Plots for Survival Regression

David J. Olive*

February 9, 2011

In the most used survival regression models, the conditional distribution of a time to event Y is independent of the predictors given a linear combination of predictors called the sufficient predictor. Slicing the estimated sufficient predictor into J groups and plotting the Cox survival curve versus the Kaplan Meier estimator for each group can be used to visualize the Cox regression model and as a diagnostic for goodness of fit. Plots for Weibull and Exponential regression are also given.

Key Words: Accelerated failure time model; Cox proportional hazards regression; Goodness of fit; Outliers.

*David J. Olive is Associate Professor, Department of Mathematics, Southern Illinois University, Mailcode 4408, Carbondale, IL 62901-4408, USA. E-mail address: dolive@math.siu.edu.

1. INTRODUCTION

Regression models are used to study the conditional distribution $Y|\boldsymbol{x}$ given the $p \times 1$ vector of nontrivial predictors $\boldsymbol{x}$. In survival regression, Y is the time until an event such as death. For many of the most important survival regression models, the nonnegative response variable Y is independent of $\boldsymbol{x}$ given $\boldsymbol{\beta}^T\boldsymbol{x}$, written $Y \perp\!\!\!\perp \boldsymbol{x}|\boldsymbol{\beta}^T\boldsymbol{x}$. Let the sufficient predictor $SP = \boldsymbol{\beta}^T\boldsymbol{x}$, and the estimated sufficient predictor $ESP = \hat{\boldsymbol{\beta}}^T\boldsymbol{x}$. The ESP is sometimes called the estimated risk score.

The conditional distribution $Y|\boldsymbol{x}$ is completely determined by the probability density function $f_{\boldsymbol{x}}(t)$, the distribution function $F_{\boldsymbol{x}}(t)$, the survival function

$$S_{\boldsymbol{x}}(t) \equiv S_{Y|SP}(t) = P(Y > t|SP = \boldsymbol{\beta}^T\boldsymbol{x}),$$

the cumulative hazard function $H_{\boldsymbol{x}}(t) = -\log(S_{\boldsymbol{x}}(t))$ for $t > 0$, or the hazard function $h_{\boldsymbol{x}}(t) = \frac{d}{dt}H_{\boldsymbol{x}}(t) = f_{\boldsymbol{x}}(t)/S_{\boldsymbol{x}}(t)$ for $t > 0$. High hazard implies low survival times while low hazard implies long survival times.

Survival data is usually right censored so Y is not observed. Instead, the survival time $T_i = \min(Y_i, Z_i)$ where $Y_i \perp\!\!\!\perp Z_i$ and Z_i is the censoring time. Also $\delta_i = 0$ if $T_i = Z_i$ is censored and $\delta_i = 1$ if $T_i = Y_i$ is uncensored. Hence the data is $(T_i, \delta_i, \boldsymbol{x}_i)$ for $i = 1, ..., n$.

The *Cox proportional hazards* regression model (Cox 1972) is a semiparametric model with $SP = \boldsymbol{\beta}_C^T\boldsymbol{x}$ and

$$h_{\boldsymbol{x}}(t) \equiv h_{Y|SP}(t) = \exp(\boldsymbol{\beta}_C^T\boldsymbol{x})h_0(t) = \exp(SP)h_0(t)$$

where the baseline hazard function $h_0(t)$ is left unspecified. The survival function is

$$S_{\boldsymbol{x}}(t) \equiv S_{Y|SP}(t) = [S_0(t)]^{\exp(\boldsymbol{\beta}_C^T\boldsymbol{x})} = [S_0(t)]^{\exp(SP)}. \tag{1.1}$$

If $\boldsymbol{x} = \mathbf{0}$ is within the range of the predictors, then the baseline survival and hazard functions correspond to the survival and hazard functions of $\boldsymbol{x} = \mathbf{0}$. First $\boldsymbol{\beta}_C$ is estimated by the maximum partial likelihood estimator $\hat{\boldsymbol{\beta}}_C$, then estimators $\hat{h}_0(t)$ and $\hat{S}_0(t)$ can be found (see Breslow 1974), and

$$\hat{S}_{\boldsymbol{x}}(t) = [\hat{S}_0(t)]^{\exp(\hat{\boldsymbol{\beta}}_C^T \boldsymbol{x})} = [\hat{S}_0(t)]^{\exp(ESP)}. \tag{1.2}$$

For *parametric proportional hazards* regression models, the baseline function is parametric and the parameters are estimated via maximum likelihood. Then $SP = \boldsymbol{\beta}_P^T \boldsymbol{x}$,

$$h_{\boldsymbol{x}}(t) = \exp(\boldsymbol{\beta}_P^T \boldsymbol{x}) h_{0,P}(t),$$

the survival function is

$$S_{\boldsymbol{x}}(t) \equiv S_{Y|SP}(t) = [S_{0,P}(t)]^{\exp(\boldsymbol{\beta}_P^T \boldsymbol{x})} = [S_{0,P}(t)]^{\exp(SP)}, \tag{1.3}$$

and

$$\hat{S}_{\boldsymbol{x}}(t) = [\hat{S}_{0,P}(t)]^{\exp(\hat{\boldsymbol{\beta}}_P^T \boldsymbol{x})} = [\hat{S}_{0,P}(t)]^{\exp(ESP)}. \tag{1.4}$$

For a parametric *accelerated failure time* model,

$$\log(Y_i) = \alpha + \boldsymbol{\beta}_A^T \boldsymbol{x}_i + \sigma e_i \tag{1.5}$$

where the e_i are iid from a location scale family. The parameters are again estimated by maximum likelihood and the survival function is

$$S_{\boldsymbol{x}}(t) \equiv S_{Y|\boldsymbol{x}}(t) = S_0\left(\frac{t}{\exp(\boldsymbol{\beta}_A^T \boldsymbol{x})}\right),$$

and

$$\hat{S}_{\boldsymbol{x}}(t) = \hat{S}_0\left(\frac{t}{\exp(\hat{\boldsymbol{\beta}}_A^T \boldsymbol{x})}\right)$$

where $\hat{S}_0(t)$ depends on $\hat{\alpha}$ and $\hat{\sigma}$.

The following univariate results will be useful for Exponential and Weibull regression. If Y has a Weibull distribution, $Y \sim W(\gamma, \lambda)$, then $S_Y(t) = \exp(-\lambda t^\gamma)$ where t, λ and γ are positive. If $\gamma = 1$, then Y has an Exponential distribution, $Y \sim EXP(\lambda)$ where $E(Y) = 1/\lambda$. Now V has a smallest extreme value distribution, $V \sim SEV(\theta, \sigma)$, if

$$S_V(t) = P(V > t) = \exp\left(-\exp\left(\frac{t-\theta}{\sigma}\right)\right)$$

where $\sigma > 0$ while t and θ are real. If $Z \sim SEV(0, 1)$, then $V = \theta + \sigma Z \sim SEV(\theta, \sigma)$ since the SEV distribution is a location scale family. Also, $V = \log(Y) \sim SEV(\theta = -\sigma \log(\lambda), \sigma = 1/\gamma)$, and $Y = e^V \sim W(\gamma = 1/\sigma, \lambda = e^{-\theta/\sigma})$.

For Weibull regression, the Weibull proportional hazards model (1.3) is valid if and only if the Weibull accelerated failure time model (1.5) is valid, where $e_i \sim SEV(0, 1)$. Hence a goodness of fit plot for the Weibull proportional hazards model is also a goodness of fit plot for the Weibull accelerated failure time model. Now $\log(Y)|\boldsymbol{x} \sim SEV(\alpha + \boldsymbol{\beta}_A^T \boldsymbol{x}, \sigma)$, and as a proportional hazards model, $Y|\boldsymbol{x} \sim W(\gamma = 1/\sigma, \lambda_{\boldsymbol{x}})$ where

$$\lambda_{\boldsymbol{x}} = \exp\left[-\left(\frac{\alpha}{\sigma} + \frac{\boldsymbol{\beta}_A^T \boldsymbol{x}}{\sigma}\right)\right] = \lambda_0 \exp(\boldsymbol{\beta}_P^T \boldsymbol{x})$$

with $\lambda_0 = \exp(-\alpha/\sigma)$ and $\boldsymbol{\beta}_P = -\boldsymbol{\beta}_A/\sigma$. Thus for $t > 0$, $P(Y > t|\boldsymbol{x}) = S_{\boldsymbol{x}}(t)$

$$= \exp(-\lambda_{\boldsymbol{x}} t^\gamma) = \exp(-\lambda_0 \exp(\boldsymbol{\beta}_P^T \boldsymbol{x}) t^\gamma) = [\exp(-\lambda_0 t^\gamma)]^{\exp(\boldsymbol{\beta}_P^T \boldsymbol{x})} = [S_{0,P}(t)]^{\exp(\boldsymbol{\beta}_P^T \boldsymbol{x})}.$$

Exponential regression is the special case where $\sigma = 1$. See, for example, Collett (2003, pp. 176-178) and Hosmer and Lemeshow (1999, pp. 108, 290).

The literature for checking the goodness of fit of the proportional hazards model is fairly large. See, for example, Arjas (1988), Gill and Schumaker (1987), Kay (1984), Lin

and Wei (1991), Marzec and Marzec (1997), Ng'andu (1997), Quantin, Moreau, Asselain, Maccario and Lellouch (1996) and Yu, Chappell, Wong, Hsu and Mazur (2008). May and Hosmer (1998) show how to implement the Grϕnnesby and Borgan (1996) test which partitions the ESP.

Grambsch and Therneau (1994) give a useful graphical check. Suppose the ith case had an uncensored survival time t_i. Let the scaled Schoenfeld residual for the ith observation and jth variable x_j be $r^*_{P,j}(t_i)$. For each variable, plot the t_i versus the $r^*_{P,j}(t_i) + \hat{\beta}_j$ and add the loess curve. If the loess curve is approximately horizontal for each of the p plots, then the proportional hazards assumption is reasonable. Alternatively, fit a line to each plot and test that each of the p slopes is equal to 0. The *R/Splus* function `cox.zph` makes both the plots and tests. See MathSoft (1999, pp. 267, 275). Hosmer and Lemeshow (1999, p. 211) suggest also testing whether the interactions $x_i \log(t)$ are significant for $i = 1, ..., p$.

If the Y_i are iid but censored data (T_i, δ_i) is observed, then the Kaplan Meier (1958) product limit estimator $\hat{S}_{KM}(t)$ is used to estimate $S_Y(t) = P(Y > t)$. This estimator is often used in graphical diagnostics for survival regression models, e.g., the cumulative hazard plot and log cumulative hazard plot. Allison (1995, p. 96) and Collett (2003, pp. 123, 182, 236) suggest that these two plots are not very useful in practice.

Section 2 suggests 4 new plots: the slice survival plot, the censored response plot, the log censored response plot and the EE plot. Inference should not be performed unless the survival regression model has been shown to be a useful approximation for the data. Section 3 shows how to assess the adequacy of the survival regression model with the new plots.

2. ADDITIONAL PLOTS

The *slice survival plot* divides the ESP into J groups of roughly the same size. For each group j with n_j cases, the model estimated survival function $\hat{S}_j(t)$ is computed using the $\boldsymbol{x}$ corresponding to the "median ESP" of the group (the kth order statistic of the ESP in group j, where $k = 1 + floor[(n_j - 1)/2]$). Let $\hat{S}_{KMj}(t)$ be the Kaplan Meier estimator computed from the survival times (T_i, δ_i) in the jth group. For each group, $\hat{S}_j(t)$ is plotted and $\hat{S}_{KMj}(t_i)$ is plotted as circles at the uncensored event times t_i. The survival regression model is reasonable if the circles "tracks $\hat{S}_j$ well" in each of the J plots.

If the slice widths go to zero, but the number of cases per slice increases to ∞ as $n \to \infty$, then the Kaplan Meier estimator and the model estimator converge to $S_{Y|SP}(t)$ if the model holds. Simulations suggest that the two survival functions are "close" for moderate n and nine slices. For small n and skewed predictors, some slices may be too wide in that the model is correct but $\hat{S}_{KMj}(t)$ is not a good approximation of $S_{Y|SP}(t)$ where SP corresponds to the $\boldsymbol{x}$ used to compute $\hat{S}_j(t)$.

For the Cox model, if pointwise confidence interval (CI) bands are added to the plot, then $\hat{S}_{KMj}$ "tracks $\hat{S}_j$ well" if most of the plotted circles do not fall very far outside the pointwise CI bands since these pointwise bands are not as wide as simultaneous bands. Collett (2003, pp. 241-243) places several observed Kaplan Meier curves with fitted curves on the same plot.

Survival regression is the study of the conditional survival $S_{Y|SP}(t)$, and the slice survival plot is a useful tool for visualizing $S_{Y|SP}(t)$ in the background of the data. Suppose

the jth slice is narrow so that $ESP \approx w_j$. If the model is reasonable, $ESP \approx SP$, and the number of uncensored cases in the jth slice is not too small, then $S_{Y|SP=w_j}(t) \approx \hat{S}_j(t) \approx \hat{S}_{KMj}(t)$. (These quantities approximate $[\hat{S}_0(t)]^{\exp(w_j)}$ for the Cox model.) Thus the nonparametric Kaplan Meier estimator is used to check the model estimator $\hat{S}_j(t)$ in each slice.

The slice survival plot tailored to the Cox model is closely related to May and Hosmer (1998) test. Van Houwelingen, Bruinsma, Hart, Veer and Wessels (2006) use similar ideas, but place the J Kaplan Meier curves on one plot and the J Cox survival curves on another plot. A similar plot has been suggested by several authors with $\boldsymbol{x}$ divided into J groups instead of the ESP. For example, see Miller (1981, p. 168). Hosmer and Lemeshow (1999, pp. 141–145) suggests making plots based on the quartiles of the ith predictor x_i, and note that a problem with Cox survival curves (1.2) is that they may use inappropriate extrapolation. Using the ESP results in narrow slices with many cases, and adding Kaplan Meier curves shows if there is extrapolation.

The following two plots can be regarded as the model checking plots of Cook and Weisberg (1997) extended to censored data, but the main use of the two plots is to check for cases with unusual survival times. A *censored response plot* is a plot of the ESP versus T with plotting symbol 0 for censored cases and + for uncensored cases. Slices in this plot correspond to the slices used in the slice survival plot. Gentleman and Crowley (1991) give a similar plot for models with a single predictor x.

Suppose the ESP is a good estimator of the SP. Consider a narrow vertical slice taken in the censored response plot about $ESP = w$. The points in the slice are a censored sample with $S_{Y|SP}(t) \approx S_{Y|w}(t)$. For proportional hazards models, $h_{Y|SP}(t) \approx$

$\exp(ESP)h_0(t)$, and the hazard increases while the survival decreases as the ESP increases.

Let $\log(T_i) = \hat{\alpha}+\hat{\boldsymbol{\beta}}_A^T\boldsymbol{x}_i+r_i$. For accelerated failure time models, a *log censored response (LCR) plot* is a plot of $\hat{\alpha}+\hat{\boldsymbol{\beta}}_A^T\boldsymbol{x}_i$ versus $\log(T_i)$ with plotting symbol 0 for censored cases and + for uncensored cases. The identity line with unit slope and zero intercept is added to the plot, and the vertical deviations from the identity line $= r_i$. Collett (2003, p. 231) defines a standardized residual $r_{Si} = r_i/\hat{\sigma}$.

For Weibull regression, $\log(Y)|(\alpha + \boldsymbol{\beta}_A^T\boldsymbol{x}) \sim SEV(\alpha + \boldsymbol{\beta}_A^T\boldsymbol{x}, \sigma)$. Thus points in a narrow vertical slice about $\hat{\alpha} + \hat{\boldsymbol{\beta}}_A^T\boldsymbol{x} = w$ are approximately a censored sample from an $SEV(w, \hat{\sigma})$ distribution if the fitted model is a good approximation to the data.

Censoring causes the bulk of the data to be below the identity line. For example, Hosmer and Lemeshow (1998, p. 226) state that for the Exponential regression model, $\hat{\alpha}$ forces

$$\sum_{i=1}^{n} \delta_i = \sum_{i=1}^{n} \frac{T_i}{\exp(\hat{\alpha} + \hat{\boldsymbol{\beta}}_A^T\boldsymbol{x}_i)}.$$

Hence $\hat{T}_i = \exp(\hat{\alpha}+\hat{\boldsymbol{\beta}}_A^T\boldsymbol{x}_i) \approx (n/\sum_{i=1}^{n}\delta_i)T_i$ (roughly). With no censoring, the bulk of the data will still be lower than the identity line if the e_i are left skewed as for the Weibull regression model where the $e_i \sim SEV(0, 1)$.

For parametric proportional hazards models, an *EE plot* is a plot of the parametric ESP $\hat{\boldsymbol{\beta}}_P^T\boldsymbol{x}$ versus the Cox semiparametric ESP $\hat{\boldsymbol{\beta}}_C^T\boldsymbol{x}$. If the parametric proportional hazards model is good, then the plotted points should track the identity line with unit slope and zero intercept. As $n \to \infty$, the correlation of the plotted points goes to 1 in probability for any finite interval, e.g., from the 1st percentile to the 99th percentile

of $\hat{\boldsymbol{\beta}}_C^T \boldsymbol{x}$. Lack of fit is suggested if the plotted points do not cluster tightly about the identity line. For the Exponential regression model, $\sigma = 1$ and $\boldsymbol{\beta}_C = -\boldsymbol{\beta}_A$, and the Exponential EE plot is a plot of

$$ESPE = -\hat{\boldsymbol{\beta}}_A^T \boldsymbol{x} \text{ versus } \text{ESPC} = \hat{\boldsymbol{\beta}}_{\text{C}}^{\text{T}} \boldsymbol{x}.$$

For the Weibull regression model, $\boldsymbol{\beta}_C = -\boldsymbol{\beta}_A/\sigma$, and the Weibull EE plot is a plot of

$$ESPW = \frac{-1}{\hat{\sigma}} \hat{\boldsymbol{\beta}}_A^T \boldsymbol{x} \text{ versus } \text{ESPC} = \hat{\boldsymbol{\beta}}_{\text{C}}^{\text{T}} \boldsymbol{x}.$$

Suppose the plotted points cluster tightly about the identity line in the EE plot with $corr(\hat{\boldsymbol{\beta}}_C^T \boldsymbol{x}_i, \hat{\boldsymbol{\beta}}_P^T \boldsymbol{x}_i) > 0.99$. Thus $\hat{\boldsymbol{\beta}}_C^T \boldsymbol{x} \approx \hat{\boldsymbol{\beta}}_P^T \boldsymbol{x}$ for the observed $\boldsymbol{x}_i$, and slicing on the Cox ESP is nearly the same as slicing on the parametric ESP. Make the slice survival plot for the Cox model and add the estimated parametric survival function (1.4) as crosses. If the parametric proportional hazards model holds, then (1.1) = (1.3). Thus if (1.2) $\approx$ (1.4) for any $\boldsymbol{x}_i$, then $S_{0,P}(t) \approx S_0(t)$, (1.2) $\approx$ (1.4) for all $\boldsymbol{x}_i$, and the parametric proportional hazards model is reasonable.

Thus checking parametric proportional hazards models has 3 steps: i) check that the proportional hazards assumption is reasonable with the slice survival plot for the Cox model, ii) check that the parametric and semiparametric ESPs are approximately the same, $\hat{\boldsymbol{\beta}}_P^T \boldsymbol{x} \approx \hat{\boldsymbol{\beta}}_C^T \boldsymbol{x}$ with the EE plot, and iii) using the slice survival plot, check that (1.2) $\approx$ (1.4) for the $\boldsymbol{x}$ used in each of the J slices. Since the Weibull proportional hazards model (1.3) is valid for $(Y, \boldsymbol{x})$ if and only if the Weibull accelerated failure time model (1.5) is valid for $(\log(Y), \boldsymbol{x})$, the above procedure can be used to simultaneously check the goodness of fit of both models. The slice survival plot for the Cox model is used because of the ease of making pointwise CI bands.

3. EXAMPLES

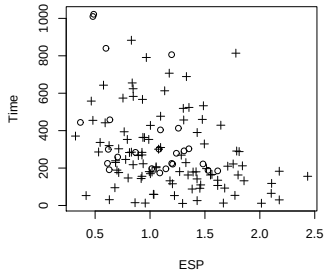

Figure 1: Censored Response Plot for R Lung Cancer Data

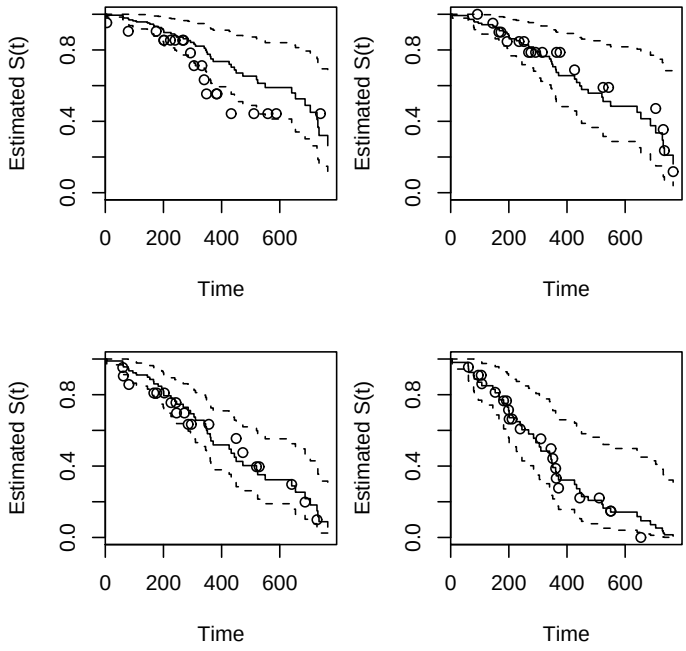

Figure 2: Slice Survival Plots for R Lung Cancer Data

Example 1. *R* and *Splus* contain a data set *lung* where the response variable Y is the time until death for patients with lung cancer. See MathSoft (1999, p. 268). Consider the data set for males with predictors *ph.ecog* = Ecog performance score 0-4, *ph.karno* =

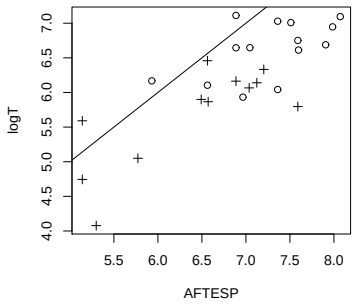

Figure 3: LCR Plot for Ovarian Cancer Data

a competitor to ph.ecog, *pat.karno* = patient's assessment of their karno score and *wt.loss* = weight loss in last 6 months. Figure 1 shows the censored response plot. Notice that the survival times decrease rapidly as the ESP increases and that there is one time that is unusually large for ESP ≈ 1.8. If the Cox regression model is a good approximation to the data, then the response variables corresponding to the cases in a narrow vertical strip centered at $ESP = w$ are approximately a censored sample from a distribution with hazard function $h_{\boldsymbol{x}}(t) \approx \exp(w)h_0(t)$. Figure 2 shows the slice survival plots. The ESP was divided into 4 groups and the ESP increases from the upper left, upper right, lower left and lower right corners of the plot where $\hat{S}(400) \approx (0.70, 0.60, 0.55, 0.30)$. The circles corresponding to the Kaplan Meier estimator are "close" to the Cox survival curves in that the circles do not fall very far outside the pointwise CI bands.

Example 2. The ovarian cancer data is from Collett (2003, pp. 187-190) and Edmunson et al. (1979). The response variable is the survival time of $n = 26$ patients in days with predictors *age* in years and *treat* (1 for cyclophosphamide alone and 2 for cyclophosphamide combined with adriamycin). Figure 3 shows that most of the plotted

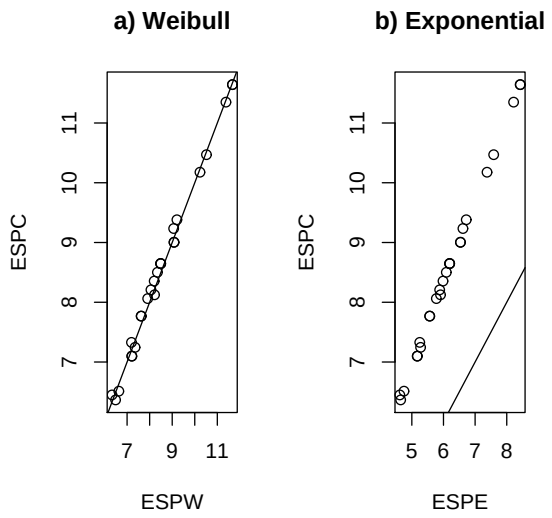

Figure 4: EE Plots for Ovarian Cancer Data

points in the LCR plot are below the identity line. If a Weibull regression model is a good approximation to the data, then the plotted points in a narrow vertical slice centered at $\hat{\alpha} + \hat{\boldsymbol{\beta}}^T \boldsymbol{x} = w$ are approximately a censored sample from an $SEV(w, \hat{\sigma})$ distribution. Figure 4 shows the Weibull and Exponential regression EE plots. Notice that the estimated risk scores from the Cox regression and Weibull regression are nearly the same with correlation = 0.997. The points from the Exponential regression do not cluster about the identity line. Hence Exponential regression should not be used. Figure 5 gives the slice survival plot for the Cox model with the Weibull survival function $\hat{S}_{\boldsymbol{x}}(t) = \exp[-\exp(-\hat{\gamma}\hat{\boldsymbol{\beta}}_A^T \boldsymbol{x})\ \exp(-\hat{\gamma}\hat{\alpha})\ t^{\hat{\gamma}}\ ]$ represented by crosses where $\hat{\gamma} = 1/\hat{\sigma}$. Notice that the Weibull and Cox estimated survival functions are close and thus similar. Again the circles corresponding to the Kaplan Meier estimator are "close" to the Cox survival curves in that the circles do not fall very far outside the pointwise CI bands.

Example 3. *R* contains a data set *nwtco* where the response variable Y is the time

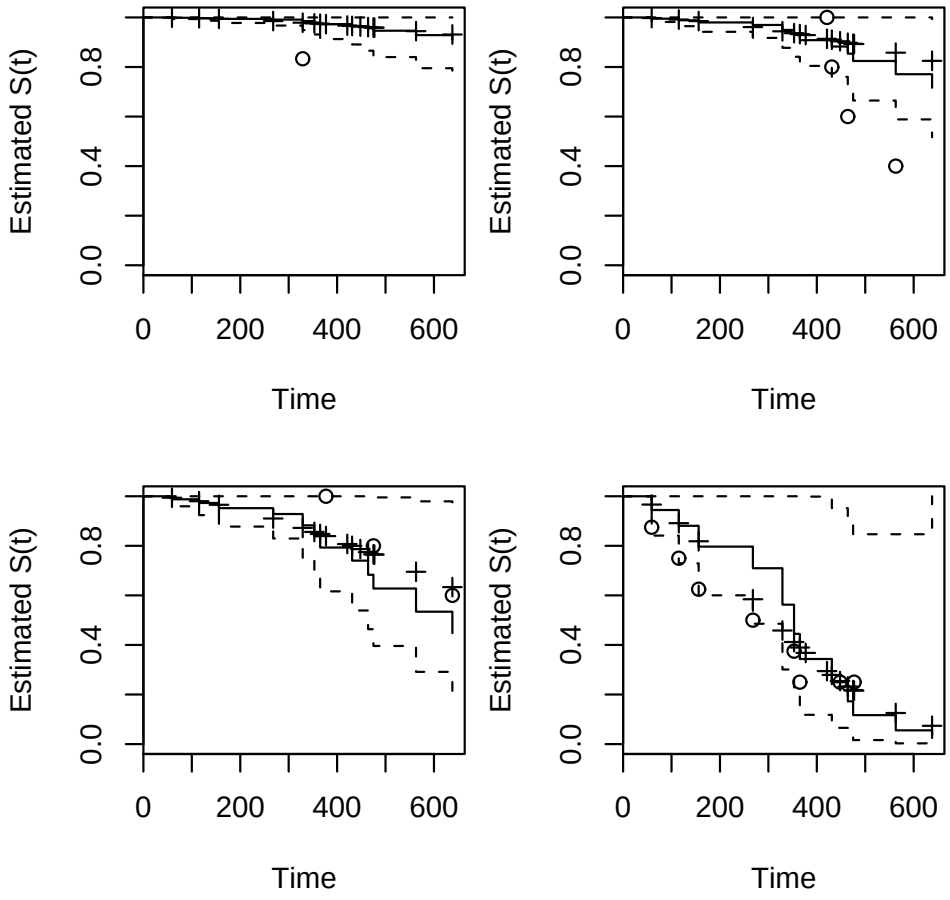

Figure 5: Slice Survival Plots for Ovarian Cancer Data

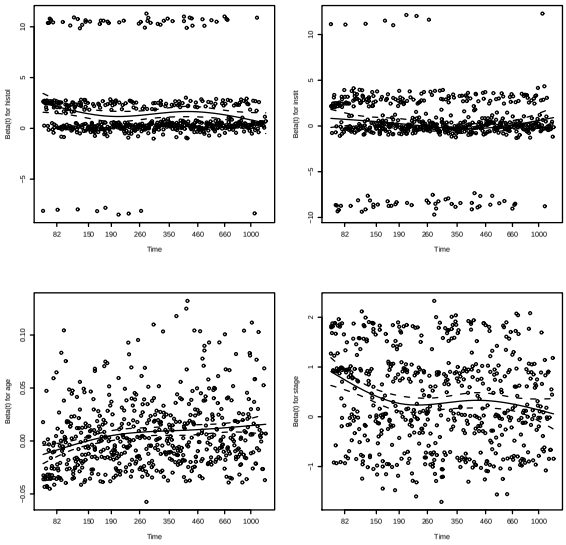

Figure 6: Grambsch and Therneau Plots for NWTCO Data

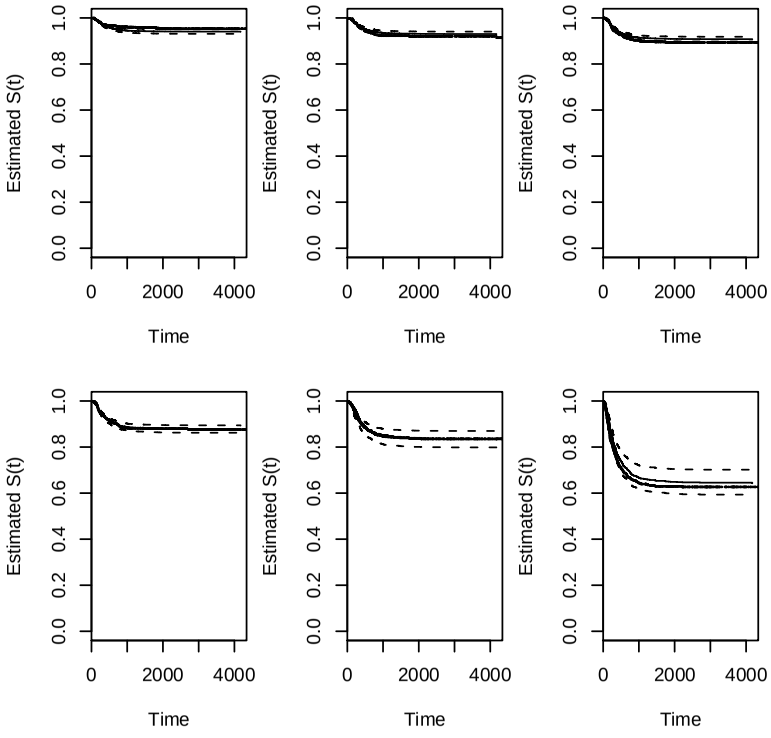

Figure 7: Slice Survival Plot for NWTCO Data

until relapse with $n = 4028$. The model used predictors *histol* = tumor histology from central lab, *instit* = tumor histology from local institution, *age* in months, and *stage* of disease from 1 to 4 (treated as an continuous variable). In Figure 6, the Grambsch and Therneau (1994) plots suggest that the Cox model is not valid since not all of the loess curves are flat, and the global test has p-value $\approx 5.66 \times 10^{-11}$. The slice survival plot in Figure 7 shows that the Cox survival estimators and Kaplan Meier estimators are nearly identical in the six slices, suggesting that the Cox model is a reasonable approximation to the data. The greatest contributors to lack of fit seem to be the predictors age and stage corresponding to the bottom two plots of Figure 6, and survival for small ESP corresponding to the upper left plot in Figure 7.

4. DISCUSSION

The slice survival plot is useful for visualizing $S_{Y|SP}(t)$ in the background of the data. The default $\boldsymbol{x}$ for R is the sample mean $\overline{\boldsymbol{x}}$ which may not be a valid $\boldsymbol{x}$ if some of the predictors are indicator variables. Choosing a representative $\boldsymbol{x}$ corresponding to the "median ESP" in each slice displays survival for the entire range of observed predictors $\boldsymbol{x}$. See Figures 2, 5 and 7. Pointwise CI bands can be used to determine whether the nonparametric Kaplan Meier estimator is close to the model estimator for each representative $\boldsymbol{x}$. If the two estimators are close for each slice, then the graph suggests that the model is giving a useful approximation to $S_{Y|SP}(t)$ for the observed data if the number of uncensored cases is large compared to the number of predictors p. The plots are also useful for teaching survival regression to students and for explaining the models to consulting clients.

Emphasis was on proportional hazards models since pointwise CI bands are available for the Cox proportional hazards model. Thus the slice survival plot can be made for the Cox model, and then the estimated survival function from a parametric proportional hazards model can be added as crosses for each slice if points in the EE plot cluster tightly about the identity line. Since the Weibull proportional hazard model is valid if and only if the Weibull accelerated failure time model is valid, the EE and slice survival plots can be used to simultaneously check the goodness of fit of both the proportional hazards and accelerated failure time models for Weibull and Exponential regression. Stratified proportional hazards models can be checked by making one slice survival plot per stratum. EE plots can be made for parametric models if software for a semiparametric analog is available. See Bennett (1983), Yang and Prentice (1999), Wei (1992) and Zeng and Lin (2007).

The four new plots are not used for models where some of the explanatory variables are time varying. If pointwise bands are not available for the parametric or semiparametric model, but the number of cases in each slice is large, then simultaneous or pointwise CI bands for the Kaplan Meier estimator could be added for each slice.

Plots were made in R using the survival library, and the function `coxph` produces the survival curves for Cox regression. See R Development Core Team (2008). The collection of R functions *regpack* available from (www.math.siu.edu/olive/regpack.txt) contains functions for reproducing simulations and some of the plots. The functions `vlung2`, `vovar` and `vnwtco` were used to produce plots in Examples 1, 2 and 3. The function `bphsim3` shows that the Kaplan Meier estimator was close to the Cox survival curves for 2 groups (a single binary predictor) when censoring was light and $n = 10$.

Simulated Weibull proportional hazards regression data was made following Zhou (2001) but with three iid N(0,1) covariates. The function `phsim5` showed that for 9 groups and $p = 3$, the Kaplan Meier and Cox curves were close (with respect to the pointwise CI bands) for $n \geq 80$. The function `wphsim` showed a similar result for Kaplan Meier curves (circles), and the function `wregsim2` shows that for $n \geq 30$, the plotted points in an EE plot cluster tightly about the identity line with correlation greater than 0.99 with high probability.

ACKNOWLEDGMENTS

This research was supported by NSF grant DMS-0600933.

REFERENCES

Allison, P.D. (1995), *Survival Analysis Using SAS: A Practical Guide,* SAS Institute, Cary, NC.

Arjas, E. (1988), "A Graphical Method for Assessing Goodness of Fit in Cox's Proportional Hazards Model," *Journal of the American Statistical Association*, 83, 204-212.

Bennett, S. (1983), "Analysis of Survival Data by the Proportional Odds Model," *Statistics in Medicine*, 2, 273-277.

Breslow, N.E. (1974), "Covariance Analysis of Censored Survival Data," *Biometrics*, 30, 89-100.

Collett, D. (2003), *Modelling Survival Data in Medical Research*, 2nd ed., Chapman & Hall/CRC, Boca Raton, FL.

Cook, R.D., and Weisberg, S. (1997), "Graphics for Assessing the Adequacy of Regression Models," *Journal of the American Statistical Association,* 92, 490-499.

Cox, D.R. (1972), "Regression Models and Life-Tables," *Journal of the Royal Statistical Society, B,* 34, 187-220.

Edmunson, J.H., Fleming, T.R., Decker, D.G., Malkasian, G.D., Jorgenson, E.O., Jeffries, J.A., Webb, M.J., and Kvols, L.K. (1979), "Different Chemotherapeutic Sensitivities and Host Factors Affecting Prognosis in Advanced Ovarian Carcinoma Versus Minimal Residual Disease," *Cancer Treatment Reports*, 63, 241-247.

Gentleman, R., and Crowley, J. (1991), "Graphical Methods for Censored Data," *Journal of the American Statistical Association,* 86, 678-683.

Gill, R.D., and Schumaker, M. (1987), "A Simple Test of the Proportional Hazards Assumption," *Biometrika*, 74, 289-300.

Grambsch, P.M., and Therneau, T.M. (1994), "Proportional Hazards Tests and Diagnostics Based on Weighted Residuals," *Biometrika*, 81, 515-526.

Grϕnnesby, J.K., and Borgan, Ø. (1996), "A Method for Checking Regression Models in Survival Analysis Based on the Risk Score," *Lifetime Data Analysis*, 2, 315-328.

Hosmer, D.W., and Lemeshow, S. (1999), *Applied Survival Analysis: Regression Modeling of Time to Event Data*, Wiley, NY.

Kaplan, E.L., and Meier, P. (1958), "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association,* 53, 457-481.

Kay, R. (1984), "Goodness-of-Fit Methods for the Proportional Hazards Model: a Review," *Reviews of Epidemiology Santé Publications*, 32, 185-198.

Lin, D.Y., and Wei, L.J. (1991), "Goodness-of-Fit Tests for the General Cox Regression Model," *Statistica Sinica*, 1, 1-17.

Marzec, L., and Marzec, P. (1997), "Generalized Martingale-Residual Processes for

Goodness-of-Fit Inference in Cox's Type Regression Models," *The Annals of Statistics*, 25, 683-714.

MathSoft (1999), *S-PLUS 2000 Guide to Statistics, Vol. 2,* Data Analysis Products Division, MathSoft, Seattle, WA. (MathSoft is now Insightful.)

May, S., and Hosmer, D.W. (1998), "A Simple Method for Calculating a Goodness-of-Fit Test for the Proportional Hazards Model," *Lifetime Data Analysis*, 4, 109-120.

Miller, R. (1981), *Survival Analysis*, Wiley, NY.

Ng'andu, N.H. (1997), "An Empirical Comparison of Statistical Tests for Assessing the Proportional Hazards Assumption of Cox's Model," *Statistics in Medicine*, 16, 611-626.

Quantin, C., Moreau, T., Asselain, B., Maccario, J., and Lellouch, J. (1996), "A Regression Survival Model for Testing the Proportional Hazards Hypothesis," *Biometrics*, 52, 874-885.

R Development Core Team (2008), "R: a Language and Environment for Statistical Computing," R Foundation for Statistical Computing, Vienna, Austria (www.R-project.org).

van Houwelingen, H.C., Bruinsma, T., Hart, A.A.M. Veer, L.J., and Wessels, L.F.A. (2006), "Cross-validated Cox Regression on Microarray Gene Expression Data," *Statistics in Medicine*, 25, 32013216.

Wei, L.J. (1992), "The Accelerated Failure Time Model: a Useful Alternative to the Cox Regression Model in Survival Analysis," *Statistics in Medicine*, 11, 1871-1879.

Yang, S., and Prentice, R.L. (1999), "Semiparametric Inference in the Proportional Odds Regression Model," *Journal of the American Statistical Association,* 94, 124-136.

Yu, Q., Chappell, R., Wong, G.Y.C., Hsu, Y., and Mazur, M., (2008), "Relationship Between Cox, Lehmann, Weibull, and Accelerated Lifetime Models," *Communications in Statistics Theory and Methods*, 37, 1458-1470.

Zeng, D., and Lin, D.Y. (2007), "Efficient Estimation for the Accelerated Failure Time Model," *Journal of the American Statistical Association*, 102, 1387-1396.

Zhou, M. (2001), "Understanding the Cox Regression Models with Time–Change Covariates," *The American Statistician*, 55, 153-155.