

total_tube_growth_results

Gabrielle Languedoc

25/11/2021

```
library(mgcv)
library(fitdistrplus)
library(goft)
library(gamlss)
library(here)
library(patchwork)
library(LambertW)
library(scales)
library(viridis)
library(performance)
library(dplyr)
library(tidyverse)
library(readr)

# import data set
tube_growth <- read_csv("tube-growth.csv")
```

Load required libraries and data set

```
# Manipulated data for intial visualization
clean_growth <- tube_growth %>%
select(total_growth, treatment)

# Manipulated data for GAMLSS Model (Important note; Could not get the model
# to run using worm_id and string_loss as factors.
#This is likely because of the small sample size (N=90) creating to much noise)
model_growth <- tube_growth %>%
  dplyr::select(c(treatment,total_growth,mortality, amt_cut_mm, worm_id)) %>%
  mutate(worm_id = as.factor(worm_id))
```

Cleaning Data Sheet

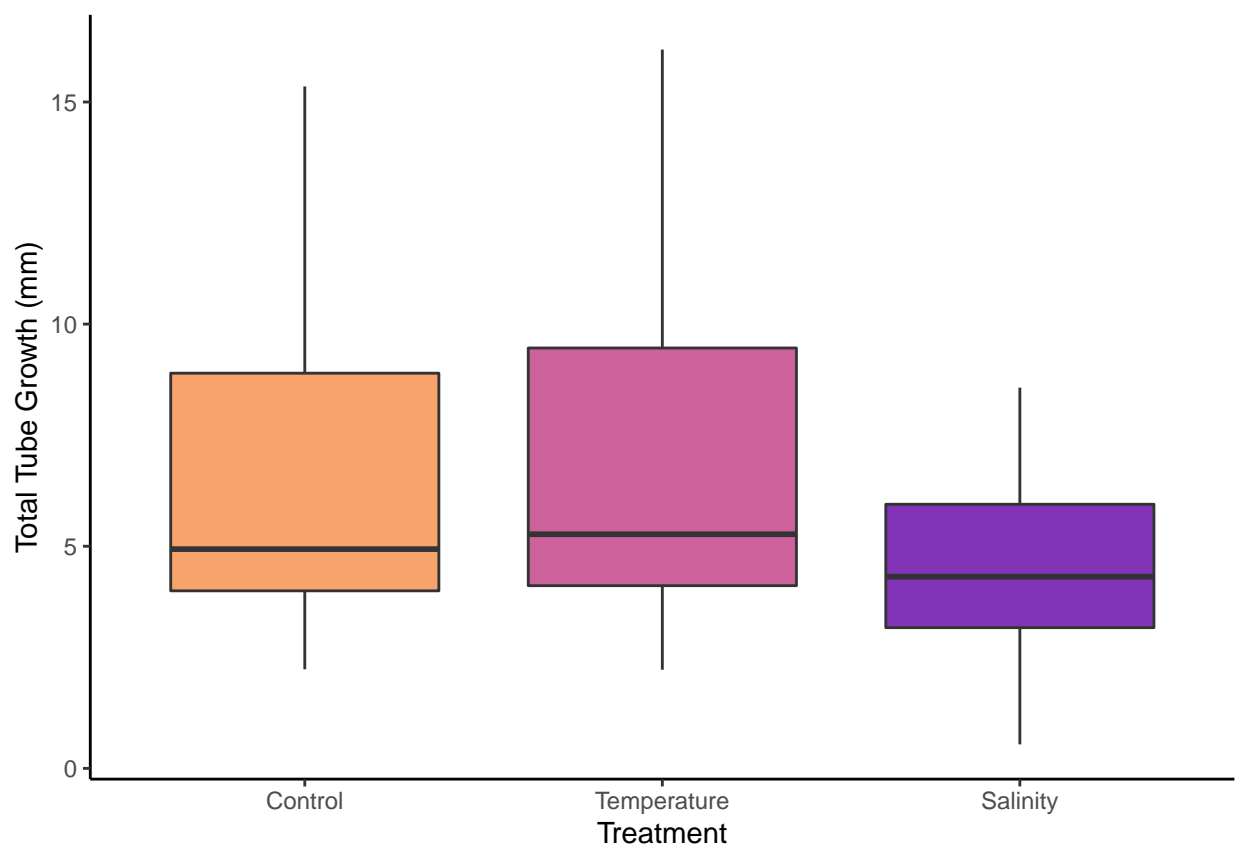
Visualizing Total Tube Growth Data with a boxplot

```

#colour version
my_cols = c("low_salinity" = "#6300A7FF", "high_temp" = "#C03A83FF", "control"
            = "#F58C46FF")

ggplot(clean_growth, aes(x = treatment, y = total_growth, fill = treatment)) +
  geom_boxplot(alpha = 0.8) +
  theme_classic() +
  ylab("Total Tube Growth (mm)") +
  xlab("Treatment") +
  scale_fill_manual(values = my_cols) +
  scale_colour_manual(values = my_cols) +
  theme(legend.position="none") +
  scale_x_discrete(labels = c('Control', 'Temperature', 'Salinity'))

```



```

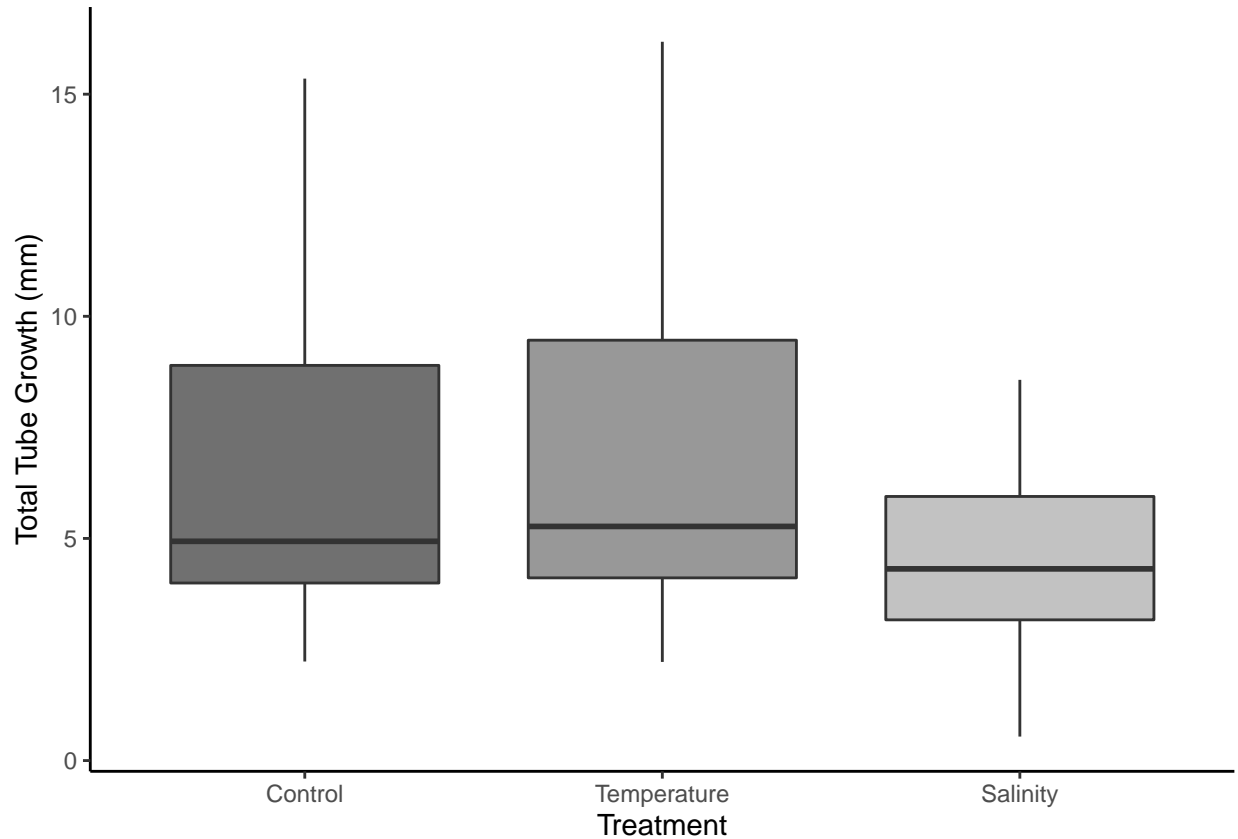
ggsave(plot = last_plot(), filename =
       "total_tube_growth_boxplot_colourized.pdf")

#grey scale version
my_cols_grey = c("low_salinity" = "grey70", "high_temp" = "grey50", "control"
                = "grey30")

ggplot(clean_growth, aes(x = treatment, y = total_growth, fill = treatment)) +
  geom_boxplot(alpha = 0.8) +
  theme_classic() +
  ylab("Total Tube Growth (mm)") +
  xlab("Treatment") +

```

```
scale_fill_manual(values = my_cols_grey) +
scale_colour_manual(values = my_cols_grey) +
theme(legend.position="none") +
scale_x_discrete(labels = c('Control', 'Temperature', 'Salinity'))
```



```
ggsave(plot = last_plot(), filename =
       "total_tube_growth_boxplot_greyscale.pdf")
```

Finding Mean Values for Each Treatment

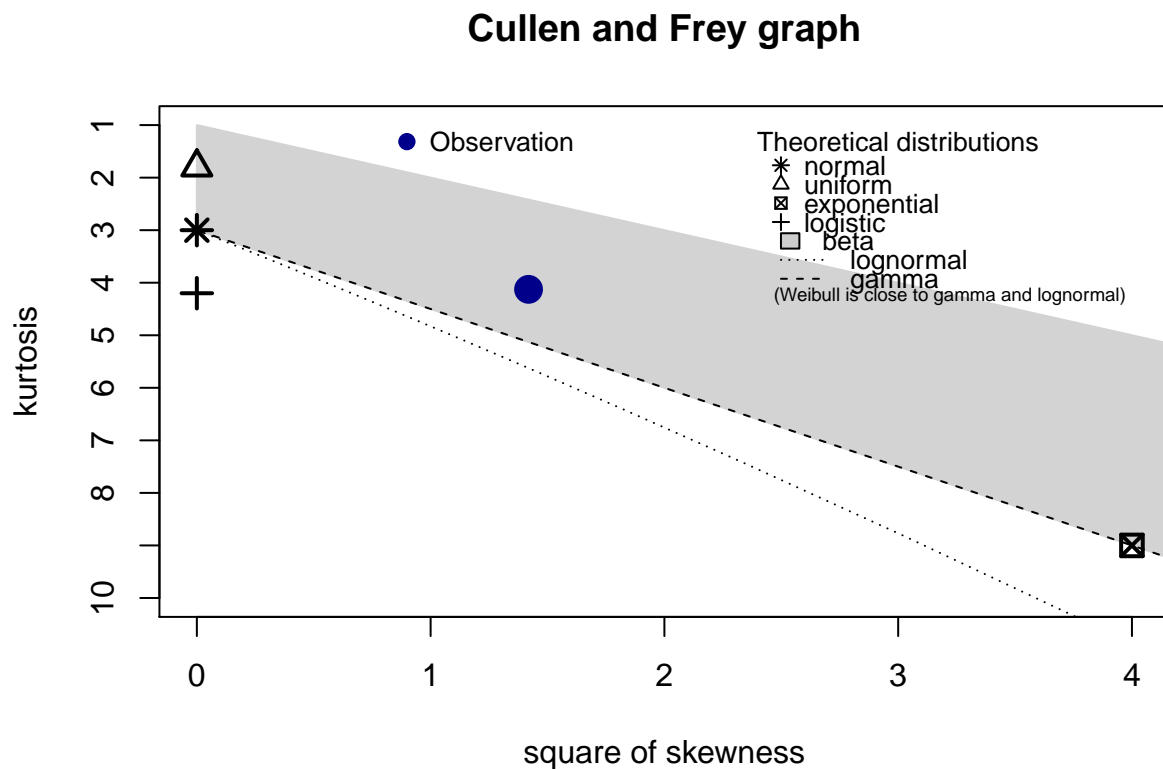
```
mean_growth <- clean_growth %>%
group_by(treatment) %>%
summarise_at(vars(total_growth), list(mean = mean))
```

Checking Distributions For Normality

```
# Original
shapiro.test(clean_growth$total_growth) # P value = 2.292e-06
```

```
##
## Shapiro-Wilk normality test
##
## data: clean_growth$total_growth
## W = 0.89439, p-value = 2.292e-06
```

```
descdist(clean_growth$total_growth)
```



```
## summary statistics
## -----
## min: 0.54 max: 16.18
## median: 4.92
## mean: 5.879944
## estimated sd: 3.207908
## estimated skewness: 1.191188
## estimated kurtosis: 4.126008
```

```
# trying log transform
```

```
clean_log <- clean_growth %>%
  mutate(loggrowth = log2(total_growth))
shapiro.test(clean_log$loggrowth) # P value 0.006626
```

```
##
```

```
## Shapiro-Wilk normality test
##
## data: clean_log$loggrowth
## W = 0.95936, p-value = 0.006626
```

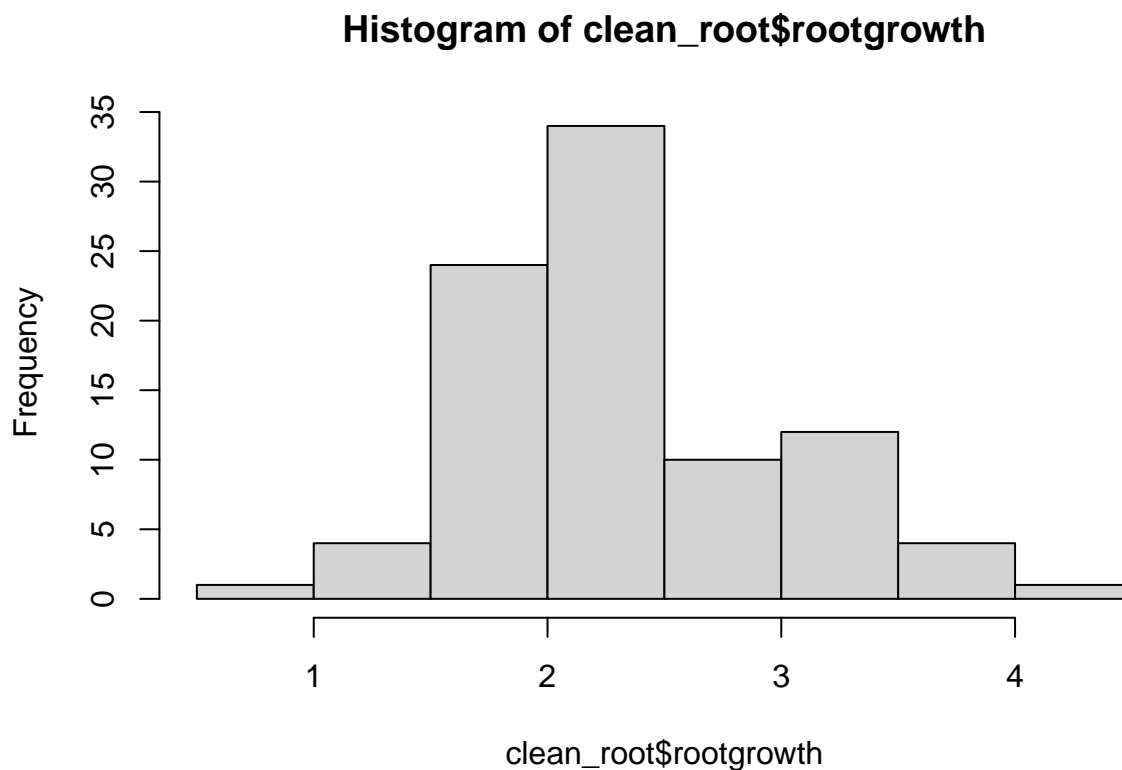
```
# trying to square root transform
```

```
clean_root <- clean_growth %>%
  mutate(rootgrowth = sqrt(total_growth))
shapiro.test(clean_root$rootgrowth) # P value = 0.008449
```

```
##
## Shapiro-Wilk normality test
##
## data: clean_root$rootgrowth
## W = 0.96095, p-value = 0.008449
```

```
# All attempts show our data varies significantly from normal,
#now trying fist dist function to find the best distribution.
```

```
hist(clean_root$rootgrowth)
```



```
fitDist(rootgrowth, data = clean_root, type = "realAll", try.gamlss = T)
```

```
## |
```

```
|
```

```
##
## Family: c("SEP2", "Skew Exponential Power type 2")
## Fitting method: "nlminb"
##
## Call: gamlssML(formula = y, family = DIST[i])
##
## Mu Coefficients:
## [1] 1.728
## Sigma Coefficients:
## [1] -0.2874
## Nu Coefficients:
## [1] 1.427
## Tau Coefficients:
## [1] 0.2958
##
## Degrees of Freedom for the fit: 4 Residual Deg. of Freedom 86
## Global Deviance: 161.898
## AIC: 169.898
## SBC: 179.897

fitDist(total_growth, data = clean_growth, type = "realAll", try.gamlss = T)
```

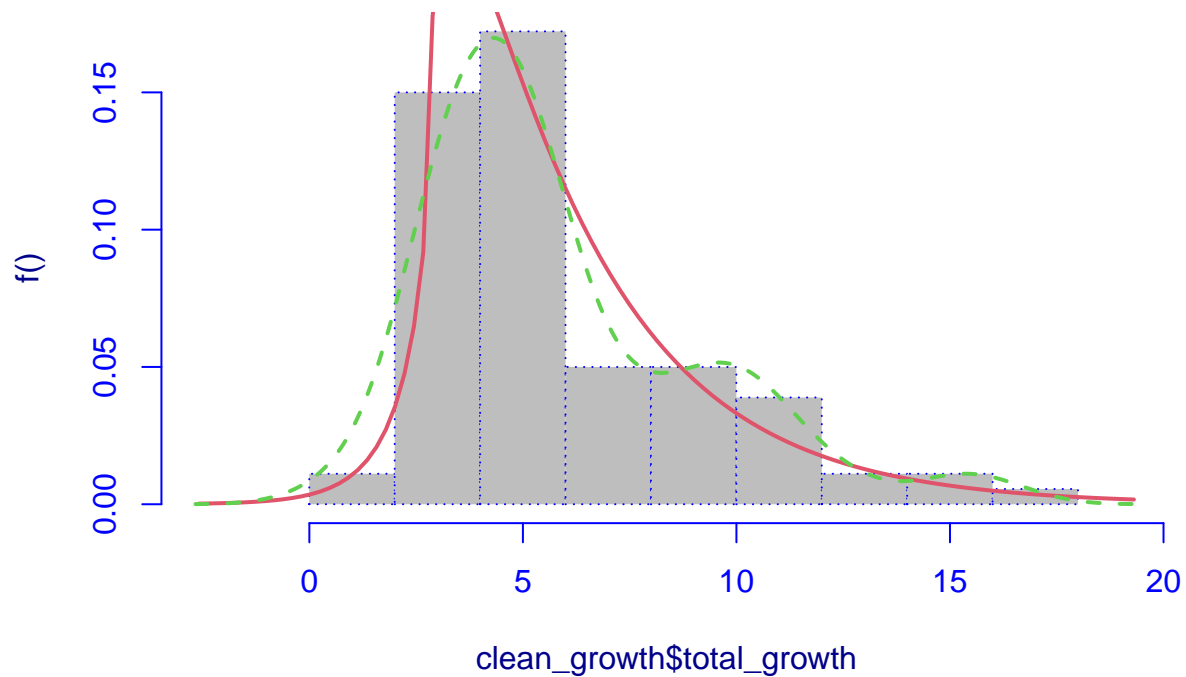
```
## |
```

```
|
```

```
### Both show the best distribution is Skew Exponential Power type 2

mSEP2 <- histDist(clean_growth$total_growth, "SEP2", density = T, main =
  "Skew exp type 2")
```

Skew exp type 2



Running a Gamlss Model

```
gamlss.mod <- gamlss(total_growth ~ treatment + amt_cut_mm + mortality,
  family = SEP2(),
  data = model_growth,
  control = gamlss.control(n.cyc = 25))
```

Modelling

```
## GAMLSS-RS iteration 1: Global Deviance = 453.4377
## GAMLSS-RS iteration 2: Global Deviance = 453.1597
## GAMLSS-RS iteration 3: Global Deviance = 452.7278
## GAMLSS-RS iteration 4: Global Deviance = 452.0642
## GAMLSS-RS iteration 5: Global Deviance = 450.5416
## GAMLSS-RS iteration 6: Global Deviance = 446.3317
## GAMLSS-RS iteration 7: Global Deviance = 440.7388
## GAMLSS-RS iteration 8: Global Deviance = 436.1565
## GAMLSS-RS iteration 9: Global Deviance = 434.69
## GAMLSS-RS iteration 10: Global Deviance = 434.5907
## GAMLSS-RS iteration 11: Global Deviance = 431.9027
## GAMLSS-RS iteration 12: Global Deviance = 431.7732
## GAMLSS-RS iteration 13: Global Deviance = 431.7615
## GAMLSS-RS iteration 14: Global Deviance = 431.7486
```

```
## GAMLSS-RS iteration 15: Global Deviance = 431.7276
## GAMLSS-RS iteration 16: Global Deviance = 430.3352
## GAMLSS-RS iteration 17: Global Deviance = 429.6759
## GAMLSS-RS iteration 18: Global Deviance = 429.2198
## GAMLSS-RS iteration 19: Global Deviance = 429.195
## GAMLSS-RS iteration 20: Global Deviance = 429.1835
## GAMLSS-RS iteration 21: Global Deviance = 429.1772
## GAMLSS-RS iteration 22: Global Deviance = 429.1733
## GAMLSS-RS iteration 23: Global Deviance = 429.1709
## GAMLSS-RS iteration 24: Global Deviance = 429.169
## GAMLSS-RS iteration 25: Global Deviance = 429.1683
```

```
# ran after 25 iterations when we removed worm_id
# as a random effect as per Dans advice
```

```
# Checking model
```

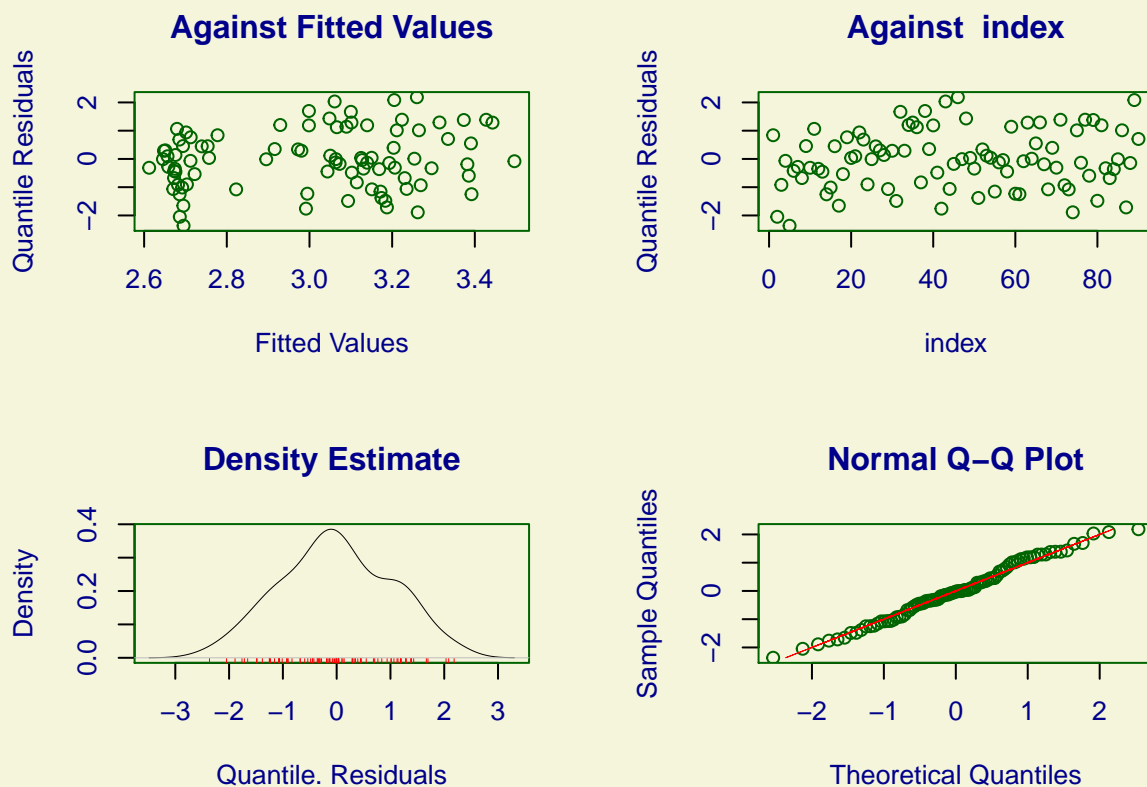
```
summary(gamlss.mod)
```

```
## *****
## Family:  c("SEP2", "Skew Exponential Power type 2")
##
## Call:  gamlss(formula = total_growth ~ treatment + amt_cut_mm +
##      mortality, family = SEP2(), data = model_growth,
##      control = gamlss.control(n.cyc = 25))
##
## Fitting method: RS()
##
## -----
## Mu link function:  identity
## Mu Coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.53051    3.19413   0.792   0.431
## treatmenthigh_temp -0.22516    0.57646  -0.391   0.697
## treatmentlow_salinity -0.58452    0.48038  -1.217   0.227
## amt_cut_mm         0.04937    0.18162   0.272   0.786
## mortality         -0.47564    0.86336  -0.551   0.583
##
## -----
## Sigma link function:  log
## Sigma Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.1062    0.1435   7.706 2.66e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## -----
## Nu link function:  identity
## Nu Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.3141    0.3875   3.391 0.00107 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```



```
## -----
## Tau link function:  log
## Tau Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.009221  0.187175  -0.049   0.961
## -----
## No. of observations in the fit:  90
## Degrees of Freedom for the fit:   8
##      Residual Deg. of Freedom:  82
##                               at cycle: 25
##
## Global Deviance:      429.1683
##           AIC:        445.1683
##           SBC:        465.1668
## *****
```

```
plot(gamlss.mod)
```



```
## *****
##      Summary of the Quantile Residuals
##              mean    = -0.008825433
##              variance = 1.044282
##              coef. of skewness = 0.01144696
##              coef. of kurtosis = 2.371234
```

```
## Filliben correlation coefficient = 0.9951015
```

```
## *****
```

```
residuals(gamlss.mod)
```

```
## [1] 0.838960862 -2.047339013 -0.916710769 -0.067464758 -2.364455626
## [6] -0.422950781 -0.281456827 -0.677165248 0.455626250 -0.312873678
## [11] 1.064725857 -0.349493941 -0.452182053 -1.252259762 -1.012323587
## [16] 0.449801457 -1.654619011 -0.536965177 0.766453321 0.034539369
## [21] 0.095120663 0.938065669 0.683733039 -0.901385708 -0.007972099
## [26] 0.445195612 0.303373896 0.139121611 -1.066117676 0.288316415
## [31] -1.488011156 1.667333949 0.290758974 1.200868837 1.287646114
## [36] 1.122586402 -0.830306879 1.697027090 0.349462926 1.187410707
## [41] -0.484304664 -1.759936020 2.034002880 -1.061255026 -0.175730681
## [46] 2.183857071 -0.007199967 1.431848792 0.035371520 -0.342589088
## [51] -1.381315876 0.337992601 0.115772439 0.041008263 -1.155764022
## [56] -0.122999227 -0.040834372 -0.445376254 1.139933093 -1.230697496
## [61] -1.248460965 -0.081508799 1.285579078 0.008589324 0.550034181
## [66] 1.290809698 -0.188914759 -1.076372566 0.393228763 -0.310015635
## [71] 1.386082658 -0.931032063 -1.077950260 -1.889465248 1.011927085
## [76] -0.134408652 1.388776962 -0.594228105 1.378184718 -1.481703121
## [81] 1.191151024 -0.329595938 -0.679137247 -0.362066468 -0.011803696
## [86] 1.013277669 -1.715732592 -0.144736767 2.082785294 0.706558254
```