Effects of leaf emergence order on leaf lifespan are independent of life-form and successional status

Roger J. Dungan, Marie-Laure Navas, Richard P. Duncan, Eric Garnier

APPENDIX 1

Analysis of leaf survival distributions

Failure time analysis can be used to estimate functions which describe lifetime changes in either survival or mortality risk for individual leaves. The survival function, $S(t)$, describes the probability that a leaf will live longer than t without dying. The hazard function, $h(t)$, is a conditional failure rate, defined as the probability that a leaf will die during the interval t to $t + \Delta t$, given that it has survived to t (Fox, 2001). For a Weibull distribution the survival function is

$$S(t) = \exp(-\lambda t^{\alpha})$$

and the hazard function is

$$h(t) = \lambda\alpha(\lambda t)^{\alpha-1}.$$

For a log-normal distribution the survival function is

$$S(t) = 1 - \Phi(\alpha \log(\lambda t)),$$

where Φ is the standard normal distribution function. In its general form, the hazard function is

$$h(t) = \frac{f(t)}{S(t)},$$

where $f(t)$, the probability density function which describes the probability that a leaf will die in the interval t to $t + \Delta t$. For a log-normal distribution, this is

$$f(t) = (2\pi)^{-\frac{1}{2}} \alpha t^{-1} \exp\left(\frac{-\alpha^2 (\log(\lambda t))^2}{2} \right).$$

R code for calculating these functions, together with estimates of the median survival time, from the outputs from the `survival` package and derived from the formulations of Tableman and Kim (2004), is listed below. The expected lifespan can be calculated from $S(t)$ and corresponds to the time when $S(t) = 0.5$. Although the shape of $S(t)$ is broadly similar for a Weibull and log-normal distribution, shapes for $h(t)$ differ significantly (Figure 1), emphasising the importance of systematically selected the appropriate underlying survival distribution. In any case, either of these distributions tended to closely match the observed non-parametric distribution calculated with non-parametric maximum likelihood methods (Grooneboom and Wellner, 1992; Figure 1).

Survival analysis with R

Analysis of leaf lifespans with failure time analysis using the R software package requires the contributed package `survival` (available for download at http://cran.r-project.org/). For this analysis data are structured as a matrix with each leaf represented as a row. For each leaf, the minimum (MinAge) and maximum (MaxAge) possible leaf age, calculated as described by Dungan *et al*. (2003), are given in columns, together with a code signifying the kind of censoring the data are subject to (in the case of interval censoring, this is 3). Subsequent columns contain variables describing constraints on leaf lifespan, such as leaf emergence order:

```
MinAge MaxAge CensorCode LeafOrder ...
80        147          3         1
162       222          3         2
2          61          3         3
30         89          3         4
30         89          3         4
...
```

Once the survival package is loaded, a parametric survival regression model can be constructed, specifying the underlying parametric distribution (Tableman and Kim, 2004). The `survreg` syntax requires the construction of a survival object using the `Surv()` function, which becomes the response variable on the left hand side of the model. For a Weibull distribution this is given as:

```
require(survival)
{
surv.mod<-survreg(Surv(MinAge, MaxAge, CensorCode, type =
"interval")~1,
dist = "weibull")
}
```

The scale (α) and shape (λ) parameters can be calculated from the `survreg` attributes `scale` and `intercept` as:

```
alpha <- 1/(as.numeric(surv.mod$scale))
lambda <- exp(-1*as.numeric(surv.mod$coeff[1]))
```

which allow calculation of survival ($S(t)$) and hazard ($h(t)$) functions. For a leaf with as assumed maximum lifespan of $t = 150$ days, $S(t)$ and $h(t)$ for a Weibull distribution are

```
t <- seq(1, 150)
s.t <- exp(-(lambda*t)^alpha)
h.t <-lambda*alpha*(lambda*t)^(alpha-1)
```

For a log-normal distribution, $S(t)$ is calculated as

```
s.t <- 1-pnorm(alpha*log(lambda*t))
```

but $h(t)$ first requires calculation of $f(t)$, the probability density function

```
f.t <- ((2*pi)^-0.5)*(alpha*t^-1)*exp((-
alpha^2*(log(lambda*t))^2)/2)
h.t <- f.t/s.t
```

The median leaf lifespan, together with 95% confidence intervals, can be calculated as

```
med.ls<-predict(surv.mod, type = "uquantile", p = 0.5, se.fit = T)
med.ls1<-med.ls$fit[1]
med.se<-med.ls$se.fit[1]
```

```
median.lifespan<-exp(med.ls1)
CILower<-exp(med.ls1-1.96*med.se)
CIUpper<-exp(med.ls1+1.96*med.se)
```

Variables affecting leaf lifespan, such as LeafOrder, can be incorporated as

```
{
surv.mod1<-survreg(Surv(MinAge, MaxAge, CensorCode, type =
"interval")~1+LeafOrder,
dist = "weibull")
}
```

Coefficients for the parameter LeafOrder which we described as E_O in our analyis, and for the log-likelihood values of the initial model and the model incorporating LeafOrder can be determined as

```
EO<-surv.mod1$coeff[2]
LogLik.initial<-surv.mod1$loglik[1]
LogLik.LeafOrder<-surv.mod1$loglik[2]
```

with the effect of leaf order tested with a likelihood ratio test as

```
test<-(-2*LogLik.initial) - (-2*LogLik.LeafOrder)
critval<-qchisq(0.95,1)
```

If `test>critval`, there is evidence that leaf order has a significant effect on leaf lifespan.

References

Dungan RJ, Duncan RP, Whitehead D. 2003. Investigating leaf lifespans with interval-censored failure-time analysis. *New Phytologist* **158**: 593–600.

Fox GA. 2001. Failure-time analysis: Studying times to events and rates at which events occur. In: Scheiner SM, Gurevitch J. eds. Design and analysis of ecological experiments. 2nd. Edition. Oxford University Press, Oxford. 235-266.

Groeneboom P, Wellner JA. 1992. *Information bounds and non-parametric maximum likelihood estimation*. Basel: Birkhäuser Verlag.

Tableman M. Kim JS. 2004. *Survival analysis using S: analysis of time-to-event data.* Texts in Statistical Sciences. Chapman & Hall/CRC, Boca Raton, FL, USA.

Figure 1. Survival ($S(t)$) and hazard ($h(t)$) functions for two species with contrasting leaf longevity, *Aegilops geniculata* (a perennial grass with L_{FT} = 60 days) and *Juniperus oxycedrus*, an evergreen tree with L_{FT} = 528 days. For each species the functions are plotted assuming a Weibull distribution (solid line) and a log-normal distribution (dashed line), and vertical lines indicate values for L_{FT} calculated for each distribution. The step function is $S(t)$ calculated from the non-parametric maximum likelihood estimate of the underlying distribution.