

**A GENERALIZED METHODOLOGY FOR DEVELOPING
FLEXIBLE WHOLE STAND SURVIVAL MODELS**

Plantation Management Research Cooperative

Daniel B. Warnell School of Forest Resources

University of Georgia

PMRC Technical Report 2002 - 2

April 6, 2002

Abstract

Survival is generally the least understood and hardest to model within a whole-stand forestry growth and yield system. The largest source of variability in yield predictions is the estimation of future surviving trees per acre. Previous whole-stand survival modeling efforts have been based upon differential equation solutions of changes in survival rates through time. Hazard functions, a survival analysis technique commonly used in fields such as medicine and sociology, can be applicable to plantation survival estimation. We offer a generalized method for deriving whole-stand survival models that are capable of modeling complex underlying hazard curves. We used our knowledge of the empirical hazard curve to limit our selection to appropriate functions and the integration of the final selected function results in an initial condition difference equation. Which, when fitted to our data, provides flexible whole-stand survival solutions and adequately represents the underlying hazard curve. Our method is relatively easy to implement and can model a whole-stand survival curve that has a complex underlying hazard function.

INTRODUCTION

Many forest growth and yield systems are inherently reliant on mortality prediction, yet it is generally the least understood and has the most variability of all the forest growth and yield models (Hamilton 1986). Furthermore, it is critical in accurately predicting growth and yield for plantations because of the sensitivity of the basal area growth model to the underlying mortality. Forest mortality has been usually classified as either natural or irregular (Staebler 1953). Natural mortality occurs because of competition for light, nutrients, and water and irregular mortality is caused by insects, diseases, and other catastrophic events. Our study focuses on plantation whole stand natural forest survival.

Forest survival, or complementary, mortality, can be analyzed using traditional survival analysis. The distribution of the random variable T from initialization to the event of interest can be characterized by the survival and hazard functions. Let the distribution of T be $F(t) = \Pr[T \leq t]$, then $f(t)$ is the corresponding density function and the survival function is defined as $S(t) = \Pr(T > t) = 1 - F(t)$. The hazard function, which is the instantaneous rate of mortality assuming the individual has survived to time t and has the dimensional quantity of number of events per interval of time (Collet 1994), is defined as $h(t) = f(t)/1 - F(t)$. Analogous to the continuous time hazard function, the discrete time hazard function is defined as $q(t) = [F(t + 1) - F(t)]/1 - F(t)$ (Wilson 1972). A nonparametric estimation of the survival function is the Kaplan-Meier product limit estimator (KM) (Kaplan and Meier 1958), which is defined as

$$\hat{S}(t) = \prod_{i:t_i \leq t} \left[1 - \frac{d_i}{n_i} \right]$$

for $t_l \leq t \leq t_j$. Where n_i and d_i are the subjects at risk (n_i) and that die (d_i) at time t_i . An empirical estimate of the hazard function is simply the number of trees that died in the interval divided by the total length of time all the trees were observed. Survival curves are non-increasing over time, whereas the hazard function can increase, decrease, remain constant, or assume a combination of these shapes.

A common cumulative distribution function (*cdf*) used in survival analysis is the Weibull because it is capable of describing the three most common types of hazard curves, which are monotonically increasing or decreasing, and constant. However, the Weibull distribution imposes strong restrictions on the data and is unable to model complex hazard shapes such as a bathtub shape (Hjorth 1980). Increasing monotonic hazard functions are the most common because many studies focus on a snapshot of the subject's lifespan in which gradual aging takes place (Lawless 1984). Because of the snapshot focus, there is usually no empirical motivation to find distributions that are capable of producing bathtub shaped hazard functions. However, a bathtub shaped hazard function is a reasonable assumption when viewing the entire lifespan of many, if not most, biological organisms. Consider the human lifespan. There are typically three distinct phases: Infant (mortality decreases), juvenile to adulthood (mortality generally stable), and mature adults (mortality increases). Taking a snapshot from any of these phases will likely result in a specific type of survival and hazard curve, but when viewed as an entity it is reasonable to assume a bathtub shaped curve. In forestry, a study by Lorimer and Frelich (1984) illustrated that the diameter-specific mortality rates for a given stand could be bathtub shaped. Moreover, mortality has been modeled as bathtub shaped with respect to diameter in individual-tree models (e.g., Buchman et al. 1983,

Monserud and Sterba 1999), but mortality has not been explicitly modeled as bathtub shaped for whole-stand models.

Whole-stand survival models have commonly been developed using a derivative of the generalized Gamma *cdf* (e.g., Weibull and exponential *cdf*'s are special cases) or the difference equation approach. Both of these approaches, either implied or explicitly stated, use presuppositions about the relative rate of instantaneous mortality that is based upon empirical evidence. The Weibull *cdf* is a flexible distribution that has been widely used for whole-stand survival models (e.g., Pinder et al. 1978, Glover and Hool 1979, Somers et al. 1980, Pienaar and Shiver 1981, Belli and Ek 1988, Amateis et al. 1997). As discussed, the Weibull *cdf* hazard function is capable of describing the three most common hazard shapes, hence, its popularity. Oftentimes the hazard function shapes that are capable when using the Weibull distribution are applicable since the study data, in which plots are often established after the initial seedling mortality, only contain a snapshot of a stand's lifespan. The hazard function for the two-parameter Weibull *cdf* survival model is $h(t) = \frac{c}{b} \left(\frac{t}{b} \right)^c$, where t is time and b, c are parameters. Cumulative distribution based whole-stand survival models use suppositions with respect to the distribution's ability to model the empirical survival trends. Conversely, difference equation models use suppositions about the relative rate of instantaneous mortality change. Although it appears these two approaches are distinct, oftentimes the difference equation supposition leads, after integration, to a *cdf* based whole-stand survival model. This is illustrated by using a simple difference equation for the relative rate of mortality, which assumes that the instantaneous mortality rate is constant, i.e.,

$$\frac{1}{N} \frac{dN}{dA} = \beta.$$

Where N is the number of trees per unit area, A is age, and β is a parameter. After integration and using the initial conditions that when $A_2 = A_1$ then $N_2 = N_1$, the result is

$$\frac{N_2}{N_1} = S(A_2) = e^{\beta(A_2 - A_1)}.$$

Thus, this difference equation results in the exponential distribution and implies that the instantaneous mortality rate is constant for all ages, densities, and site indices (Clutter et al. 1983). Exponential *cdf*'s have been used in several whole-stand survival studies (e.g., Martin's et al. 1999, Devine and Clutter 1985), however it imposes a strong assumption of a constant hazard rate. Nonetheless, the constant hazard rate is oftentimes reasonable because of the age range for the study data. For example, Devine and Clutter (1985) used survival data from 161 plots, of which only two had measurement data less than five-years of age. Hence, we wouldn't expect to detect early stand survival trends. Clutter and Jones (1980) presented a more flexible difference equation in which they assumed that the relative rate of instantaneous mortality is proportional to age and initial trees per acre, which are raised to a power, i.e.,

$$\frac{1}{N} \frac{dN}{dA} = \alpha A^\delta N^\phi.$$

Integrating and using the initial condition that when $A_2 = A_1$, then $N_2 = N_1$ yields

$N_2 = \left[N_1^\beta + \eta (A_2^\phi - A_1^\phi) \right]^{\frac{1}{\beta}}$. This flexible whole-stand survival model has subsequently been used with slight modifications in several other studies (e.g., Martin and Brister 1999 Pienaar and Rheney 1993).

It has become common to model whole-stand mortality for a stand's lifespan using a system of equations (e.g., Matney and Farrar 1992, Amateis et al. 1997). These systems disaggregate the lifespan of a stand's survival into distinct phases, typically some combination of the seedling, juvenile, adult, and mature phases. Disaggregating whole-stand survival into phases is primarily conducted because of the difficulty in developing a flexible biologically reasonable function that can model survival throughout stand development and data limitations.

We demonstrate a method for deriving flexible biologically reasonable whole-stand survival models, which are capable of modeling complex underlying hazard functions. Furthermore, it is hypothesized, for our data, that the continuum of whole-stand forestry survival has an underlying bathtub shaped hazard function.

DATA

Data were obtained from the Consortium for Accelerated Pine Production Studies (CAPPS), which is overseen by the Warnell School of Forestry at the University of Georgia. CAPPS was initiated to investigate the effects of intensive forest management on the productivity of loblolly pine plantations in the Southeastern United States.

Loblolly pine plantations were established throughout Georgia at Athens, Dawsonville, Eatonton, Thompson, Tifton, and Waycross. The study called for two complete blocks to be established at each location with each block containing four 0.375-acre treatment plots. A 0.375-acre treatment plot was established at each location using bare-root seedlings on an eight by eight spacing. A 0.125-acre measurement plot was centered

within each of the treatment plots. Each of four cultural treatments was randomly assigned to the blocks at each location. These cultural treatments are

- 1) Herbicide: spray plot with non-soil active herbicide as needed to maintain complete control of woody and herbaceous vegetation,
- 2) Fertilization: apply recommended rates of fertilizer annually, if necessary, to ensure that nutrients are not the limiting factor,
- 3) Herbicide – Fertilization: apply both herbicide and fertilization treatments, and
- 4) Control treatment: no cultural treatment other than mechanical site preparation.

The original study called for a replication of all treatment plots every two years for the first ten years of the study. This protocol would have resulted in five complete sets of experimental plots at all installations, where the plots have a staggered establishment time. Hence, this study has a $2^2 \times 5$ factorial experimental design, i.e., each location has two treatments with two levels (herbicide versus no herbicide, fertilization versus no fertilization) and the stand replication treatment at five levels. The actual study varies from the protocol because of funding limitations and the replications have been repeated at different intervals for different locations. The plots have been measured annually with available survival data beginning at age two. Plot survival data are summarized by plot age and treatment (Table 1).

Table 1. The CAPPS study plot distribution by year planted for the spectrum of plots and by treatment.

Year Planted	Plot Age	Plots	Plots by Treatment			
			Control	Fertilized	Herbicide	HF*
1986	14	26	8	4	8	6
1987	13	28	8	6	8	6
1988	12	36	10	8	10	8
1989	11	24	8	4	8	4
1992	8	20	6	4	6	4
1994	6	12	4	2	4	2
Total		146	44	28	44	30

* HF is the fertilizer and herbicide treatment.

MODEL DEVELOPMENT

To develop whole-stand survival models, which are capable of reflecting complex underlying hazard functions, we first computed the KM survival estimates and the corresponding discrete empirical hazard function (Table 2).

Table 2. The Kaplan-Meier product limit estimator for the interval survival, cumulative survival and the hazard for the CAPPS study data.

Age	N	Alive	Dead	Censored	Survival	Hazard
2	11956	11425	532	0	0.9556	0.0444
3	11424	11269	156	4	0.9426	0.0136
4	11264	11193	72	1	0.9367	0.0063
5	11191	11134	62	3	0.9319	0.0051
6	11126	11061	70	950	0.9265	0.0058
7	10106	10034	76	15	0.9199	0.0071
8	10015	9940	79	1594	0.9130	0.0075
9	8342	8274	72	73	0.9055	0.0082
10	8197	8115	86	868	0.8965	0.0100
11	7243	7171	76	1341	0.8876	0.0099
12	5826	5745	81	1899	0.8752	0.0139
13	3846	3772	74	1943	0.8584	0.0192
14	1829	1788	41		0.8391	0.0224

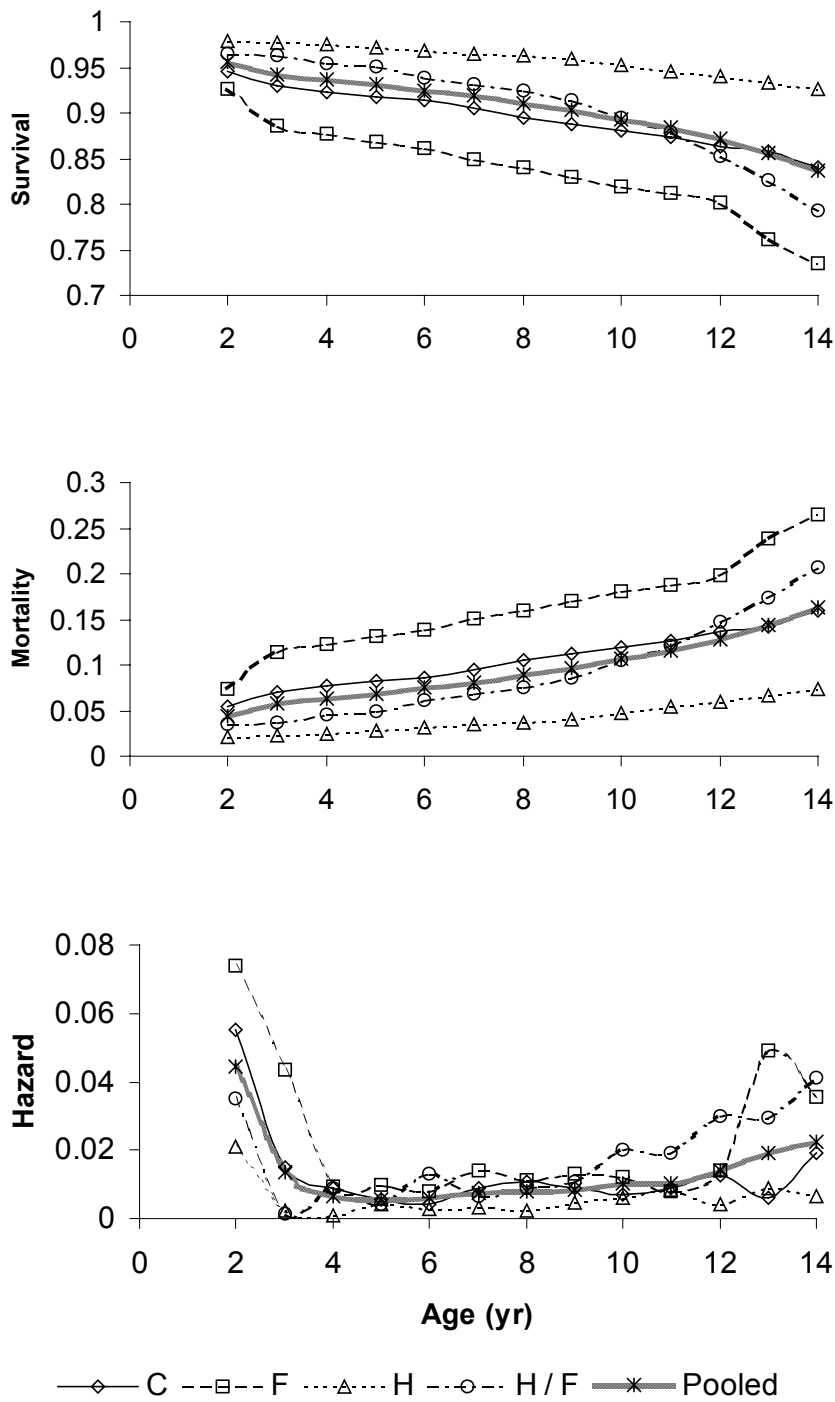


Figure 1. The CAPPS study empirical survival, mortality, and hazard functions by treatment (C = control, F = fertilized, H = herbicide, and HF = herbicide and fertilized) and for the spectrum of plots (pooled).

The KM survival estimates for the spectrum of plots illustrates that the underlying discrete hazard decreases from age 2-5 and then increases from ages 5-14. Kaplan-Meier survival and mortality estimates, and the corresponding discrete hazard function were computed by treatment (Figure 1). Survival curves by treatment illustrate that the herbicide and fertilizer treatments result in the most favorable and unfavorable survival. In addition, the herbicide/fertilizer treatment has favorable early survival but mortality increases rapidly after about age eight. The hazard function trend for the spectrum of plots appears to be bathtub shaped, but the oldest plots are 14 years and it is difficult to infer the future trend of the hazard function. Nevertheless, it is reasonable to assume that the hazard function will continue to increase with time, which is consistent with most biological populations (Pinder 1978). We assumed that whole-stand plantation survival could be modeled using a generalized differential equation to describe the relative rate of mortality, i.e.,

$$\frac{1}{N} \frac{dN}{dt} = f(t) X$$

Where N is the number of trees per unit area, $f(t)$ is a function of time, and X can be a function of any whole-stand attribute. We narrowed the scope for viable survival models by assuming X equals one or N^d . Our search for a viable $f(t)$ function began by assuming using empirical evidence (Figure 1), that the function should be flexible enough to model a bathtub shape hazard function. This criterion led us to the functional form of

$$f(t) = \frac{1}{1+at} + \frac{b}{1+t} + ct.$$

Where t is time, and a, b, c are parameters. This function has the flexibility to model both monotonically increasing or decreasing hazards as well as a bathtub shape hazard function (Figure 2).

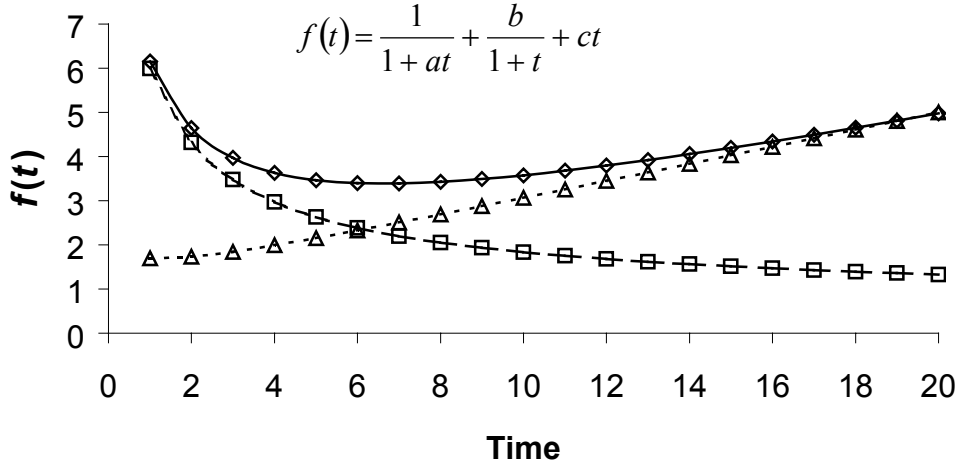


Figure 2. Examples of viable shapes for the chosen function $f(t)$ used in the differential equation for the whole-stand survival model.

Substituting $f(t)$ into the differential equation, and integrating for both values of X using the initial condition that when $N_2 = N_1$ then $t_2 = t_1$ results in model (1)

$$N_2 = N_1 \left(\frac{1+at_2}{1+at_1} \right)^{\frac{1}{a}} \left(\frac{1+t_2}{1+t_1} \right)^b e^{\frac{c}{2}(t_2^2-t_1^2)}, X=1 \quad (1)$$

and after re-parameterization, model (2)

$$N_2 = \left[N_1^d + a \ln \left(\frac{1+at_2}{1+at_1} \right) + b \ln \left(\frac{1+t_2}{1+t_1} \right) + c(t_2^2 - t_1^2) \right]^{\frac{1}{d}}, X=N^d. \quad (2)$$

Both models possess the desirable properties of path invariance and when $t_2 \rightarrow t_1$ then $N_2 \rightarrow N_1$. In addition, their respective lower asymptotes are zero, if c for model (1) and d for model (2) are negative. Model (2) also has the ability to distinguish among the density classes for survival if the data exhibit this trend.

Our motive for choosing $f(t)$ is its ability to model bathtub shapes. Moreover, we can establish if the resulting survival function is capable of modeling an underlying bathtub shaped hazard function by noting the concavity and number of inflection points in the survival curve. To obtain a bathtub shaped hazard curve, it is obvious that the hazard curve must first decrease over time, level off, and then increase. A bathtub shaped hazard curve corresponds to a survival curve with at least two inflection points. Furthermore, the slope of the survival curve (which is negative) between any two successive intervals must first decrease to the first inflection point, then increase from the first inflection to the second inflection point, and then decrease after the second inflection point. Models (1) and (2) have the flexibility to model multiple inflection points.

Models (1) and (2) were fitted to the CAPPS study data and evaluated by examining the residuals, fit index (defined as one minus the error sum of squares divided by total sum of squares), mean square error (MSE), root mean square error (RMSE), and error sum of squares (SSE). In addition, the behavior of the fitted functions was examined, both within the range of the data and extrapolating to a reasonable age.

RESULTS

Models (1) and (2) asymptotes were modified to allow for a biologically reasonable lower asymptote. This has been demonstrated to be a reasonable assumption for

plantation loblolly pine of the Southeastern United States (Harrison and Borders 1996, Martin and Brister 1999). In addition, models (1) and (2) were initially fit separately by treatment. Then the estimated parameters by treatment for models (1) and (2) were plotted and linear trends were detected for each parameter. However, some treatment parameters were not substantially different. Therefore, we refitted models (1) and (2) and allowed the fertilizer, herbicide, and herbicide/fertilizer cultural treatments to vary systematically from the baseline parameters. The cultural treatment effects were coded as $fert = 1$ if fertilized and zero otherwise, similarly for the herbicide ($herb$) and herbicide/fertilizer (hf) treatments. This resulted in a re-parameterization of the a , b , c , and d parameters where applicable for models (1) and (2) as

$$\left. \begin{aligned} a &= a_0 + a_1(fert) + a_2(herb) + a_3(hf) \\ b &= b_0 + b_1(fert) + b_2(herb) + b_3(hf) \\ c &= c_0 + c_1(fert) + c_2(herb) + c_3(hf) \\ d &= d_0 + d_1(fert) + d_2(herb) + d_3(hf) \end{aligned} \right\} \quad (3)$$

Cultural treatment parameters were removed from models (1) and (2) using a stepwise procedure ($\alpha = 0.05$). Models (1) and (2) achieved convergence easily during the cultural treatment parameter elimination process. Residual plots for the models were examined and since there was no evidence of heteroscedasticity, no weighting or transformations were necessary. The fitted models (1a) and (2a) that allows for systematic cultural treatment effects and a modified lower asymptote are

$$N_2 = N_{\min} + (N_1 - N_{\min}) \left(\frac{1 + a t_2}{1 + a t_1} \right)^{\frac{1}{a}} \left(\frac{1 + t_2}{1 + t_1} \right)^b e^{\frac{c}{2}(t_2^2 - t_1^2)}, \quad (1a)$$

Where the parameters that allow for cultural treatment are defined as

$$\left. \begin{aligned} a &= a_0 + a_1(\text{fert}) + a_2(\text{herb}) + a_3(\text{hf}) \\ b &= b_0 + b_1(\text{fert}) + b_2(\text{herb}) \\ c &= c_0 + c_1(\text{fert}) + c_3(\text{hf}) \end{aligned} \right\}$$

$N_i = TPH/100$, N_{\min} equals 2.5, which is the lower asymptote for $TPH/100$, t_i is plot age at time i , and a, b, c are parameters. The N_{\min} of 2.5 corresponds to approximately 100 trees per acre, which has been deemed a reasonable lower limit for loblolly pine of this region (Harrison and Borders 1996). Model (2a) is

$$N_2 = N_{\min} + \left[(N_1 - N_{\min})^d + a \ln\left(\frac{1 + a t_2}{1 + a t_1}\right) + b \ln\left(\frac{1 + t_2}{1 + t_1}\right) + c(t_2^2 - t_1^2) \right]^{\frac{1}{d}}, \quad (2a)$$

Where the final parameters that allow for systematic cultural treatment effects are

$$\left. \begin{aligned} a &= a_0 + a_2(\text{herb}) + a_3(\text{hf}) \\ b &= b_0 + b_2(\text{herb}) + b_3(\text{hf}) \\ c &= c_0 + c_1(\text{fert}) \\ d &= d_0 + d_2(\text{herb}) + d_3(\text{hf}) \end{aligned} \right\}$$

Models (1) and (2) and their respective parameter estimates, stand errors, and p-values are presented in Table 3. Summary fit statistics reveal that model (2) explains more of the variation in survival. Models (1) and (2) RMSE are 0.4590 and 0.4570, respectively. This means that for the average TPH (approximately 1600), there is less than a three-percent error. The fit index for models (1) and (2) are 0.9507 and 0.9511, respectively. There is no substantial difference between models (1) and (2) for these criteria. Hence, to further assess model performance, the mean survival and corresponding hazard functions were computed for the spectrum of plots and stratified by treatment.

Table 3. The estimated parameters and their associated standard errors and p-values for models (1) and (2) when fitted to the CAPPS study data.

Parameter	Model 1			Model 2		
	Estimate	Standard Error	Pr > t	Estimate	Standard Error	Pr > t
a	0.6843	0.0258	<0.0001	0.8316	0.0619	<0.0001
a _F	-0.1200	0.0433	0.0056			
a _H	0.0920	0.0386	0.0173	0.2170	0.0519	<0.0001
a _{HF}	-0.0500	0.0117	<0.0001	0.1104	0.0350	0.0016
b	-1.3539	0.0281	<0.0001	-0.03438	0.0045	<0.0001
b _F	-0.1759	0.0533	0.0010			
b _H	0.1548	0.0414	0.0002	0.2054	0.0737	0.0054
b _{HF}				0.1194	0.0540	0.0273
c	-0.00118	0.000271	<0.0001	-2.8565	0.3884	<0.0001
c _F	-0.00172	0.000666	0.0098	-0.3898	0.1308	0.0029
c _{HF}	-0.00219	0.000449	<0.0001			
d				-0.05453	0.00328	<0.0001
d _H				0.04294	0.00640	<0.0001
d _{HF}				0.03039	0.00942	0.0013

Cultural treatments are F = fertilizer, H = herbicide, and HF = herbicide and fertilizer.

Models (1) and (2) predicted mean survival and corresponding discrete hazard functions for the spectrum of plots illustrate that both models adequately mirror the empirical survival and hazard functions trends (Figure 3). Note that small differences in the survival curves can have a profound impact on the shape of the hazard function. Model (1) more closely mirrors the empirical hazard function for the early ages but model (2) exhibits more overall flexibility. Mean survival and their corresponding hazard function were computed and stratified by treatment for both models (Figures 4 and 5). Both models adequately reflect the underlying hazard function associated with the survival curves.

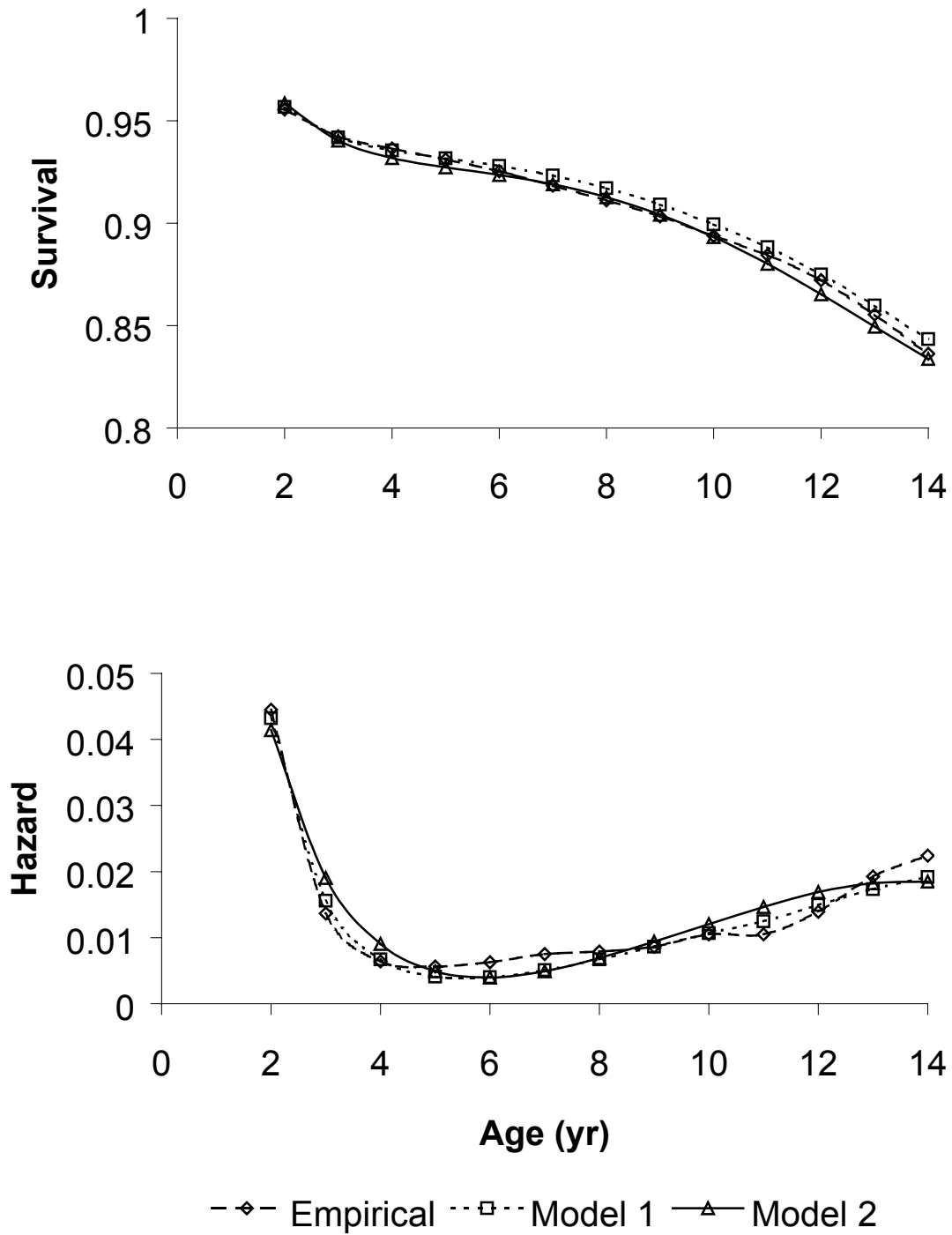


Figure 3. Models (1) and (2) fitted, and the empirical hazard and survival functions using the CAPPS study data.

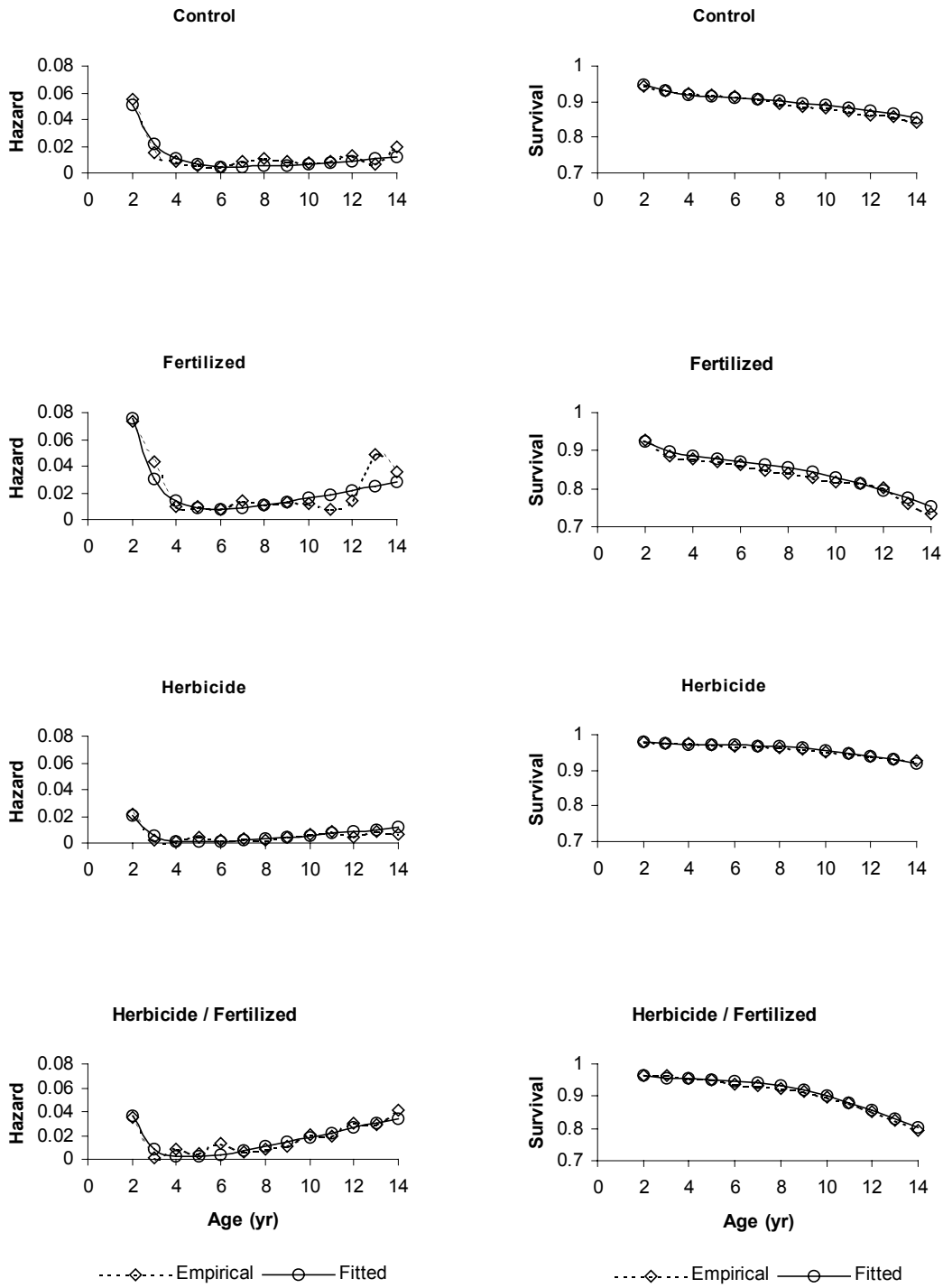


Figure 4. Model (1) fitted and the empirical survival and hazard functions by treatment and age.

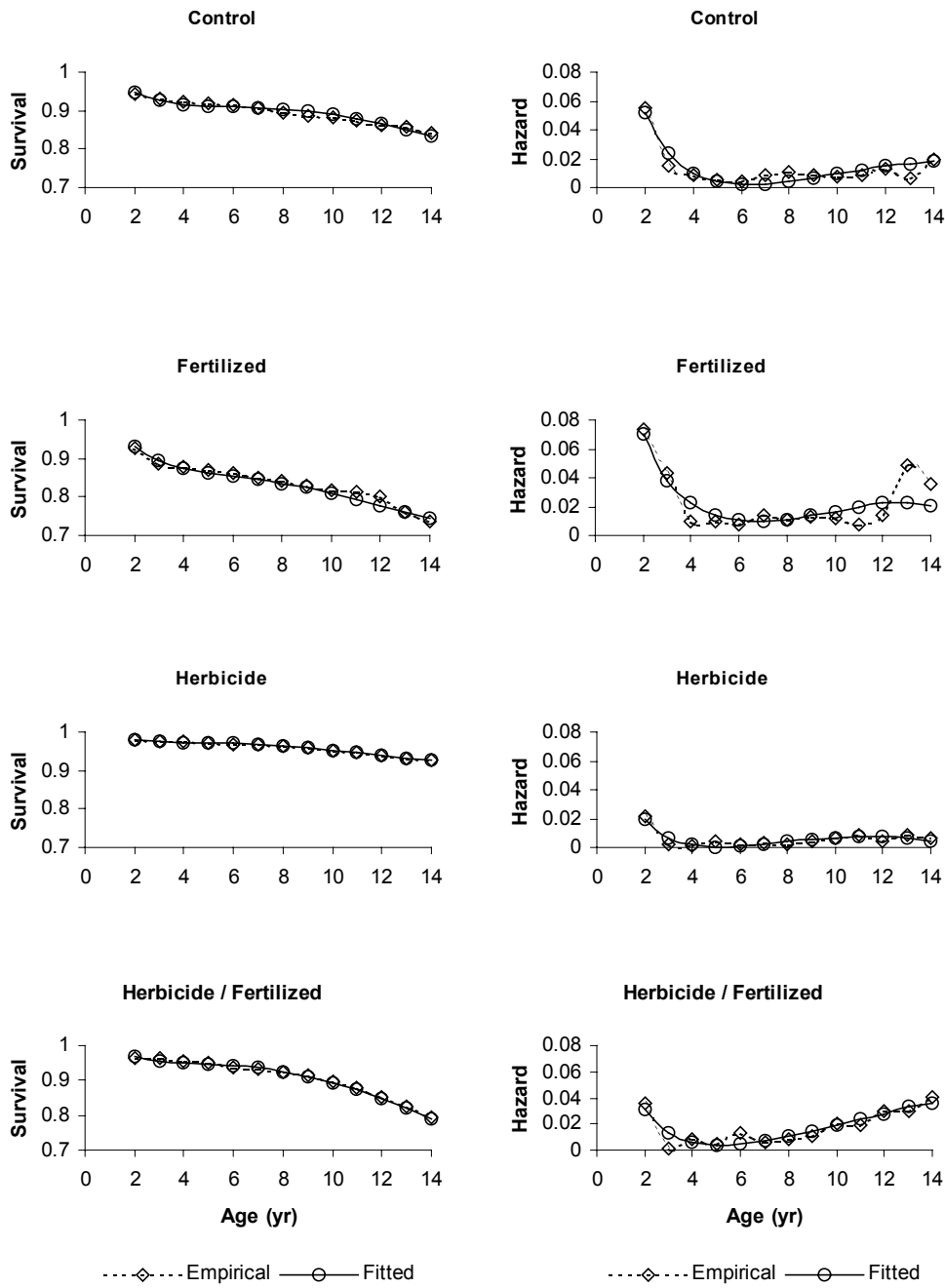


Figure 5. The CAPPS study model (2) fitted and the empirical survival and hazard functions by treatment and age.

Further assessment was conducted by examining the behavior of the fitted models to predict survival by treatment for ages 1-30 (Figure 6). Both models behave similarly within the data range (age 2-14). However, model extrapolation properties are substantially different for these models when fitted to our data. Model (1) provides reasonable survival curves by treatment and has a lower asymptote of 250 *TPH*. In contrast, model (2) has reasonable extrapolation properties only for the herbicide treatment and the herbicide/fertilizer treatment declines rapidly beyond the data range. Additionally, the control and fertilizer treatments are only able to predict to age 17 using the estimated parameters because the term in the bracket for model (2) when predicting survival for these treatments is negative after age 16, which is raised to a negative fractional power. Model (1) provides more reasonable extrapolation predictions when fitted to our data. However, both models behave adequately and provide flexible solutions within our data range.

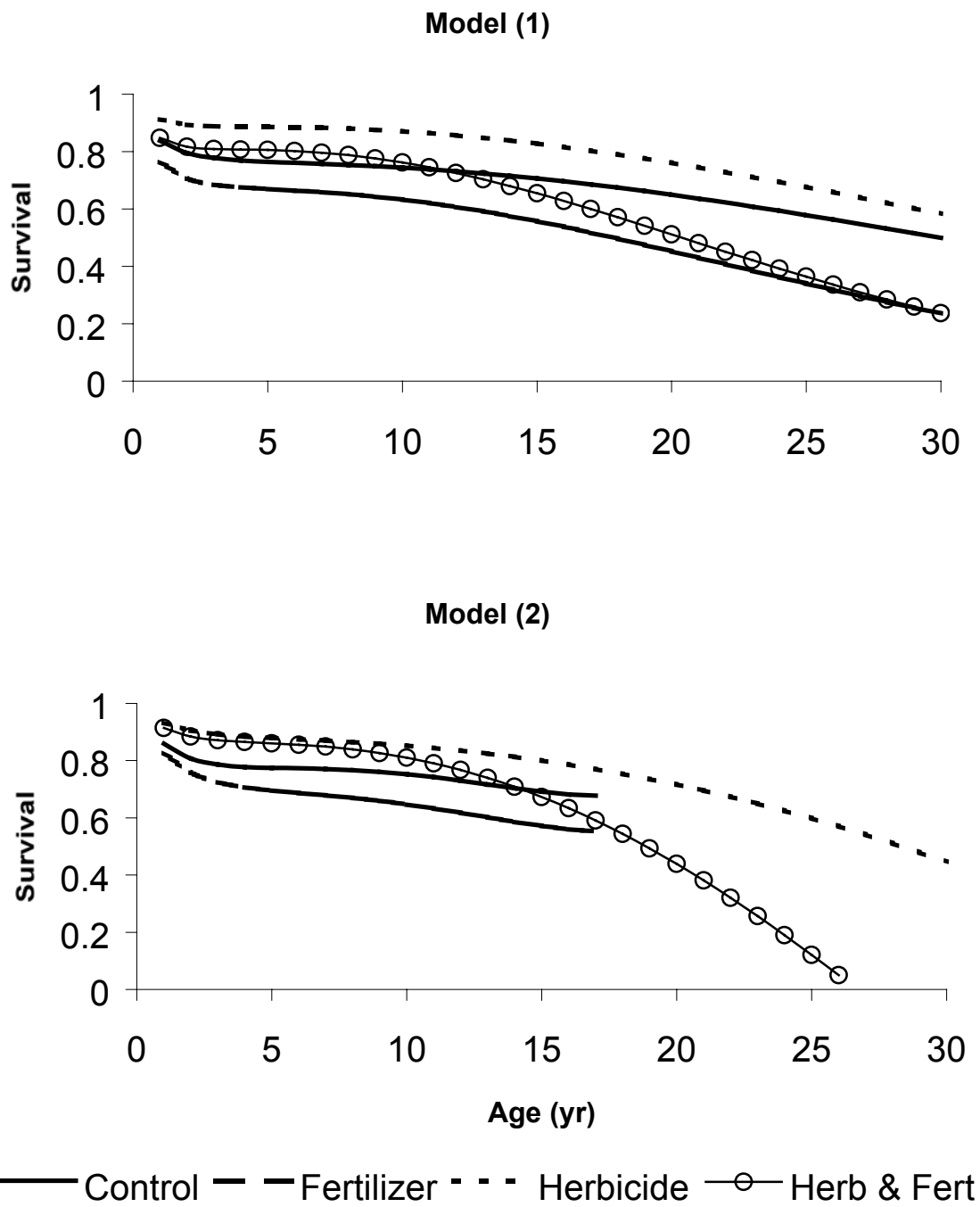


Figure 6. Predicted survival for ages 1-30 by treatment using models (1) and (2).

DISCUSSION

Flexible whole-stand survival models were developed and demonstrated to provide biologically reasonable solutions for complex underlying hazard functions. In addition, our empirical hazard curves illustrated that non-proportionality exists among the treatments, which was easily modeled by including, where necessary, treatment parameters. It is easily demonstrated that functional forms such as $f(t) = at^b$, that are simple linear or nonlinear functions of time are incapable of modeling complex hazard functions. These types of functions can be expanded to include other covariates, such as site index, that impact survival. However, the additional covariates are not necessarily enough in and of themselves to produce bathtub shaped hazard functions. In contrast, the addition of two (or more) linear/nonlinear functions can be developed that will model bathtub shaped hazard curves. Furthermore, the functional form does not need to have numerous parameters in order to model complex hazard functions, e.g., model (1) has only three baseline parameters. Model (2) demonstrated that different assumptions placed on the generalized differential equation form using the same $f(t)$ function as model (1) can produce more flexible hazard curves. Nevertheless, the less restrictive assumption used by model (2) for X provided only slightly more additional flexible solutions than those exhibited using just the chosen $f(t)$ function. Moreover, this increased flexibility has an extrapolation cost when model (2) is fitted to our data. Additional covariates in the general difference equation form may provide more flexibility in the solutions. However, the flexibility from using additional covariates is unlikely to approach the flexible whole-stand survival solutions gained from a flexible $f(t)$ function.

Our demonstrated method is relatively straightforward for developing a whole-stand survival model that is capable of modeling a complex underlying hazard function. The Weibull distribution and the Clutter and Jones (1980), hereafter referred to as C&J, models are probably the most widely used whole-stand survival models; therefore, these models were fitted to our data for comparison purposes. Neither model was able to achieve convergence for our data, probably because of the large decrease in early stand survival (Figure 1). Since we could not achieve convergence and because it is common to have whole-stand plantation survival data that begins at age four or five, we fitted these models and our model (1) after eliminating all data prior to age four, i.e., our first measurements for survival began at age five. All three models easily achieved convergence using the age 5-14 survival data. The results revealed that all three models fit well, but the C&J model has a more favorable fit index of 0.9842 than the Weibull (0.9834) and model (1) (0.9834). Our motive for fitting these models to the age 5-14 data was not to determine which model fit the data better, but to establish which model would more accurately predict the empirical hazards for ages 2-4. The predicted survival curves and corresponding hazard curves for these models, using $N_l = TPH = 1600$ (approximately the average) for the ages 2-14 reveal that model (1) is able to extrapolate extremely well for these data (Figure 7). All three models adequately mirror the empirical hazard for the range of the data (5-14 years). However, the Weibull and C&J models extrapolate poorly when predicting the age two and three hazards. Model (1) performs excellent when its predicted hazards are compared to the ages 2, 3, and 4 empirical hazards.

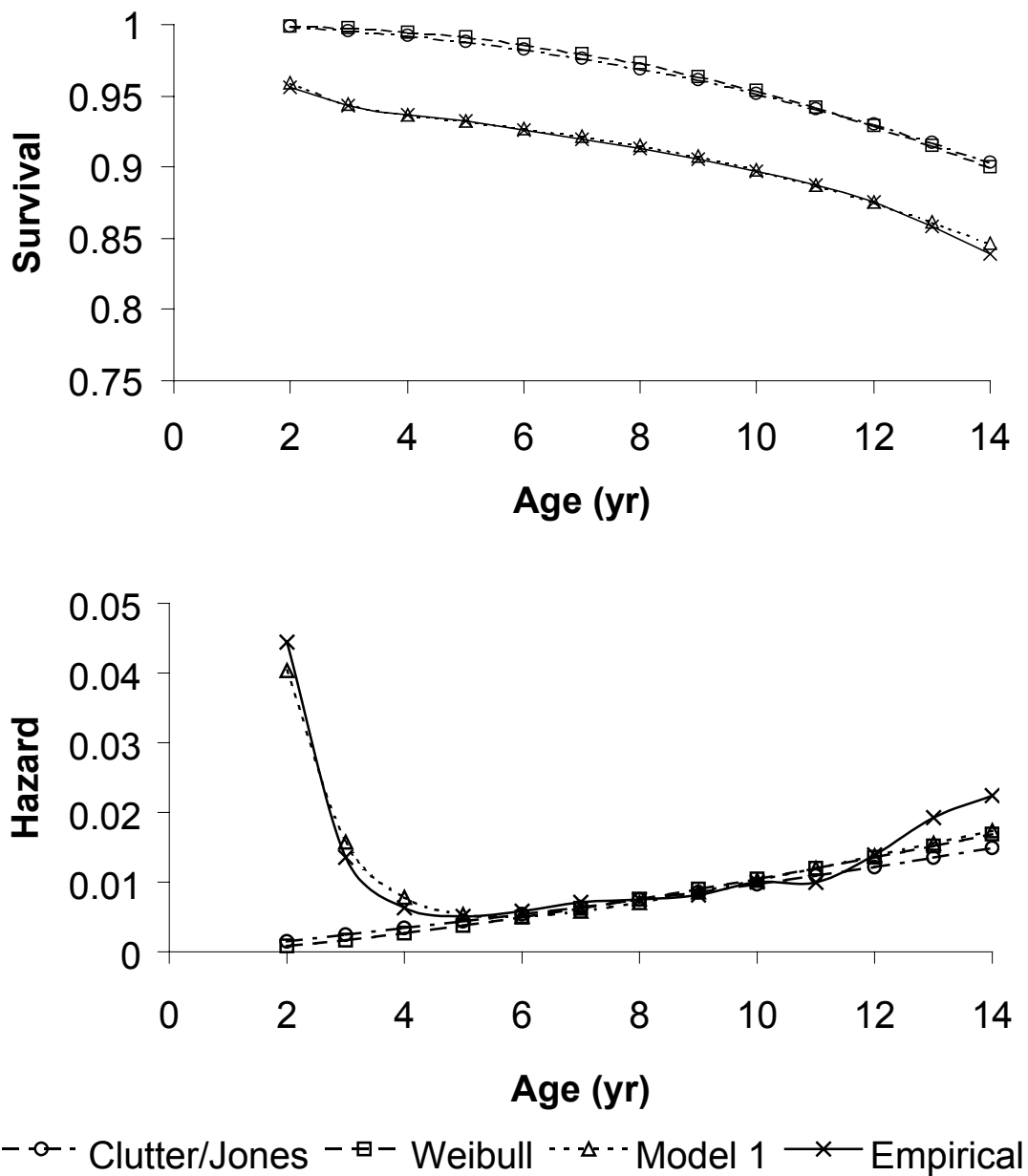


Figure 7. Predicted survival and their respective hazard curve for empirical, and the Weibull, Clutter and Jones (1980), and Model (1) models. The models were fitted excluding the data from ages 2-4.

This demonstrates that although the C&J model fits marginally better than model (1) when fitted to the age 5-14 data, the extrapolation predictions for the age 2-4 hazards are substantially improved using model (1). Furthermore, the age one hazard for the C&J and model (1) are 0.0005 and 0.1387, respectively. Hence, model (1) is behaving more biologically reasonable with respect to early whole-stand loblolly pine survival of this region. This demonstrates that a model may provide adequate future extrapolation predictions, but may not behave reasonable for early survival extrapolation predictions because its hazard function is more restrictive. Hence, extrapolating both forward and backwards (if the data set warrants) are equally important in establishing if a model has biologically reasonable behavior. Oftentimes the model developer is only interested in predicting survival within the range of the data. Regardless of the data time frame for a given study, plotting the empirical hazard function can aid in establishing the complexity of the function that may be necessary to adequately model whole-stand survival.

CONCLUSION

Forestry survival is a complex and difficult process to model; but the difficulty can be reduced by using the empirical hazard function behavior to aid in selecting an appropriate survival function. Use of the hazard function to aid in forestry survival model selection is not novel. Preisler and Slaughter (1997) demonstrated that they could limit their individual-tree survival model selection to a model that was capable of reflecting the empirical hazard function behavior. However, their study did not consider any survival functions capable of having bathtub shaped behavior. Our study established that improved whole-stand survival models could result by considering the underlying

hazard function. We demonstrated this by using the empirical hazard function to limit our selection to an appropriate function that could model the bathtub shape. Note that although our fitted models (1) and (2) consist of 10 and 11 parameters, respectively, it is not the number of parameters that allows our model to reflect an underlying bathtub shaped hazard function behavior. Our models when fitted to the data set without cultural treatment parameters still exhibit the underlying bathtub hazard function behavior. The cultural treatment parameters allow additional model flexibility, but it is the $f(t)$ function that allows our model to exhibit bathtub shaped trends if the data warrants. Whole-stand lifespan survival is commonly modeled using a system of equations, which oftentimes creates a cumbersome and difficult system to implement. It was demonstrated that one equation might provide the desired flexibility when derived from knowledge about the underlying hazard function. Furthermore, we demonstrated the ability of a simple yet flexible function to be integrated to obtain an initial condition equation. Survival is generally the least understood and hardest to model within a whole-stand forestry growth and yield system, however our method is relatively easy to implement and can model a whole-stand survival curve that exhibits a complex underlying hazard function behavior.

LITERATURE CITED

- Amateis, R.L., H.E. Burkhart, and L. Jiping. 1997. Modeling survival in juvenile and mature loblolly pine plantations. *For. Ecol. Manage.* 90:51-58.
- Belli, K.L., and A.R. Ek. 1988. Growth and survival modeling for planted conifers in the great lakes region. *For. Sci.* 34(2):458-473.
- Buford, M.A., and W.L. Hafley. 1985. Probability distributions as models for mortality. *For. Sci.* 31(2):331-341.
- Collet, D. 1994. Modelling survival data in medical research. Chapman & Hall. London. 347 p.
- Clutter, J.L., J.C. Fortson, L.V. Pienaar, G.H. Brister, and R.L. Bailey. 1983. Timber management a quantitative approach. John Wiley & Sons, Inc. New York. 333 p.
- Clutter, J.L., and E.P. Jones. 1980. Predictions of growth after thinning in old-field slash pine plantations. USDA For. Serv. Res. Pap. SE-217. 14 p.
- Devine, O.J., and J.L. Clutter. 1985. Predictions of survival in slash pine plantations infected with fusiform rust. *For. Sci.* 31(1):88-94.
- Glover, G.R., and J.N. Hool. 1979. A basal area ratio predictor of loblolly pine plantation mortality. *For. Sci.* 25(2):275-282.
- Hamilton, D.A., 1986. A logistic model of mortality in thinned and unthinned mixed conifer stands of northern Idaho. *For. Sci.*, 32:989-1000.
- Harrison, W.M., and B.E. Borders. 1996. Yield prediction and growth projection for site-prepared loblolly pine plantations in the Carolinas, Georgia, Alabama, and Florida. The University of Georgia, PMRC Tech. Rep. 1996-1.

- Hjorth, U. 1980. A reliability distribution with increasing, decreasing, constant and bathtub-shaped failure rates. *Technometrics*. 22(1):99-107.
- Kaplan, E.L., and P. Meier. 1958. Nonparametric estimation from incomplete observations. *JASA*. 53:457-481.
- Lawless, S.W. 1982. *Statistical models and methods for lifetime data*. New York: John Wiley & Sons, Inc. 580 p.
- Lorimer, C.G., and L.E. Frelich. 1984. A simulation of equilibrium diameter distribution of sugar maple (*Acer saccharum*). *Bull. Torrey Bot. Club* 111:193-199.
- Martin, S.W., R.L. Bailey, and E.J. Jokela. 1999. Growth and yield predictions for lower coastal plain slash pine plantations fertilized at mid-rotation. *SJAF*. 23(1):1-7.
- Martin, S.W., and G.H. Brister. 1999. A growth and yield model incorporating hardwood competition for natural loblolly pine stands in the Georgia piedmont. *SJAF*. 23(3):179-185.
- Matney, T.G., and R.M. Farrar Jr. 1992. A thinned/unthinned loblolly pine growth and yield simulator for planted cutover site-prepared land in the mid-gulf south. *SJAF*. 16(2):70-75.
- Monserud, R.A., and H. Sterba. 1999. Modeling individual tree mortality for Austrian forest species. *For. Ecol. Manage.* 113:109-123.
- Pienaar, L.V., and J.W. Rheney. 1993. Yield prediction for mechanically site-prepared slash pine plantations in the southeastern coastal plain. *SJAF*. 17(4):163-173.
- Pienaar, L.V., and B.D. Shiver. 1981. Survival functions for site-prepared slash pine plantations in the flatwoods of Georgia and northern Florida. *SJAF*. 5:59-62.

- Pinder III, J.E., J.G. Wiener, and M.H. Smith. 1978. The Weibull distribution: a new method of summarizing survivorship data. *Ecology*. 59(1):175-179.
- Preisler, H.K., and G.W. Slaughter. 1997. A stochastic model for tree survival in stands affected by annosum root disease. *For. Sci.* 43(1):78-85.
- Somers, G.L., R.G. Oderwald, W.H. Harms, and O.G. Langdon. 1980. Predicting mortality with a Weibull distribution. *For. Sci.* 26(2):291-300.
- Staebler, G.R., 1953. Mortality estimation in fully stocked stands of young growth Douglas-fir. USDA For. Serv. Pac Northwest Forest and Range Exp. Stn. Res. Pap. 4. 8 p.