Physical Model for Tree Volume

Jonathon Hirschi

Intro

It is a common practice in forestry to estimate tree volume from measurements of tree diameter (DBH) and height. This topic is of key interest to researchers and industry. In forestry, it is economically valuable to have a reliable estimate of usable timber in a stand of trees that might take years to mature. In researching climate change, it is valuable to know how much carbon trees can sequester, which is related to the volume of the tree.

Data

```r
# Environment
library(dplyr); library(ggplot2); library(latex2exp)

# Read Data
data(trees, package = "datasets")

# Rename columns
trees <- trees |>
  rename(
    d = Girth, h = Height, v = Volume
  )

# Transform Units
trees <- trees |>
  mutate(
    d = d/12
  )
```

Modeling

```r
lm_fit <- lm(v ~ d+h, data = trees)
print(paste0('Loss = ', round(deviance(lm_fit), 2)))
```

Standard linear regression:

```
## [1] "Loss = 421.92"
```

Cone

$$V = \pi r^3 \frac{h}{3}$$

```
trees$r = trees$d/2 # calculate radius
r = trees$r
h = trees$h

v_cone = pi*r^2*h/3
ss_cone = sum((v_cone-trees$v)^2)
print(paste0('Loss = ', round(ss_cone, 2)))
```

```
## [1] "Loss = 865.06"
```

That made things quite a bit worse. But trees are not perfect cones, so we might be able to find a more realistic physical model.

Truncated cone A truncated cone is defined by a lower radius and an upper radius.

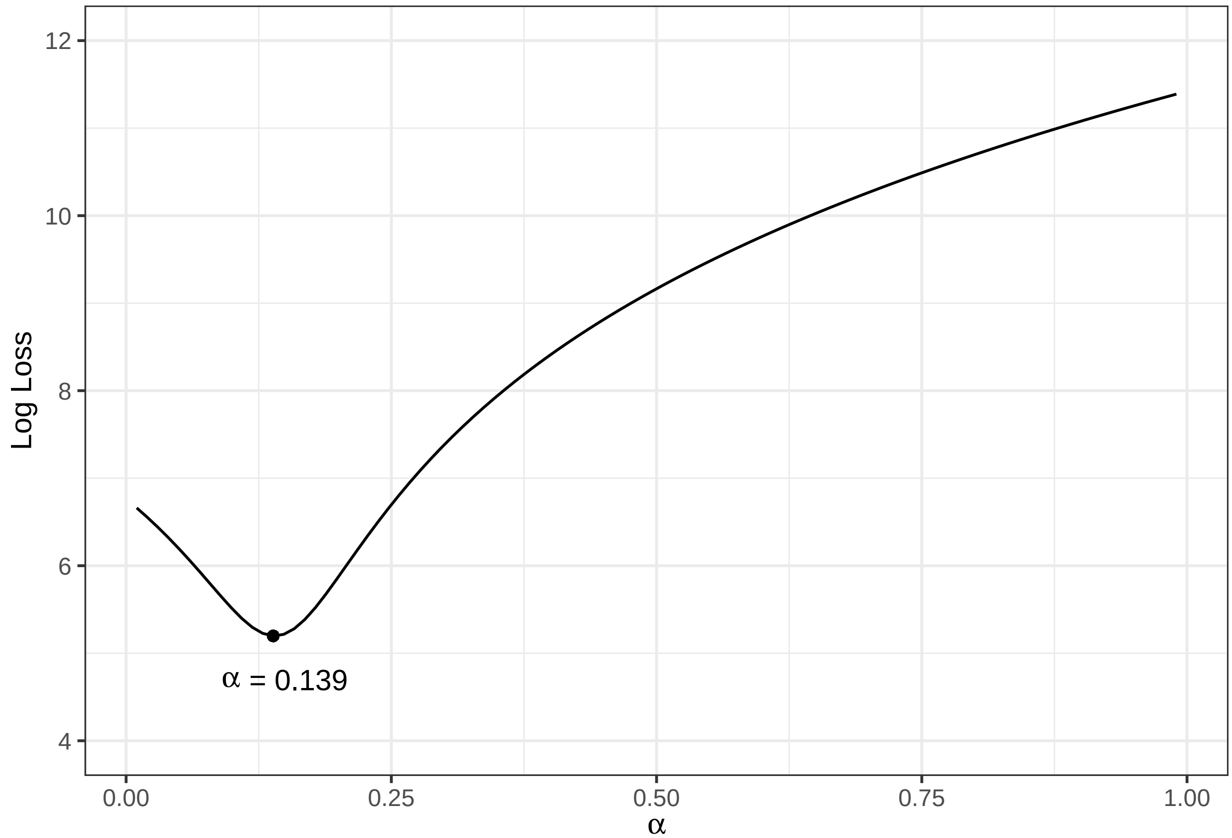
$$V = \frac{1}{3}\pi(r_1^2 + r_1 r_2 + r_2^2)h$$

Where r_1 and r_2 are the radii of the base and top of the cone, respectively.

We will consider the upper radius to be a fraction of the lower radius. So, we introduce a parameter α where $r_2 = \alpha r_1,$ for $\alpha \in (0, 1)$. We will use a grid search to find an alpha that minimizes the fitted sum of squares.

```
alpha_grid = seq(from = 0.01, to = .99, length.out=100)

ssa = rep(NA, length(alpha_grid))

for(i in 1:length(alpha_grid)){
  a = alpha_grid[i]
  r2 = trees$r*a
  v_a = 1/3*pi*(r^2+r*r2+r2^2)*h
  ssa[i]=sum((trees$v - v_a)^2)
}
```

```
df = data.frame(a = alpha_grid, ssa = ssa)
ggplot(aes(x=alpha_grid, y = log(ssa)), data=df) +
  geom_line()+
  geom_point(
    aes(
      x = alpha_grid[which.min(ssa)],
      y = log(min(ssa)))
    ) +
  labs(x = expression(alpha), y = 'Log Loss')+
  annotate('text', x = alpha_grid[which.min(ssa)]-.04,y = log(min(ssa))-.5, label=paste0(expression(alp
  annotate('text', x = alpha_grid[which.min(ssa)]+.02,y = log(min(ssa))-.5, label=paste0(' = ', round(al
  ylim(4, 12)+
  theme_bw()
```

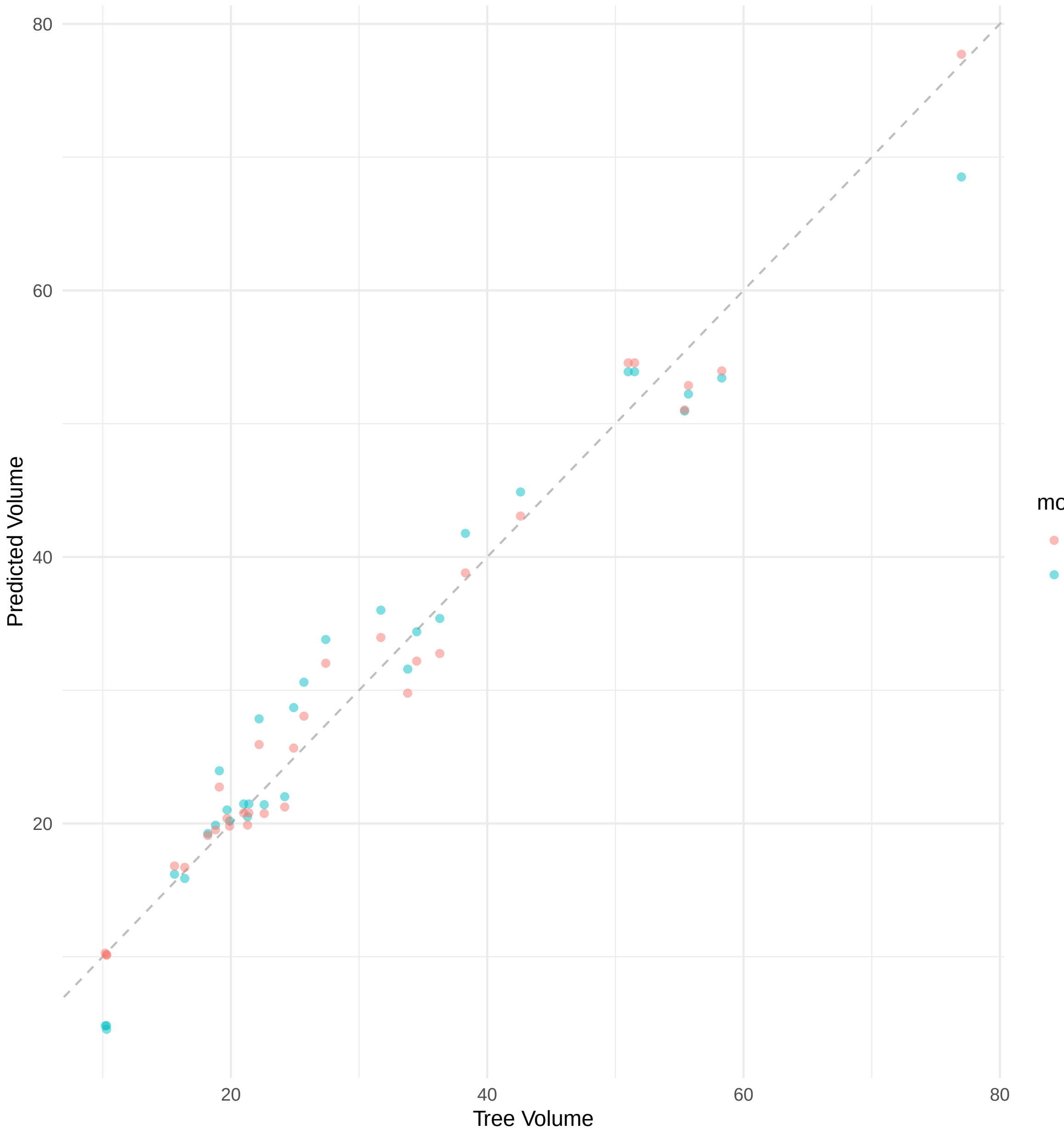



```
a <- alpha_grid[which.min(ssa)]
r = trees$r
h = trees$h
r2 <- r*a
v_a <- 1/3*pi*(r^2+r*r2+r2^2)*h
ssa <- sum((trees$v-v_a)^2)

print(paste0('Alpha = ', round(a,2))); print(paste0('Loss = ', round(ssa,2)))
```

```
## [1] "Alpha = 0.14"

## [1] "Loss = 180.9"
```

```
df = data.frame(
  v = c(trees$v, trees$v),
  pred = c(lm_fit$fitted.values, v_a),
  model = c(rep("lm", nrow(trees)), rep("cone", nrow(trees)))
)
df |>
  ggplot(aes(x=v, y=pred, color=model))+
  geom_abline(slope=1, intercept = 0, lty=2, color="grey")+
  geom_point(alpha=.5)+
```

```
  labs(x="Tree Volume", y="Predicted Volume")+
  theme_minimal()
```



Comparing Models

Model from paper Paper: https://www.researchgate.net/publication/318780019_Modeling_Height-Diameter_Relationship_and_Volume_of_Teak_Tectona_grandis_L_F_in_Central_Lowlands_of_Nepal

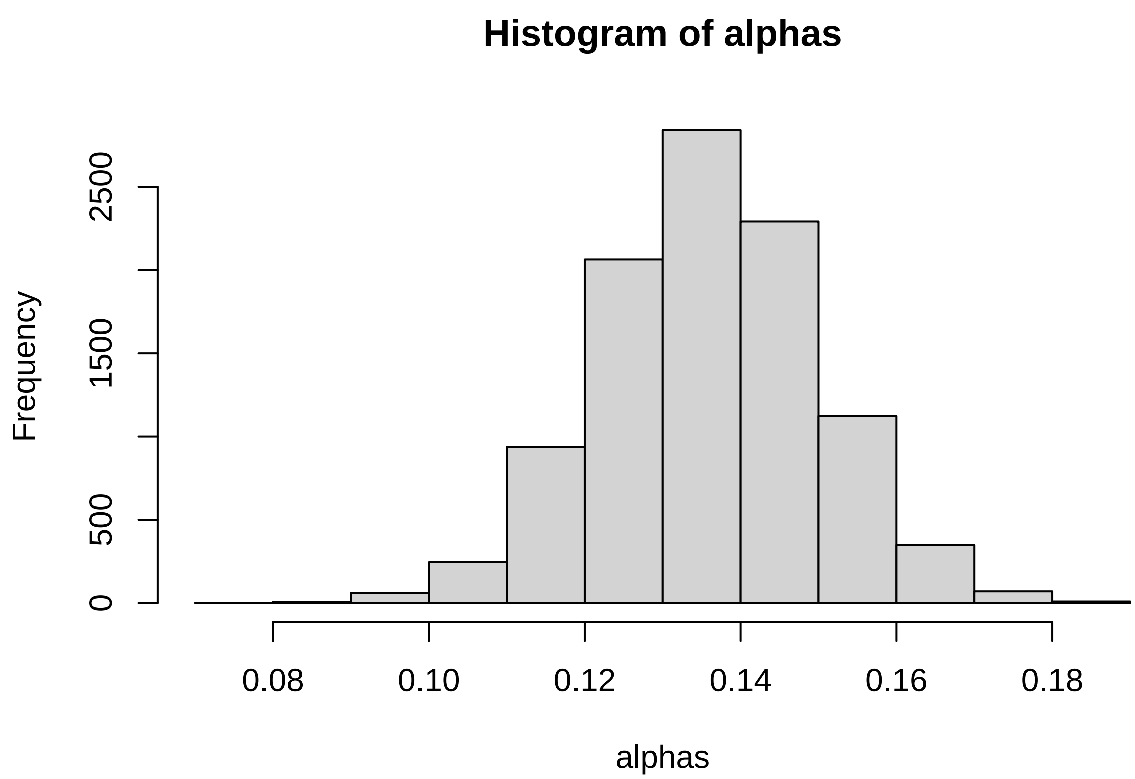
$$V = \beta_0 + \beta_1 \cdot d + \beta_2 \cdot d^2 \cdot h$$

```
trees$d2h <- trees$d^2*trees$h
fit2 <- lm(v~d+d2h, data=trees)
deviance(fit2)
```

```
## [1] 179.7975
```

Quantifying Uncertainty We employ a Monte Carlo method to estimate the uncertainty in the α parameter. For N bootstrap samples, find α_n for each $n \in N$, then calculate quantiles based on this set of values.

```
set.seed(87337) # rand num: "trees" in text
nreps <- 10000
alphas <- rep(NA, nreps)
for(j in 1:nreps){
  trees_boot <- trees[
      sample(1:nrow(trees), replace = T),
    ]
  for(i in 1:length(alpha_grid)){
    a = alpha_grid[i]
    h = trees_boot$h
    r = trees_boot$r
    r2 = r*a
    v_a = 1/3*pi*(r^2+r*r2+r2^2)*h
    ssa[i]=(trees_boot$v - v_a)%*%(trees_boot$v - v_a)
  }
  alphas[j] <- alpha_grid[which.min(ssa)]
}

hist(alphas)
```



Theoretical Calculation Let the lower tree radius be r and the upper be r_2. Since $r_2 = \alpha r$, we can write the volume as:

$$V = \frac{1}{3} \cdot h \cdot \pi \cdot r^2(1 + \alpha + \alpha^2)$$

Then, we can minimize the RSS with respect to α. Let v_i be the true volume for tree i and let $\hat{v}_i$ be the estimated tree volume for tree i using the above formula:

$$RSS = \sum_{i=1}^{N}(v_i - \hat{v}_i)^2 = \sum_{i=1}^{N}(v_i - \frac{1}{3}\pi h_i r_i^2(1 + \alpha + \alpha^2))^2$$

To simplify the notation, let $x_i = \frac{1}{3}\pi h_i r_i^2,$ for $i = 1, 2, ..., N$.

$$RSS = \sum_{i=1}^{N}(v_i - x_i(1 + \alpha + \alpha^2))^2$$

Then,

$$\begin{aligned}
\frac{\partial RSS}{\partial \alpha} &= \sum -2(1 + 2\alpha)x_i(v_i - x_i(1 + \alpha + \alpha^2)) && (1)\\
&= -2(1 + 2\alpha) \sum x_i(v_i - x_i(1 + \alpha + \alpha^2)) && (2)\\
\text{Set } \rightarrow 0 &= \sum x_i(v_i - x_i(1 + \alpha + \alpha^2)) && (3)\\
0 &= \sum v_i x_i - (1 + \alpha + \alpha^2) \sum x_i^2 && (4)
\end{aligned}$$

We employ the quadratic formula to get:

```
trees$x = 1/3*trees$h*trees$r^2*pi
```