to ensure that outputs are reasonable for the time period of the simulation.

Some processes in agriculture certainly display trends that will likely continue into the future. For example, crop yields have steadily increased over time due to technological advances. At the same time, prices received for crops have declined steadily. Therefore, if initial runs of the model deviate significantly from historical norms, the user may consider alternate autoregression-trend functions or distribution functions given above.

Input of Other Parameters

If the user has previously estimated the parameters discussed in the section above on Parameter Estimation for Returns Elements (specifically: *S*, *D*, *sigma*, *seedall*, *rh*), these may be entered directly into memory without the need to run the script estimate.m. In this case, the value for other parameters coded in estimate.m should also be entered directly into memory (specifically: *nec*, *nep*, *ne*, *q*, *trendall*, *distall*).

Apart from the parameters discussed above in the section on Parameter Estimation for Returns Elements, ALLEY Model 2.0 utilizes numerous other parameters related to management and decision making, which are primarily coded in the params.m script. A list of key parameters is given in table 1. Lists of other parameters are in the appendix (tables A.1–A.3). The default parameters can be altered by directly coding them in the script params.m. Once satisfied with the input parameters, the user can run the script params.m in the MATLAB interface to load the parameters into memory.

The params.m script also prompts the user for variables related to the three model options (see Model Options under the Model Overview section above as well as the Detailed Model Options section below). If the model options are selected, params.m will create new variables *corrprofile* and *compprofile* [table A.1 (appendix)] and alter the values of *S*, *D*, *sigma*, *seedall*, and *rh*, as relevant.

Monte Carlo Simulation

Once the returns elements parameters are estimated and/or loaded and other parameters loaded into memory, the user may begin generating Monte Carlo simulations. The Monte Carlo simulations are run by calling the scripts monocrop.m, pine.m, and alleycrop.m. The Monte Carlo simulation consists of four basic steps:

1. Generate simulated returns elements for the *m*-year time horizon.

2. Model various decision rules and other factors for the *m*-year time horizon.

3. Calculate the financial returns indicators for the *m*-year time horizon.

4. Repeat steps 1–3 for *n* iterations.

More detailed descriptions of the monocrop.m, pine.m, and alleycrop.m scripts are given in the schematic flowcharts in figures 2–4.

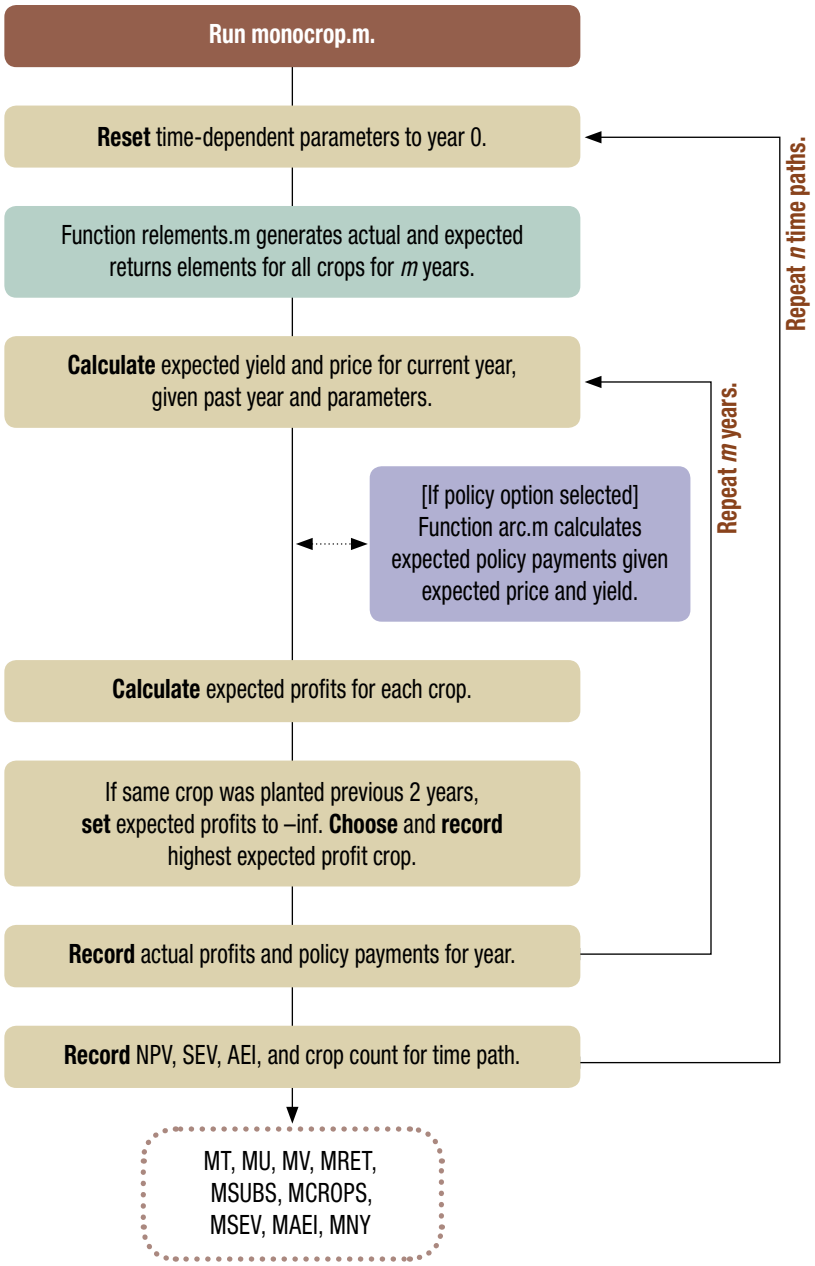

Figure 2—Basic schematic flowchart of the monocrop.m script.

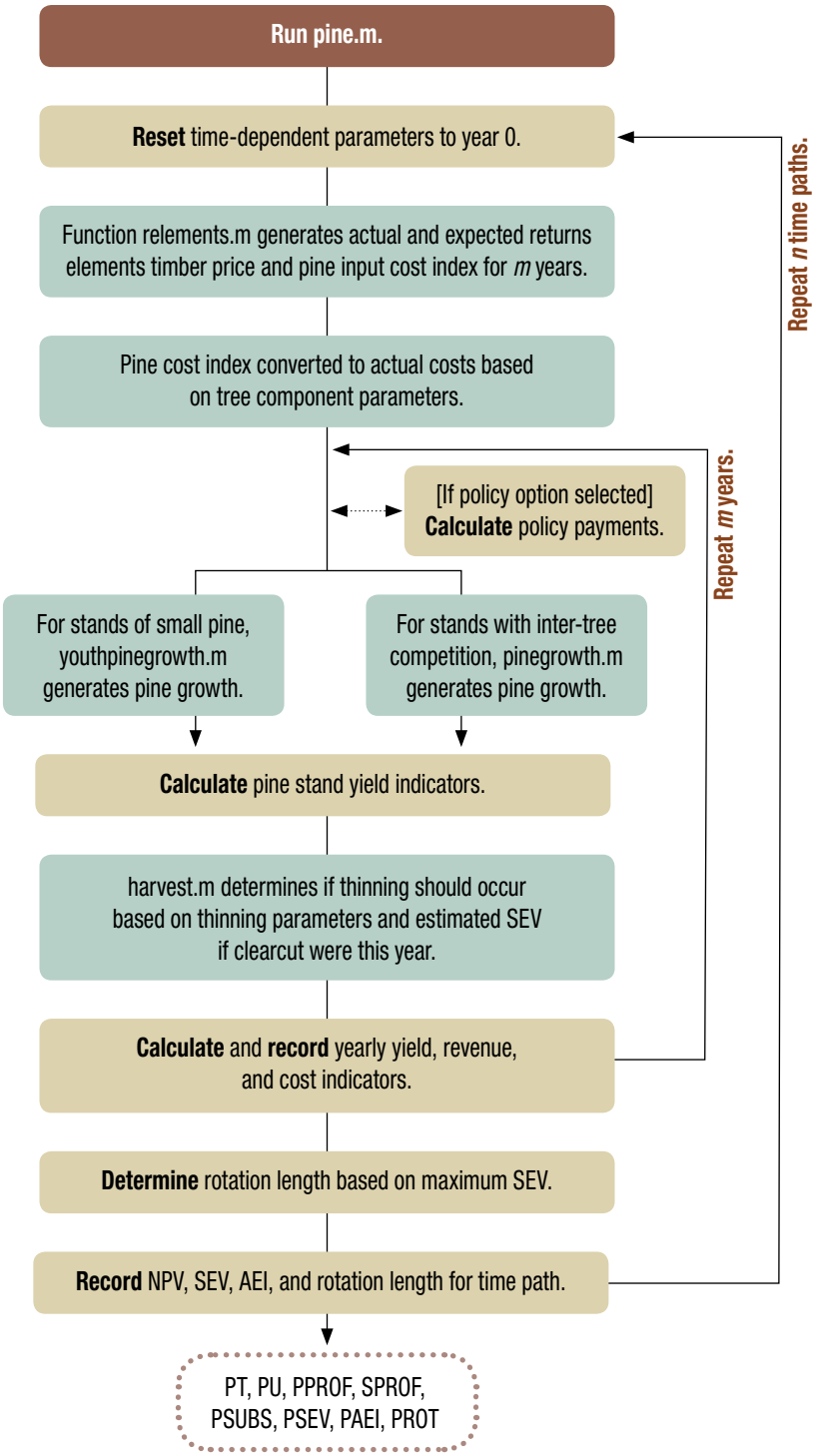

Figure 3—Basic schematic flowchart of the pine.m script.

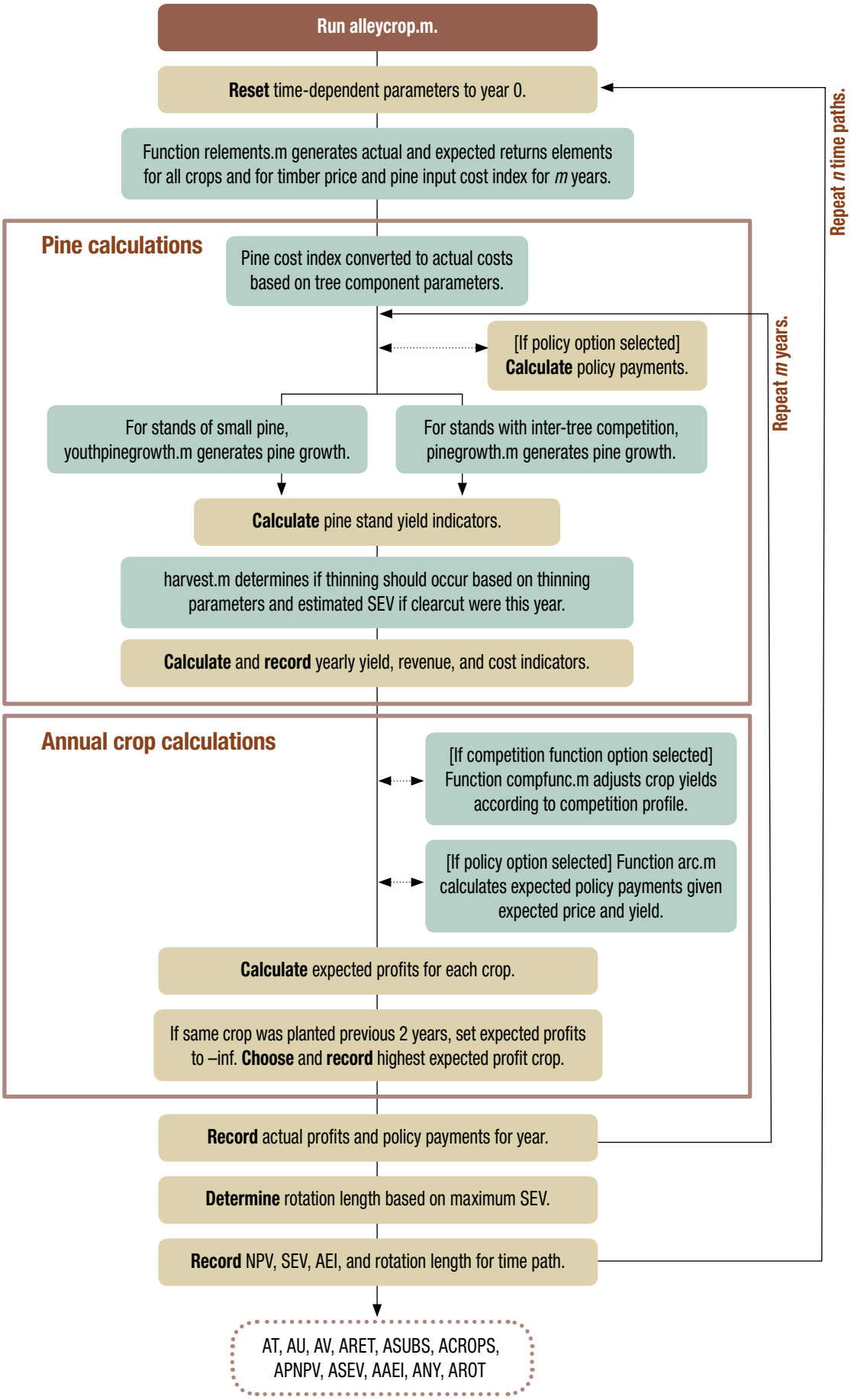

Figure 4—Basic schematic flowchart of the alleycrop.m script.

Generation of the Simulated Returns Elements for the *m*-year Time Horizon—Each of the three Monte Carlo simulation scripts (monocrop.m, pine.m, and alleycrop.m) begins each iteration by calling the function relements.m. relements.m is the function that uses the estimated or loaded returns elements parameters including *trendall*, *distall*, *seedall*, *S*, *D*, and *sigma* to generate simulated values for each of the returns elements over an *m*-year time horizon.

The approach used to generate the returns elements starts with a Gaussian copula that generates jointly distributed random shocks using a multivariate normal distribution based on a mean value that is a vector of zeros and *sigma* that is estimated and loaded based on historical data (see section above on Parameter Estimation for Returns Elements). Since the copula generates a joint return, it allows correlation between returns elements within and between crops. This is important since price and yield, for example, are well known to be negatively correlated. Also, it is likely that yields of different crops are positively correlated (the same weather is probably good for most crops, in most circumstances). These correlations are encapsulated in the shock terms in our model, as a fundamental underlying assumption. This is important to note, as it is possible to imagine other ways to model these correlations.

Once the random normal shocks, τ, are generated, they are converted to actual values of returns elements (in the same units as provided in the original data used to estimate the parameters), using a process that is essentially the reverse of the process described in the section above on Parameter Estimation for Returns Elements.

First, the random standard normal shocks, τ, are converted to shocks of the specified distribution, g, in *distall*, then adjusted to the final shock values, ε, using specified parameters in *D*:

0. $distall_k == 0$: normal distribution

$$\tau \sim N(0,1) \rightarrow \varepsilon \sim N(0, \hat{\sigma}) \quad \text{(S.D.0.1)}$$

1. $distall_k == 1$: lognormal distribution

$$\tau \sim N(0,1) \rightarrow g \sim LN(\hat{d}_1, \hat{d}_2) \quad \text{(S.D.1.1)}$$

$$\varepsilon_t = g_t + minadj \quad \text{(S.D.1.2)}$$

2. $distall_k == 2$: beta distribution

$$\tau \sim N(0,1) \rightarrow g \sim Beta(\hat{d}_1, \hat{d}_2) \quad \text{(S.D.2.1)}$$

$$\varepsilon_t = g_t \cdot (maxadj - minadj) + minadj \quad \text{(S.D.2.2)}$$

where

τ, g, and ε = simulated random normal, redistributed, and final adjusted shocks, respectively

Then, the converted shocks (now in units of each risk element) are added to the expected value of the risk element, $E(y_t)$, based on the autoregression-trend functions:

0. $trendall_k == 0$: full mean reversion plus shock

$$E(y_t) = \bar{y} \quad \text{(S.T.0.1)}$$

$$y_t = E(y_t) + \varepsilon_t \quad \text{(S.T.0.2)}$$

1. $trendall_k == 1$: random walk - no mean reversion, no time trend

$$E(y_t) = y_{t-1} \quad \text{(S.T.1.1)}$$

$$y_t = E(y_t) + \varepsilon_t \quad \text{(S.T.1.2)}$$

2. $trendall_k$ == 2 : no time trend, AR1 autoregression

$$E(y_t) = \beta_1 + \beta_2 \cdot y_{t-1} \quad \text{(S.T.2.1)}$$

$$y_t = E(y_t) + \varepsilon_t \quad \text{(S.T.2.2)}$$

3. $trendall_k$ == 3 : linear time trend, AR1 autoregression

$$E(y_t) = \beta_1 + \beta_2 \cdot t + \beta_3 \cdot y_{t-1} \quad \text{(S.T.3.1)}$$

$$y_t = E(y_t) + \varepsilon_t \quad \text{(S.T.3.2)}$$

4. $trendall_k$ == 4 : logarithmic time trend, AR1 autoregression

$$E(y_t) = \beta_1 - \frac{1}{ln(1-\beta_2)} \cdot \beta_2^{\,t} + \beta_3 \cdot y_{t-1} \quad \text{(S.T.4.1)}$$

$$y_t = E(y_t) + \varepsilon_t \quad \text{(S.T.4.2)}$$

5. $trendall_k$ == 5 : increasing exponential time trend, AR1 autoregression

$$E(y_t) = \beta_1 - e^{\beta_2 \cdot t} + \beta_3 \cdot y_{t-1} \quad \text{(S.T.5.1)}$$

$$y_t = E(y_t) + \varepsilon_t \quad \text{(S.T.5.2)}$$

6. $trendall_k$ == 6 : decreasing exponential time trend, AR1 autoregression

$$E(y_t) = \beta_1 - \beta_2 \cdot e^{\beta_3 \cdot t} + \beta_4 \cdot y_{t-1} \quad \text{(S.T.6.1)}$$

$$y_t = E(y_t) + \varepsilon_t \quad \text{(S.T.6.2)}$$

The function relements.m then outputs to the script both the actual values of the risk elements, *y*, and the expected values, *E*(*y*), for the *m*-year time horizon. The expected values are used in several of the decision rules.

Sawtimber and pulpwood price and pine plantation input cost index are modeled in an identical manner to the crop returns elements.

Timber growth and yield (pine plantation and alley crop models)—While sawtimber and pulpwood price and pine plantation input cost index are modeled in an identical manner to the crop returns elements, timber growth and yield are modeled differently. Timber grows slowly over multiple years, and financial gains occur only when a periodic harvest is conducted, unlike annual crops. Since only a fraction of standing timber is harvested in any given year, timber price in a given year is not likely to be strongly correlated with the growth rates achieved in that year, unlike agriculture where there is assumed to be strong inverse correlation.

Timber yield is modeled based on biometric equations found in Westfall and others (2004) for young pine growth, prior to intraspecific competition, and Burkhart and others (2008) for older pines. Systems of equations generate individual tree growth for all trees in a representative acre. These systems of equations are found in youngpinegrowth.m and pinegrowth.m functions, respectively. These timber models that simulate incremental annual tree growth are stochastic in nature; that is, the biometric equations include a random residual factor, which is modeled using a random normal number generator in MATLAB. We do not recreate these systems of equations in this publication; the reader should refer to Westfall and others (2004) and Burkhart and others (2008).

In the growth model, variability in individual tree growth is independent of other trees in the stand. This is an unrealistic, but necessary, simplifying assumption, as yearly correlation between growth

of individual trees within stands is unknown. Such correlation would depend on the stand size and variability of the soils and topography, among other factors.

In the pine.m and alleycrop.m scripts, a matrix, *live*, of zeros and ones is created and updated in each year *t* to represent seedling planting sites on a representative acre of land, with ones representing live trees and zeros representing dead, harvested, or no trees [i.e., sites that were not originally planted (as in the alleys where the annual crops are located)]. The *live* matrix forms a basis for youngpinegrowth.m and pinegrowth.m to generate values for diameter at breast height (DBH) of individual trees in a representative acre in a separate matrix *dbh*, tree height in a matrix *height*, crown ratio in a matrix *cr*, trees per acre (TPA) in variable *tpa*, and basal area (BA) in variable *ba*. To estimate growth of trees, the youngpinegrowth.m and pinegrowth.m functions input site index parameter *si*, age of the stand, matrix *U* for young pines and matrices *dbh* and *height* from the previous year for older pines, *live* from the previous year, vector *catyear* representing the years in which catastrophic loss of pines occurs, planting distance parameter *pd*, and catastrophe probability parameters *pcatprob* and *pcateffect*. To use the tree growth to estimate yield of timber by product class, the functions also input parameters *sawmin*, *sawtop*, *chipmin*, *chiptop*, *pulpmin*, and *pulptop.* youngpinegrowth.m and pinegrowth.m input many of these same variables in the subsequent year to simulate growth.

An important note is that, although the input and output parameters for the ALLEY Model 2.0 suite are in SI units (m, ha, metric tons, etc.), all inputs and outputs from youngpinegrowth.m and pinegrowth.m are in English units (feet or inches, acres, U.S. tons, etc.). This is because Westfall and others (2004) and Burkhart and others (2008) use English units in their growth and yield equations. The ALLEY Model 2.0 suite provides all the necessary conversions for seamless use of SI units.

ALLEY Model 2.0 randomly simulates a catastrophic event for timber yield with annual probability *pcatprob*. If a random uniform (0,1) variable is selected that is less than *pcatprob*, a catastrophic event is determined to occur and the year is recorded in the vector *catyear* (additional years can be added so the vector may grow in length through the life of the stand). The variable *pcateffect* includes the mean and standard deviation of the proportion of trees killed in a catastrophic event. The killed trees are randomly selected throughout the representative acre, and the value for the killed trees are set to zero in the matrix *live*.

Modeling of Decision Rules and Other Factors for the *m*-year Time Horizon—Once the returns elements values and expected values are output from the function relements.m into the monocrop.m, pine.m, or alleycrop.m script, each of the three scripts then models decision rules and other factors over the *m*-year time horizon. For example, scripts monocrop.m and alleycrop.m simulate annual returns/ha for each individual crop, based on the three "returns elements" (output prices, yields, input prices). They then use decision rules that determine which crop is planted in any given year. The rules are based on expected value, crop rotation, etc. Once these rules are applied and the crop for each year is selected and other factors applied, as described below, the returns indicators (NPV, SEV, AEI) for the entire *m*-year time horizon are calculated. This process is repeated *n* times for the Monte Carlo simulation to create distributions of the returns indicators. This example and other rules and factors are described in the sections below related to the monocrop, pine plantation, and alley crop models, which represent the scripts monocrop.m, pine.m, and alleycrop.m, respectively.

Crop rotation and selection (monocrop and alley crop models)—Although ALLEY Model 2.0 generates values for returns for all crops each year, it is assumed that the farmer selects only one crop to plant over the entire field each year. For the selection of crop to plant in the monocrop and

alley crop in any given year, we assumed expected profit maximizing farmers and short-run risk neutrality;[9] that is, no discount is taken from crops that are inherently more risky in the short run. We made this simplifying assumption to avoid complicating the model greatly at this stage, but this could be an area for future improvement of the model. We do note, however, that various types of long-run (overall land use) risk aversion can be considered by comparing the distributions of long-run output returns distributions of the three competing land uses. Therefore, while the model assumes risk neutrality in the short run, no such assumption is made in the long run; indeed, evaluating long-run risk is a motivation of this work. While it is possible that farmers may make short-run, year-to-year decisions with a certain level of risk aversion, more likely risk aversion in the long run is a stronger factor influencing a land-use plan, such as adoption of agroforestry.

At the time of planting, a farmer does not have perfect knowledge of what this year's final price or yield will be. Because a farmer cannot predict at the time of planting which crop will generate the greatest profits at harvest, the farmer uses knowledge of last year's profits and the autoregression-trend function to calculate the expected value of returns. It is assumed the farmer does know the parameters *trendall*, *distall*, *S*, *D*, and the previous year's values for all the returns elements but not the random component, ε. Also, the farmer must take into consideration crop rotation.

To address this decision rule, consider a variable u, which has a value for each potential crop that can be selected. In most cases, u will be simply the expected value of the returns per acre for a crop; however, this can be altered in specific cases to accommodate other aspects of the decision (such as crop rotation). The basic decision is to find the maximum u, which thus maximizes expected returns. Again, this is assuming risk neutrality in the short run.

The farmer can thus determine the expected returns, u, from each crop in the current planting year. Functions for estimated values, ŷ, of the returns elements (price, yield, input cost) are given in the section on Generation of the Simulated Returns Elements for the m-year Time Horizon.

$$crop_t = \underset{j \in 1,\ldots,c}{\operatorname{argmax}}\ u_{j,t} \tag{7}$$

$$u_{j,t} = erev_{j,t} + esubs_{j,t}\left(erev_{j,t}, rev_{j,t-1,t-5}\right) - cost_{j,t} \tag{8}$$

where

u = expected returns
j = index of all possible crops
$erev_{j,t}$ = expected revenue for crop j in year t
$esubs_{j,t}$ = expected government payment for crop j in year t

The expected government payment is a function of expected revenue and historical revenues, described later in the section on Detailed Model Options. Expected revenue is in turn determined by the expected price and yield.

$$erev_{j,t} = E\left(rev_{j,t} \middle| price_{j,t-1}, yield_{j,t-1}\right) = E\left(price_{j,t} \cdot yield_{j,t} \middle| price_{j,t-1}, yield_{j,t-1}\right) \tag{9}$$

$$= E\left\{\left[E\left(price_{j,t} \middle| price_{j,t-1}\right) + \varepsilon_{price,j,t}\right] \cdot \left[E\left(yield_{j,t} \middle| yield_{j,t-1}\right) + \varepsilon_{yield,j,t}\right]\right\} \tag{10}$$

[9] While short-run risk neutrality is an assumption of the model, long-run risk neutrality is not. Viewing long-run distributions of returns and comparing them vis-à-vis standard deviation, stochastic dominance, and other risk factors is a primary motivation of this work.

Since the expected values of the shocks are zero and assumed to be uncorrelated with the risk elements, and because the expected values of the risk elements are assumed to be uncorrelated, this simplifies to:

$$erev_{j,t} = E(price_{j,t}|price_{j,t-1}) \cdot E(yield_{j,t}|yield_{j,t-1}) + Cov(\varepsilon_{price,j,t}, \varepsilon_{yield,j,t}) \quad (11)$$

The covariance matrix used to generate the random normal shocks τ is not the same as the covariance of the converted shocks, ε. That is because the variances of the converted shocks are transformed to different distribution functions (such as lognormal or beta distributions) with different variances than the standard normal, and then shifted back to the appropriate range. Given the property:

$$Cov(cX, kY) = ck * Cov(X, Y) \quad (12)$$

Then

$$Cov(\varepsilon_{price,j,t}, e_{yield,j,t}) = sd(\varepsilon_{price,j,t}) \cdot sd(\varepsilon_{yield,j,t}) \cdot Cov(\tau_{price,j,t}, \tau_{yield,j,t}) \quad (13)$$

We can calculate the covariance of the ε's based on the covariance of the τ's, the known variance formulas of the new distribution functions, and the multiplicative shifts applied.

Consider next the crop rotation component of this problem. Crop rotation is a factor that must be considered, as certain crops can heavily utilize certain soil nutrients, so typically it is recommended not to utilize the same crop on the same site for more than 2 years in succession. We accommodate this concern if the same crop (variable $crop_t$ is the index of the selected crop in year t) has been used for the previous 2 successive years, by setting u to an arbitrarily large negative number. Otherwise, the value for u remains as-is.

$$crop_{t-2} == crop_{t-1} \rightarrow u_{crop,t} := -1 \times 10^9 \quad (14)$$

Combined, the farmer thus maximizes expected returns subject to crop rotation by selecting the crop with the u. The index, *crop*, of the selected crop for year t is the argument of the maximum of u.

Timber thinning (pine plantation and alley crop models)—Timber thinning occurs based on a fixed basal area rule, given in *BAdecSI*. The function harvest.m simulates timber thinning. When basal area (*ba*) of the stand, as output by pinegrowth.m to pine.m or alleycrop.m reaches a value in m^2/ha higher than that specified in *BAdecSI*, harvest.m will trigger a thinning. harvest.m will select the smallest trees to harvest until the remaining basal area is lower than *thinlimitSI*. The trees harvested will be set to zero in the matrix *live*, and the volume of sawtimber, chip-n-saw, and pulpwood harvested [determined by volume equations from Burkhart and others (2008)] will be recorded in the matrix *harv*.

Optimal year for timber harvest (pine plantation and alley crop models)—Selection of the optimal year for timber harvest presents a challenge. Three options were considered. First, we considered the Faustmann rule, under which harvest is undertaken once the annual change in value of the forest becomes less than the value that is lost by delaying future rotations by 1 year. However, the stochastic and discrete annual model we used is not optimal for this approach. The change in value of the forest can go above and below the decision threshold numerous times before and after reaching the true maximum value. Second, we considered a Bellman rule, under which harvest is undertaken if the present value is greater than all possible expected future values, given the present state of yield and price. The authors believe this optimization approach to be the most realistic, and is an area for future improvement of the model. However, estimating expected future value given the present state is a difficult computational problem and not attempted in this model. Third, we considered a simplifying assumption of perfect knowledge of future timber growth and prices, and

simply picking the year (up to the maximum age) that generates the maximum SEV. This model uses this third approach: it simply models the growth and price through the maximum age, *m*, and selects the year for harvest that has the highest calculated SEV.

Proportion of land dedicated to alleys (alley crop model)—In the alley crop model (alleycrop.m), crop profits per acre are multiplied by a factor representing the proportion of land dedicated to alleys.

Calculation and Output of the Financial Returns Indicators and Management Variables

Financial Returns Indicators—Formulas for the financial returns indicators SEV and AEI are given in the Model Overview section under Model Outputs. Since there are *n* iterations of the *m*-year time horizon, there are various ways to present these rich output data. The entire *n* values for NPV, SEV, and AEI are output in the first three columns of the output matrices MT, AT, and PT, where the M, A, and P represent the monocrop, alley crop, and pine plantation models, respectively (table 2). The simplest way to present the distribution is to output the mean and standard deviation of the financial returns, which are given in MSEV and MAEI, ASEV and AAEI, and PSEV and PAEI, where the M, A, and P represent the monocrop, alley crop, and pine plantation models, respectively.

Another, more detailed way to present this information is through the cumulative distribution functions of returns. This is particularly helpful in understanding the long-run financial risk of particular systems. The user can utilize MATLAB's built-in ecdf and plot functions to view and compare cumulative distribution functions of the land management alternatives after running the monocrop.m, alleycrop.m, and pine.m scripts.

Other returns-related variables are the mean and standard deviation of annual returns for each crop, for each year in the *m*-year time horizon. These can be obtained in the output matrix MRET and ARET (table 2). Average price, yield, cost, revenue, and returns over time for all crops are in the matrix MCROPS and ACROPS and can be plotted using MATLAB's built-in plot function.

Other Management Variables—The model also outputs variables related to management. These include the average number of years during the *m*-year time horizon that each crop is selected for planting (MNY, ANY), payments from government (MSUBS, ASUBS, PSUBS), and tree rotation in years (AROT, PROT) (table 2).

Detailed Model Options

Table 3 lists the three modeling options in ALLEY Model 2.0, which are described in more detail below.

Specialty Crops Option—Commodity crops are those which are easily produced, stored, and traded in large quantities. The examples given in this document include corn, soybeans, and cotton, but several other crops meet this definition. For our purposes, we define commodity crops as those which are likely to have detailed historical records about local yields, prices, and costs going back many years. We define "specialty crops" as non-commodities, that is, those which do not have historical data available.

Specialty crops are important to consider since they may provide local market opportunities that make monocrop or alley crop systems more economically feasible. For alley crop systems in particular, there may be specialty crops that perform better in the microenvironment within the

alleys. ALLEY Model 2.0 includes the option to simulate specialty crops, which is activated by entering 1 when prompted by the params.m script.

Since data are not available for specialty crops, the user will have to make assumptions about parameters related to those crops. Some of those parameters, such as the functional form of autoregression-trend function and the distribution function of the shocks for each returns element are assumed to be the same for all crops, regardless of whether they are commodities or specialty crops. However, the coefficients of those functions vary by crop.

When running the params.m script, if the user inputs 1 at the prompt, he/she will be guided through entering the relevant parameters. The user is given a default option to select the average of other crops, which may be a useful guide if very little is known about the specialty crop, or if it is simply hypothetical.

The first parameter requested by params.m will be the "correlation profile." The correlation profile can take any number. The correlation between the specialty crop's returns elements and the other crops will be an average of the correlation among the commodity crops, multiplied by the number in the correlation profile.

Competition Function Option—Although data and results are limited in number and scope, some experiments in alley crop systems in the Southern and Midwestern United States have shown that interspecific competition between annual crop components and relatively young trees aged 5–10 years can reduce crop yields by 25–50 percent (Gillespie and others 2000, Miller and Pallardy 2001, Wanvestraut and others 2004, Zamora and others 2009). ALLEY Model 2.0 includes the option to simulate this competition, which is activated by entering 1 when prompted by the params.m script.

We considered an option to model the loss in yield as a simple function of the age of the agroforestry system, but this would lose some of the information from the actual growth of the trees and the competition they produce. We chose to formulate yield as a function of length of live crown (*LLC*), a parameter that factors in both the number of trees and the growth of the trees, including the amount of live branches on each tree. *LLC*, which has been found to be correlated with forage production under loblolly pine in some silvopasture systems (Fassola and others 2005), is a parameter indicating the total length from the base of the crown to the top of the tree for all trees in a ha (measured in m/ha). The following are the competition functions, which generated a scalar value directly multiplied to the yield:

Competition profile function 1 (generally for most crops that thrive in full sun, e.g., cotton, corn):

$$1.0411 - \frac{1}{1 + e^{-0.0007 \cdot (LLC - 4500)}} \tag{15}$$

where

LLC = length of live crown

Competition profile function 2:

$$1 - \frac{1}{1 + e^{-0.0005\,(LLC - 6000)}} + 0.3 \cdot e^{-\frac{(LLC - 2000)^2}{2 \cdot (1200^2)}} \tag{16}$$

These functions are depicted in figure 5. Competition profile function 2 is assumed to increase in production with moderate competition before falling off as competition with trees becomes too intense. Some crops, such as wheat (Savin and Slafer 1991), annual hay crops, certain fruits and vegetables, etc., may be less adapted to hot microclimates and may perform reasonably well in partial shade.

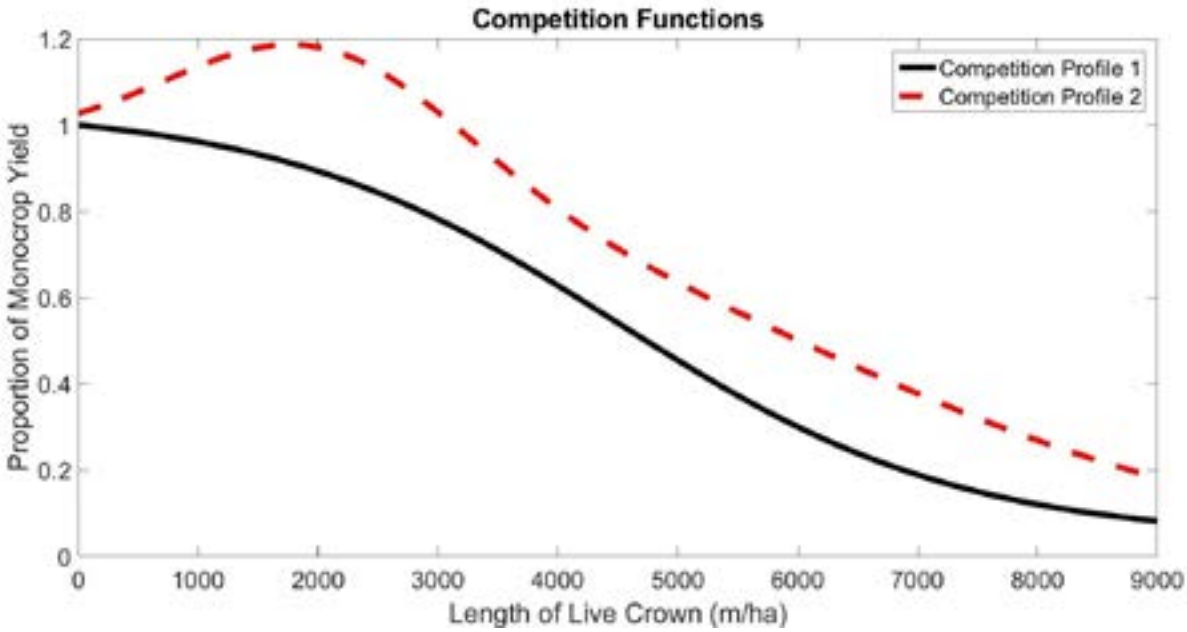

Figure 5—Competition functions presented as two alternative competition profiles of different types of annual crops. Competition profiles are given as relative production of an annual crop, as a proportion of monocrop yield, under variable length of live crown in an alley crop system.

Policy Option—Activating the policy option provides Farm Bill Agricultural Risk Coverage (ARC) payments for annual crops and cost-share payments for pine establishment.

Farm Bill ARC payments to farmers (monocrop and alley crop models)—In the United States, payments from government to farmers can make up an important share of overall revenues. Currently, the most prominent farm payments are modeled as a risk management tool, making payments to farmers in years when returns are below normal. The monocrop and alley crop models include a policy option to simulate payments from the Agricultural Risk Coverage – Individual Coverage (ARC-IC) program,[10] authorized in the 2014 Farm Bill (FSA 2014). The ARC payment option can be activated in the model by entering 1 at the relevant point when running the params.m script.

Under the ARC-IC program, a payment is made if the revenue generated from a specific crop is less than the *guarantee*—86 percent of the most recent 5-year Olympic average[11] of revenue for that specific crop for that particular farmer. In this circumstance, the ARC program will pay 65 percent of the difference between the actual revenue and the guarantee.

$$rev_{j,t} < guarantee_{jt} \rightarrow subs_{jt} = .65 \cdot (guarantee_{jt} - rev_{j,t}) \quad (17)$$

$$rev_{j,t} \geq guarantee_{jt} \rightarrow subs_{jt} = 0 \quad (18)$$

$$guarantee_{jt} = 0.86\left[\sum_{q=t-5}^{t-1}(rev_{jq}) - \max_{q}(rev_{jq}) - \min_{q}(rev_{j,q})\right]/3 \quad (19)$$

$$rev_{j,t} < guarantee_{jt} \rightarrow subs = 0.65(guarantee_{jt} - rev_{j,t}) \quad (20)$$

$$rev_{j,t} \geq guarantee_{jt} \rightarrow subs = 0 \quad (21)$$

[10] Here we use the ARC-IC program as a stand-in for potential payments from government sources under the Farm Bill, which also include the options of enrolling in ARC-County (ARC-CO) and Price Loss Coverage (PLC) programs. In reality, most soybean and corn producers utilize the ARC-CO option. However, modeling ARC-CO payments would mean separately modeling the average county yields.

[11] The Olympic average is the average of the three middle values of a group of five values; that is, the highest and lowest values are discarded.

where

$guarantee_{j,t}$ = 86 percent of the revenue Olympic average for years *t*-5 to *t*-1 for crop *j*

Likewise, the farmer will factor government payments into the decision rule about which crop to pick.

$$esubs_{j,t} = E\left(subs_{j,t} \middle| erev_{j,t}, guarantee_{j,t}\right) \tag{22}$$

$$erev_{j,t} < guarantee_{j,t} \rightarrow esubs = 0.65\left(guarantee_{j,t} - erev_{j,t}\right) \tag{23}$$

$$erev_{j,t} \geq guarantee_{j,t} \rightarrow esubs = 0 \tag{24}$$

Cost-share payments for pine establishment (pine plantation and alley crop models)—Many States, and some Federal programs, provide payments to compensate landowners for a portion of the expense related to establishing a new stand of trees. The pine plantation and alley crop models include an option to simulate cost-share payments for pine establishment, by setting the parameter *policy* to 1. The payment is equal to 50 percent of the establishment costs.

Example Application

As an example for the user's reference, we used ALLEY Model 2.0 and the sample dataset provided to generate financial returns distributions for alley crop, monocrop, and pine plantation, with and without the assumption of competition between trees and annual crops in the alley crop system (competition function option). In this simple example, we do not utilize the specialty crops or policy options. We do not provide a comprehensive review of all the possible information and results that could be gleaned from ALLEY Model 2.0 here but rather a brief sampling of some of the outputs.

Data and Method

The default parameters are based on Halifax County, NC, as an example. We chose this county because it has a relatively diversified agricultural sector. It is also within the area where loblolly pine plantations are common.

We estimated default crop parameters based on historic output price and yield data from Quick Stats (NASS 2016) and input costs from Commodity Costs and Returns (ERS 2016). We used three annual crops that are relatively common in Halifax County: 1) corn, 2) soybeans, and 3) cotton. We estimated timber price parameters from the NC Cooperative Extension's Price Reports (NCCE 2014) and pine plantation costs from Dooley and Barlow (2013) and NC Forest Service's Prevailing Rates Report (NC Forest Service 2014). Timber volume growth and yield are based on Burkhart and others (2008) and Westfall and others (2004).

These data have missing values for a few years. When this occurs, the estimate.m script prompts the user to input substitute values or, by default, replace the missing values with averages. For this example, we used the default method.

We ran the full ALLEY Model 2.0 suite two times, first without and then with the competition function option. Since the competition function only affects the alley crop estimates [function compfunc.m is called exclusively by alleycrop.m (table 4)], the outputs for monocrop and pine will be the same for these two cases. We used the default values of the parameters as described in tables 1 and A.1–A.3 (appendix), which are pre-coded into the estimate.m and params.m scripts.

For the competition function option, params.m prompts the user to input the competition profile for each crop. We chose competition profile 1 for the first and third crops (corn and cotton), and competition profile 2 for the second crop (soybeans).

Results

Using the default value of n = 10,000 iterations, running on a computer with Microsoft® Windows 10, a 64-bit operating system, 16.0 GB RAM, and an Intel® CoreTM i5 processor, the scripts took the following amounts of time to run: monocrop.m: approximately 20–30 seconds; pine.m and alleycrop.m: approximately 20 minutes, depending on what other programs or system processes were running meanwhile. These times could be reduced with a more powerful processor or by limiting the number of iterations. With n = 10,000 iterations as chosen here, results appear fairly consistent and stable. Empirical cumulative distribution functions (CDFs) are smooth at reasonable resolution, and the results are relatively stable.

The competition function only affects interspecific competition within the alley crop system (intraspecific competition amongst pines or amongst annual crop plants is not turned off with this option), so results for the monocrop and pine systems were the same for the two model runs within about 1 percent error (smaller for certain outputs) due to the random nature of the Monte Carlo simulations. In both simulations, the monocrop system mean SEV was approximately US$ 5,500/ha with a standard deviation of US$ 5,000/ha. The pine system had a lower mean SEV of approximately US$ 1,900/ha with a standard deviation of US$ 800/ha. When assuming no competition between the pine and annual crops (competition function option off), the mean SEV in the alley crop system was approximately US$ 6,700/ha with a standard deviation of US$ 4,700/ha. Therefore, when assuming no competition between trees and annual crops, the alley crop system financially outperforms monocrop by being more profitable and less risky. It is more profitable than pine but also more risky. The fact that alley crop has higher average returns than monocrop is because there are fewer trees than the pine plantation so individual trees perform much better than in the pine plantation. However, when assuming interspecific competition between trees and annual crops, SEV in the alley crop system is reduced to a mean of US$ 3,400/ha with a standard deviation of US$ 4,600/ha. This indicates a wide range of possible outcomes for farms.

The empirical CDFs describe the distribution of the SEVs for the three different land uses in much greater detail (fig. 6). The alley crop CDF assuming no competition is always to the right of the monocrop CDF. That means that at every percentile, alley crop SEVs are higher, i.e., alley crop first-order stochastically dominates monocrop. A first-order stochastically dominant distribution is preferred by all individuals who prefer more to less, regardless of risk aversion/neutrality/proneness. That is, any land manager should prefer a distribution that is first-order dominant, assuming that there are no other preferences that are not included in the estimate of profits (for example, environmental stewardship). Compared to pine plantation, both alley crop and monocrop have higher mean and median SEVs, but pine is less risky and has higher SEVs at the bottom tail of the

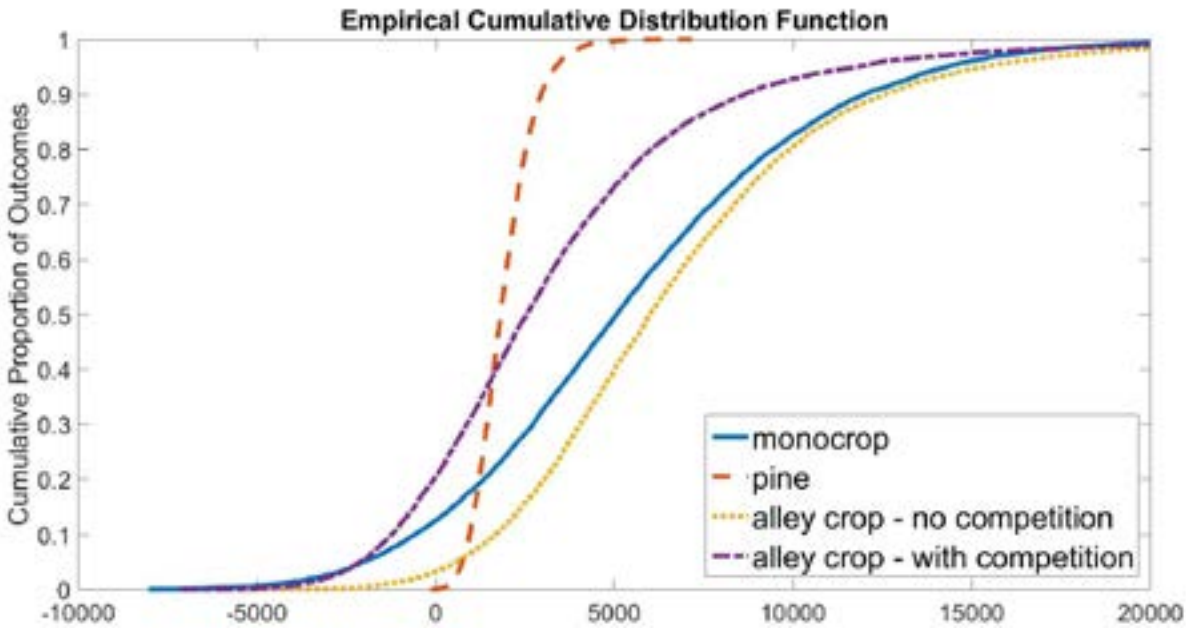

Figure 6—Empirical cumulative distribution functions of soil expectation value under three alternative land uses—monocrop, pine, and alley crop—with differing assumptions about competition between annual crops and trees in the alley crop system.

CDF. Therefore, neither alley crop nor monocrop is first-order or second-order dominant over pine plantation. It is possible that some risk-averse land managers could prefer pine plantation despite the lower average returns.

Interestingly, even when assuming interspecific competition in an alley crop system, monocrop does not first- or second-order dominate alley cropping. Alley crop systems are less risky (lower variance) than monocrop, and the CDF graph shows that alley crop has higher SEVs at the upper and lower tails of the distribution. This means that some risk-averse land managers may prefer alley crop to monocrop systems. The same is true of pine plantations. Alley crop (with interspecific competition) seems to have both returns and risk that are intermediate between monocrop and pine.

It is also interesting to note how alley crop systems could change the way system components are optimally managed. In the monocrop system, corn was chosen for production 21 percent of the time, soybeans 39 percent, and cotton 40 percent. When interspecific competition is included in the alley crop system, this changes to 23, 56, and 21 percent, respectively. This is primarily due to the fact that soybean was assumed to perform better in partial shade than corn and cotton, and was thus assigned competition profile 2 instead of competition profile 1.

In pine plantation systems, the mean optimal rotation length was estimated to be 27.9 years (standard deviation of 6.7 years). Assuming no interspecific competition, pine rotation in the alley crop system was shorter with a mean of 23.8 years (standard deviation 9.7 years). This is because individual trees grow faster and are mature earlier in alley crop settings. When including an assumption of interspecific competition, the alley crop optimal pine rotation is shorter with a mean of 21.1 years (standard deviation 11.3 years). That is because it makes sense to harvest the pines sooner to return to a situation with less shade for better annual crop production. One could presume (although not estimated here) that if all the annual crops had competition profile 1, which always has lower production under any partial shade and drops off more quickly than profile 2, then the optimal alley crop pine rotation would be even shorter.

Areas for Future Improvements to the Model

On its surface, ALLEY Model 2.0 is complex. At the heart of the model, however, is a rather simple and elegant core—drawing random samples from a joint multivariate distribution. These values are then inserted into various real-world decision rules.

There are several areas in which this model can be improved. Each area will take time and resources. It is our hope to address them in future versions of the model. These may include the following:

- Make the model more interactive and user friendly for data input, option selection, and output viewing. Create dialogue boxes for inputs.
- Simplify the model generally, including modeling processes and language, for users with little or no computer language or statistical expertise.
- Allow more flexibility in functional form of the probability distribution functions.
- Consider a bootstrapping/non-parametric approach to construction of probability distribution functions.
- Allow more flexibility in setting short-run risk preferences (risk averse vs. neutral).
- Allow for hedging of short-run risk with futures markets and contracts for purchase.
- Allow multiple fields on a farm.
- Allow perennial alley crops such as fruit crops (e.g., blueberries) or forage crops (e.g., for hay production or conversion of alley crop system to silvopasture).
- Allow the farmer to decide to fallow a field if expected returns are too low.
- Improve the pine rotation decision rule to pick the harvest year with uncertain information (the main difficulty here is computational power needed to update future expected values for entire future period, each year).
- Allow modelling of crops that generate multiple outputs.
- Allow the option of pruning or otherwise managing trees to reduce competition.

The authors welcome feedback from users and other ideas to improve ALLEY Model 2.0. Please send an email to gregory.e.frey@usda.gov telling us about the purpose for using the model, its usefulness, and suggestions for improvement.

Conclusion

This report describes a suite of computer scripts and functions called ALLEY Model 2.0 for use in MATLAB, which models stochastic prices, yields, and costs for three alternative land uses: monocrop, alley crop, and pine plantation. The model can be run numerous times over a given time horizon, to estimate expected values and distributions of potential financial returns. The model can generate parameter estimates based on user-provided historical data on crop yields and timber prices, and example data are provided from Halifax County, NC. Timber growth and yield were simulated using biometric equations from past research on loblolly pine in the U.S. Southeast. Other parameters, such as catastrophic probabilities and competition between trees and annual row crops in alleys, were based on literature and authors' estimates.

ALLEY Model 2.0 will be used for research on agroforestry decision making, testing of crop characteristics that improve returns in agroforestry settings, evaluation of current policies and future policy options, and assessment of potential changes in returns do to the effects of a changing climate.

Literature Cited

Burkhart, H.E.; Amateis, R.L.; Westfall, J.A.; Daniels, R.F. 2008. PTAEDA4.0: **Simulation of individual tree growth, stand development and economic evaluation in loblolly pine plantations**. Blacksburg, VA: Virginia Tech.

Cary, M.A.; Frey, G.E.; Mercer, D.E. 2014. **The value of versatile alley cropping in the southeast US: A Monte Carlo simulation**. In: Aguilar, F.X.; Lantiainen, S., eds. Forest economics: gateway to sustainability; inaugural symposium of the International Society of Forest Resource Economics. St. Louis, MO: 6–11.

Cubbage, F.W.; Davis, R.R.; Frey, G.E. [and others]. 2015. **Financial and economic evaluation guidelines for international forestry projects**. In: Pancel, L.; Kohl, M., eds. Tropical forestry handbook. Berlin: Springer Berlin Heidelberg.

Dooley, E.; Barlow, R. 2013. **Costs and trends of southern forestry practices 2012**. FOR-2001. Alabama Cooperative Extension System: Alabama A&M and Auburn Universities. 8 p.

ERS. 2016. **Commodity costs and returns.** Washington, DC: U.S. Department of Agriculture, Economic Research Service. http://www.ers.usda.gov/data-products/commodity-costs-and-returns.aspx. [Date accessed: May 23, 2016].

Fassola, H.E.; Lacorte, S.M.; Pachas, A.N.; Pezzuti, R. 2005. **Factores que influencian la producción de biomasa forrajera de Axonopus jesuiticus valls, bajo dosel de *Pinus taeda* L. en el nordeste de corrientes**. La Revista de Investigaciones Agropecuarias (RIA). 34(3): 21–38.

Frees, E.W.; Valdez, E.A. 1998. **Understanding relationships using copulas**. North American Actuarial Journal. 2(1): 1–25.

FSA. 2014. **Base acre reallocation, yield updates, Agriculture Risk Coverage (ARC) & Price Loss Coverage (PLC). 2014 Farm Bill Fact Sheet**. Washington, DC: U.S. Department of Agriculture, Farm Service Agency. 11 p. http://www.fsa.usda.gov/Internet/FSA_File/base_acre_reallocate_arc_plc.pdf. [Date accessed: May 23, 2016].

Gillespie, A.; Jose, S.; Mengel, D. [and others]. 2000. **Defining competition vectors in a temperate alley cropping system in the midwestern USA: 1.** Production physiology. Agroforestry Systems. 48(1): 25–40.

Goodwin, B.K.; Ker, A.P. 2002. **Modeling price and yield risk**. In: Just, R.E.; Pope, R.D., eds. A comprehensive assessment of the role of risk in U.S. agriculture. Norwell, MA: Kluwer Academic Publishers: 289–323.

LADAF. 2016. **Quarterly report of forest products.** Louisiana Department of Agriculture & Forestry. http://www.ldaf.state.la.us/forestry/reports/. [Date accessed: May 23, 2016].

Miller, A.W.; Pallardy, S.G. 2001. **Resource competition across the crop-tree interface in a maize-silver maple temperate alley cropping stand in Missouri**. Agroforestry Systems. 53(3): 247–259.

NASS. 2016. **Quick stats.** Wasington, DC: U.S. Department of Agriculture, National Agricultural Statistics Service. http://quickstats.nass.usda.gov/. [Date accessed: May 23, 2016].

NC Forest Service. 2014. **Prevailing rates for sub-practices: FY2014–FY2015. Raleigh, NC: NCFS Forest Development Program**. http://ncforestservice.gov/Managing_your_forest/pdf/fdp_prevailing_rates.pdf. [Date accessed: May 23, 2016].

NCCE. 2014. **Price report: historic price information**. Original data compiled from TimberMart-South and other sources. Raleigh, NC: North Carolina Cooperative Extension Forestry. http://www.ces.ncsu.edu/forestry/resources/price_report.php. [Date accessed: July 24, 2014].

Savin, R.; Slafer, G.A. 1991. **Shading effects on the yield of an Argentinian wheat cultivar**. The Journal of Agricultural Science. 116(01): 1–7.

Wanvestraut, R.H.; Jose, S.; Nair, P.K.R.; Brecke, B.J. 2004. **Competition for water in a pecan (*Carya illinoensis* K. Koch)–cotton (*Gossypium hirsutum* L.) alley cropping system in the southern United States**. Agroforestry Systems. 60(2): 167–179.

Westfall, J.A.; Burkhart, H.E.; Allen, H.L. 2004. **Young stand growth modeling for intensively-managed loblolly pine plantations in southeastern U.S.** Forest Science. 50(6): 823–835.

Workman, S.W.; Bannister, M.E.; Nair, P.K.R. 2003. **Agroforestry potential in the southeastern United States: perceptions of landowners and extension professionals**. Agroforestry Systems. 59(1): 73–83. doi: 10.1023/A:1026193204801.

Zamora, D.S.; Jose, S.; Napolitano, K. 2009. **Competition for 15 N labeled nitrogen in a loblolly pine–cotton alley cropping system in the southeastern United States**. Agriculture, Ecosystems & Environment. 131(1): 40–50.

Appendix: Additional Parameter Lists

Table A.1—Parameters used in the model options

Parameter	Format	Description	Units	Default value
corrprofile	Vector	Correlation profile. Value multiplied by the average of the correlation among the commodity crops, to determine a value for the correlation between a specialty crop and other crops.	Unitless	No default value
compprofile	Vector	Competition profile. Set to 1 or 2 to model interspecific completion between trees and annual crop.	Unitless	No default value

Table A.2—Tree (loblolly pine) component parameters

Parameter	Format	Description	Units	Default value
pdSI	1 x 2 vector	Planting distance between trees	m	[2.44 3.05]
siSI	Scalar	Site index at age 25	m	22.86
phdwd	Scalar	Percent of basal area in hardwood species	%	0
pcatprob	Scalar	Probability of a catastrophic forestry event occurring on a parcel in any given year	Unitless	0.01
pcateffect	1 x 2 vector	Mean and standard deviation of proportion of trees killed when a catastrophic forestry event occurs	Unitless	[0.15 0.15]
minage	Scalar	Minimum age for pine to be harvested	Years	8
BAdecSI	Scalar	Decision point at which a forest thinning occurs in the model, in terms of a maximum basal area allowed before thinning	m^2/ha	25.2
thinlimitSI	Scalar	When a thinning occurs, the reduction of basal area from *BAdecSI*	m^2/ha	16.0
sawminSI	Scalar	Minimum diameter at breast height (DBH, diameter at 1.4 m/4.5 feet above ground level) for a tree to be utilized for sawtimber	m	0.30
sawtopSI	Scalar	Minimum small-end diameter that can be used for sawtimber	m	0.20
chipminSI	Scalar	Minimum DBH for a tree to be utilized for chip-n-saw	m	0.20
chiptopSI	Scalar	Minimum small-end diameter that can be used for chip-n-saw	m	0.15
pulpminSI	Scalar	Minimum DBH for a tree to be utilized for pulpwood	m	0.13
pulptopSI	Scalar	Minimum small-end diameter that can be used for pulpwood	m	0.10
pestab	Scalar	Mean establishment cost of pine plantation	$/ha	-780
prelease	Scalar	Mean cost of mid-rotation release	$/ha	-250
pman	Scalar	Mean cost of annual management, taxes, etc.	$/ha	-25

Table A.3—Alley crop-specific parameters

Parameter	Format	Description	Units	Default value
afpdSI	1 x 2 vector	Planting distance between trees within tree rows	m	[2.44 2.44]
rows	Scalar	Number of rows of trees	Unitless	2
alleySI	Scalar	Width of the alley; must be a multiple of the second element of *afpdSI*	m	12.2
al	Scalar	Equivalent number of rows of trees before pattern repeats *alleySI*/*afpdSI*(2)+rows-1	Unitless	6
apestab	Scalar	Mean establishment cost of pines in alley crop 1.3*rows/[*alleySI*/*afpdSI*(2)+rows-1]**pestab*	$/ha	-644.79
aprelease	Scalar	Mean cost of mid-rotation release 1.3*rows/[*alleySI*/*afpdSI*(2)+rows-1]**pestab*	$/ha	-208.33
apman	Scalar	Mean cost of annual management, taxes, etc. for tree component of the alley crop *pman**rows/al*alleydiff =	$/ha	-20.83

Frey, Gregory E.; Cary, Michael A.; Goodwin, Barry K.; Mercer, D. Evan. 2018. Agroforestry land-use economic yield and risk (ALLEY) model 2.0: a computer suite to simulate and compare stochastic yield and returns of alley crop, monocrop, and pine plantation systems in the U.S. South. e-Gen. Tech. Rep. SRS-235. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. 35 p.

Alley crop systems, which combine field crops in wide alleys between rows of trees, have been presented as a sustainable land use that can generate economic benefits and ecosystem services. We created models of stochastic processes that simulate yields, prices, and costs for an alley crop system and two competing land uses: monocrop and pine plantation. The suite of models, called Agroforestry Land-use Economic Yield and Risk (ALLEY) Model 2.0, uses a Monte Carlo approach, iterating numerous times over a given time horizon, to estimate and compare expected values and distributions of potential financial returns from the three land uses with selected management variables. ALLEY Model 2.0 includes sample historical data on crop yields in Halifax County, NC, crop and timber prices in North Carolina, and costs from the U.S. Southeast. Timber growth and yield were simulated using biometric equations from past research on loblolly pine in the U.S. Southeast. Other parameters with little historical data, such as competition between trees and annual row crops in alleys, were based on literature review and authors' estimates. ALLEY Model 2.0 may be used for research on agroforestry decision making, testing of crop characteristics that improve returns in agroforestry settings, evaluation of current policies and future policy options, or assessment of potential changes in returns due to the effects of a changing climate. An example application is provided. The ALLEY Model 2.0 software suite and sample data file can be accessed at https://www.srs.fs.usda.gov/pubs/gtr/gtr_srs235/.

KEYWORDS: Agroforestry, Gaussian copula, Monte Carlo method, risk and uncertainty, stochastic modeling.

In accordance with Federal civil rights law and U.S. Department of Agriculture (USDA) civil rights regulations and policies, the USDA, its Agencies, offices, and employees, and institutions participating in or administering USDA programs are prohibited from discriminating based on race, color, national origin, religion, sex, gender identity (including gender expression), sexual orientation, disability, age, marital status, family/parental status, income derived from a public assistance program, political beliefs, or reprisal or retaliation for prior civil rights activity, in any program or activity conducted or funded by USDA (not all bases apply to all programs). Remedies and complaint filing deadlines vary by program or incident.

Persons with disabilities who require alternative means of communication for program information (e.g., Braille, large print, audiotape, American Sign Language, etc.) should contact the responsible Agency or USDA's TARGET Center at (202) 720-2600 (voice and TTY) or contact USDA through the Federal Relay Service at (800) 877-8339. Additionally, program information may be made available in languages other than English.

To file a program discrimination complaint, complete the USDA Program Discrimination Complaint Form, AD-3027, found online at http://www.ascr.usda.gov/complaint_filing_cust.html and at any USDA office or write a letter addressed to USDA and provide in the letter all of the information requested in the form. To request a copy of the complaint form, call (866) 632-9992. Submit your completed form or letter to USDA by: (1) mail: U.S. Department of Agriculture, Office of the Assistant Secretary for Civil Rights, 1400 Independence Avenue, SW, Washington, D.C. 20250-9410; (2) fax: (202) 690-7442; or (3) email: program.intake@usda.gov.

USDA is an equal opportunity provider, employer, and lender.