SOM: Diagnostics of linear models

Ben Marwick

Elspeth Hayes

Chris Clarkson

Richard Fullagar

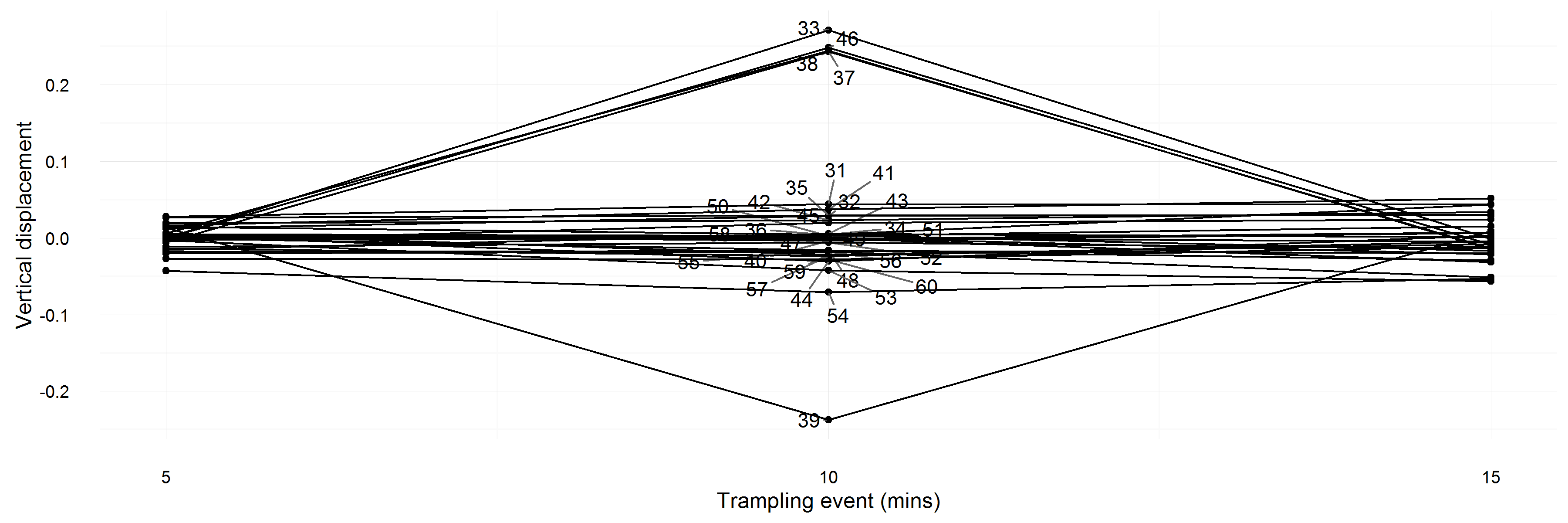
2016-10-20

This is the abstract.

## Loading required package: ggplot2

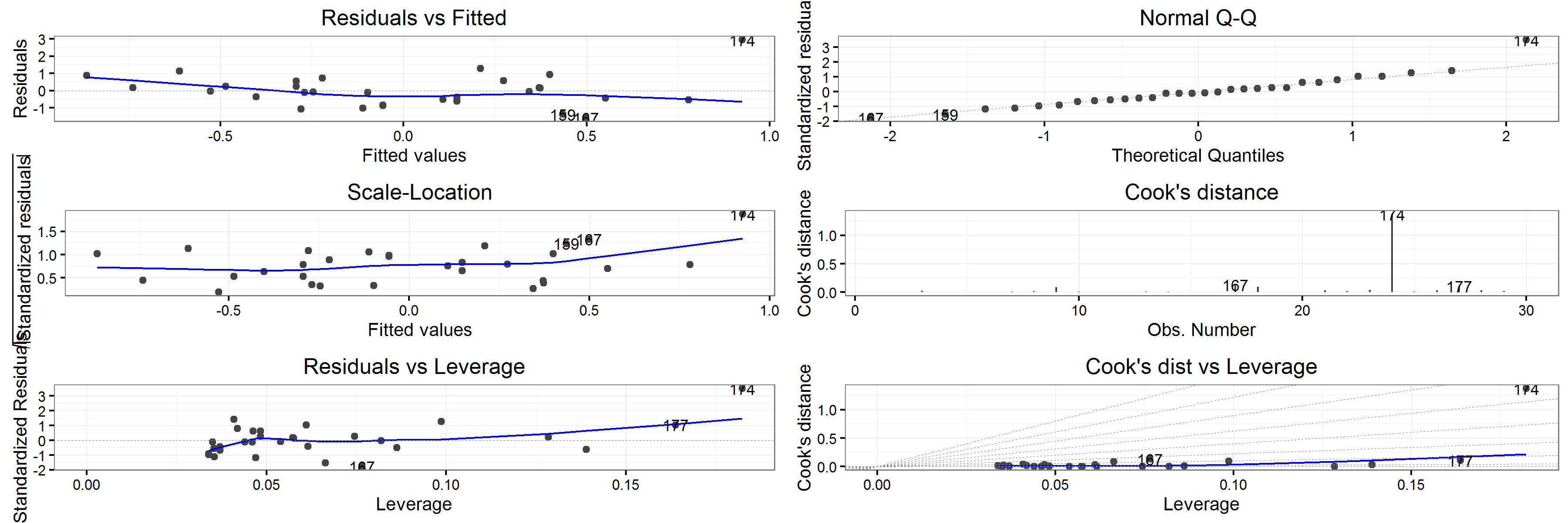
## Visualising individual artefact movements

In the plot below we can see the vertical distance each artefact has moved after each trampling event.

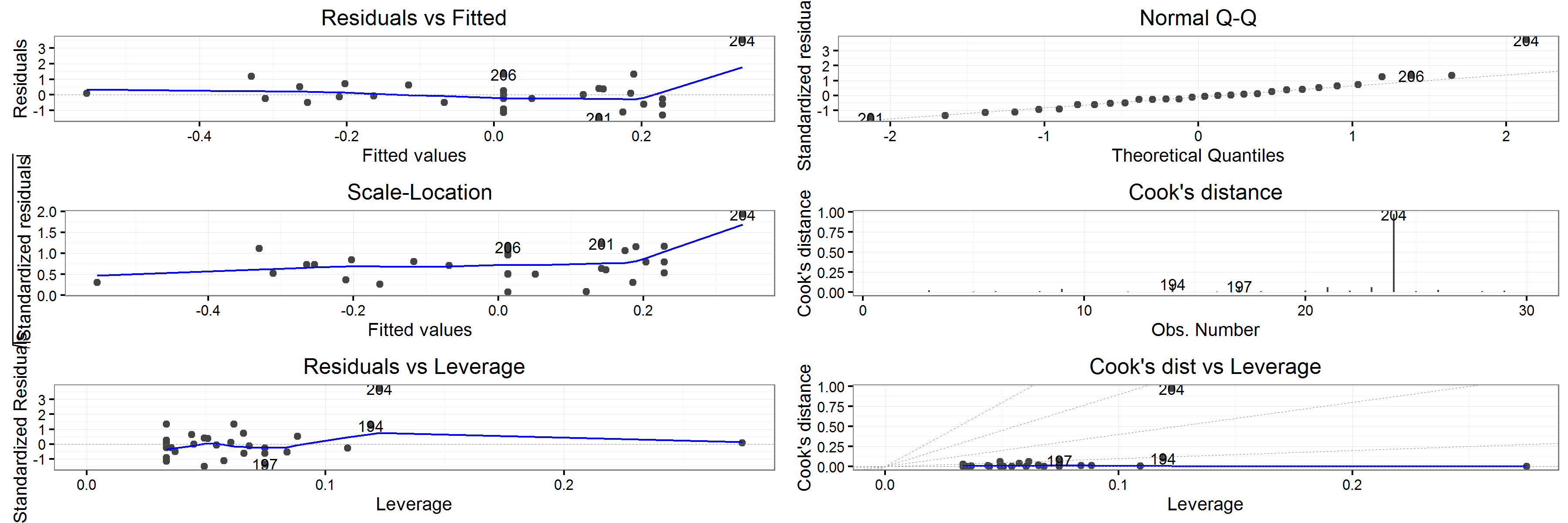


## Diagnostic plots for linear models of horizontal displacement of artefacts

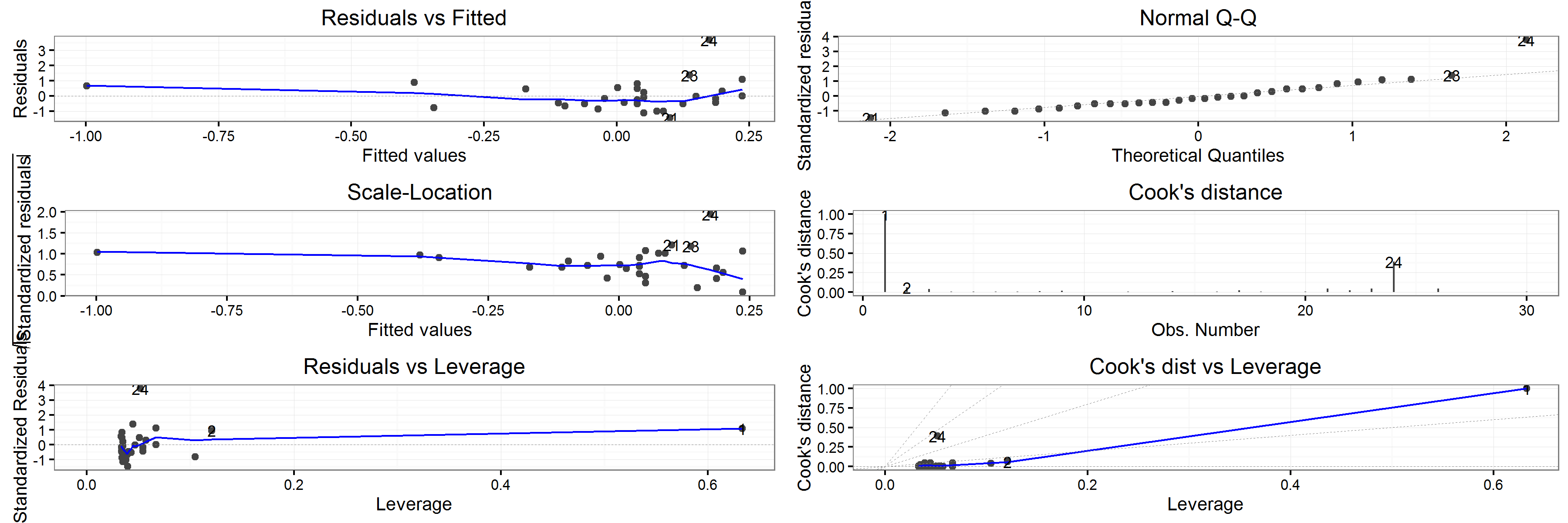
## $elongation



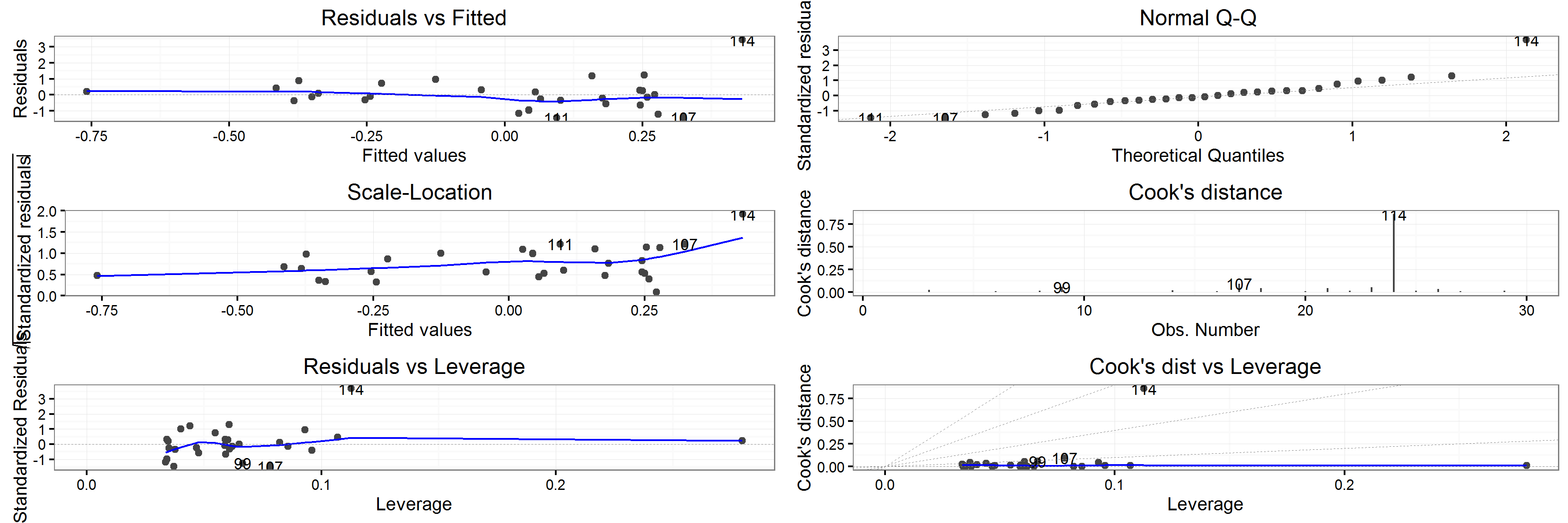
##   
## $flatness



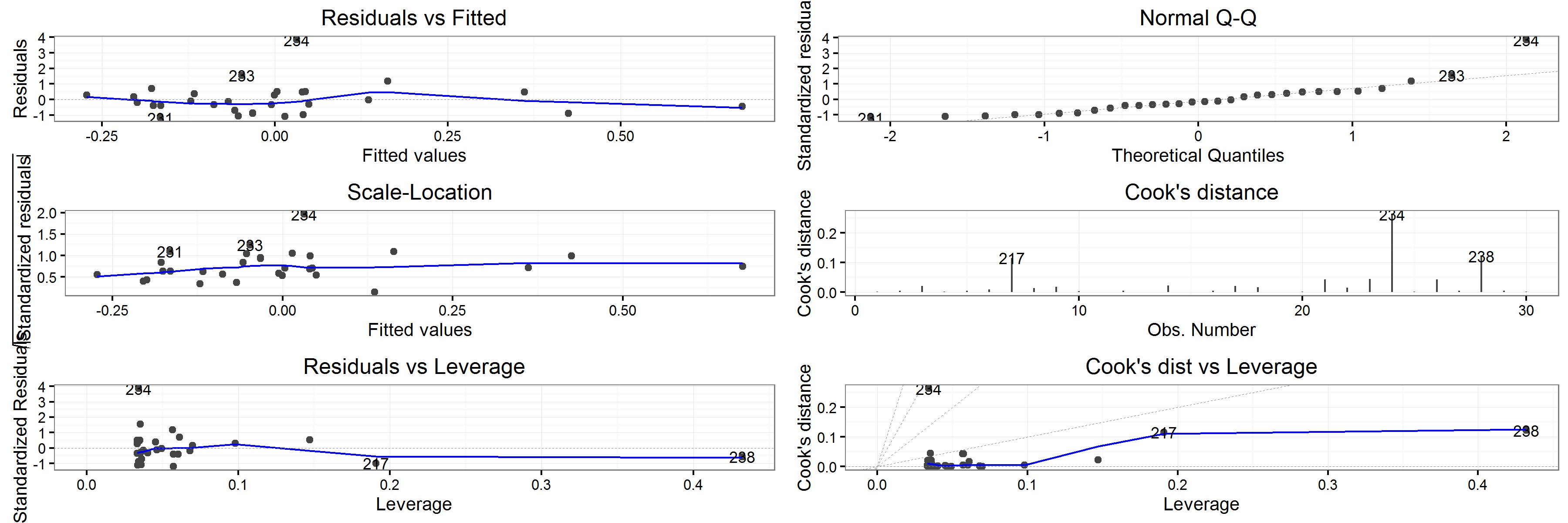
##   
## $`length(mm)`



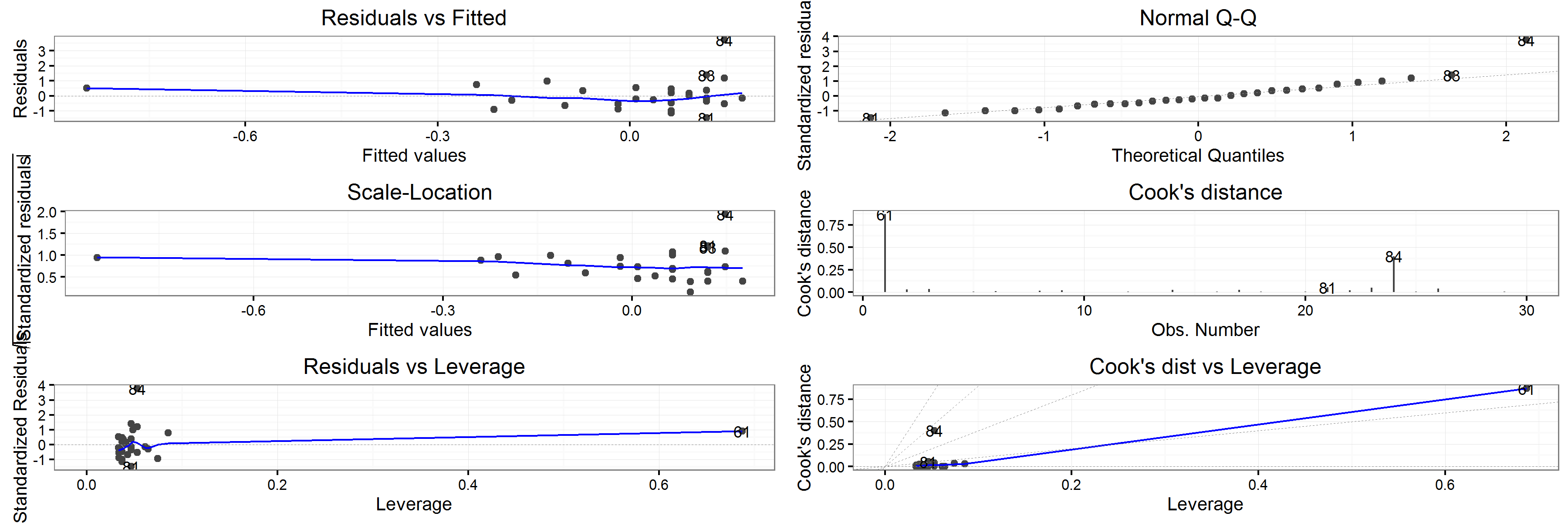
##   
## $shape



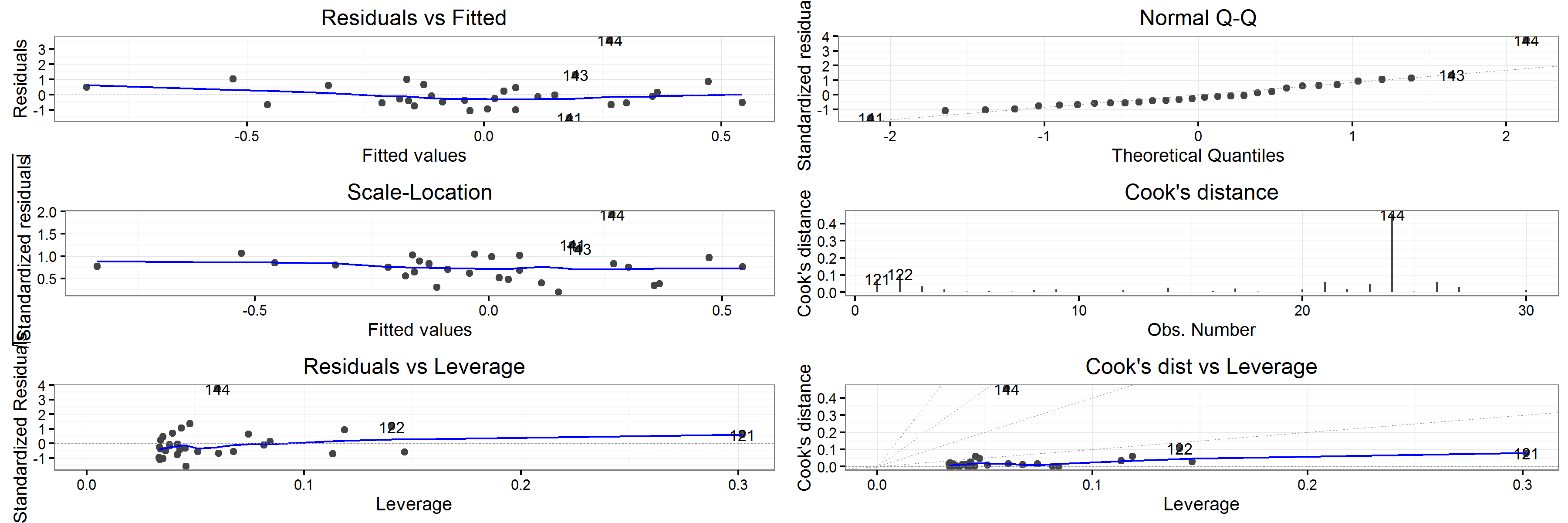
##   
## $sphericity



##   
## $`thickness(mm)`



##   
## $volume



##   
## $`width(mm)`

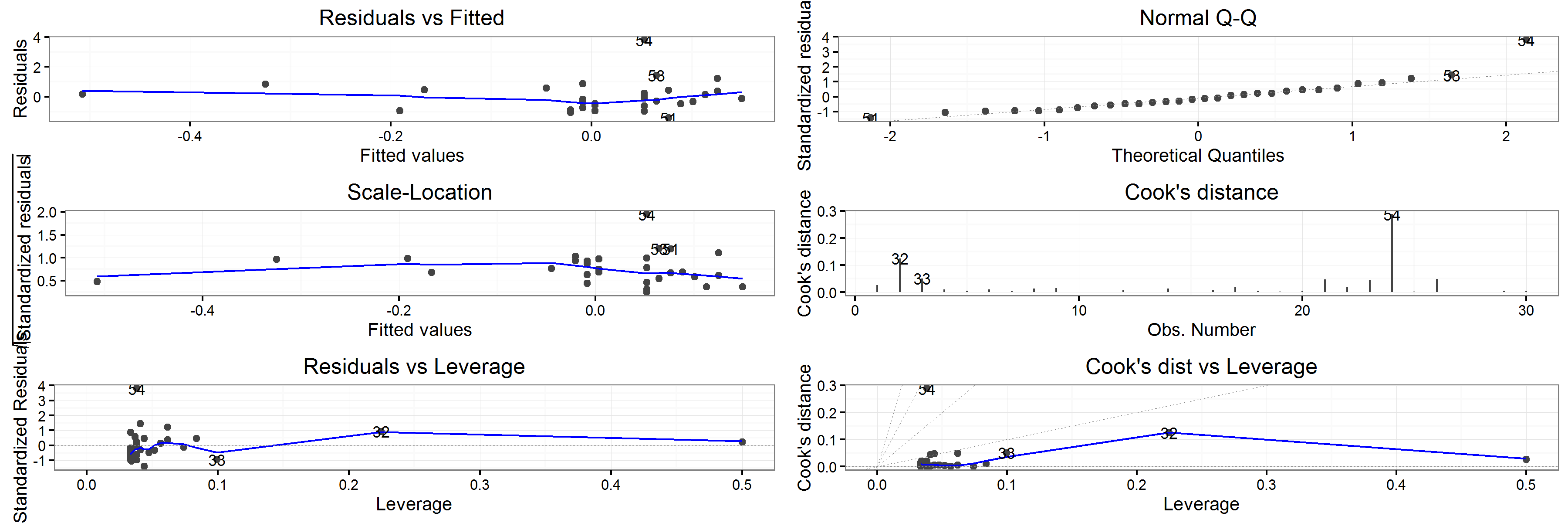
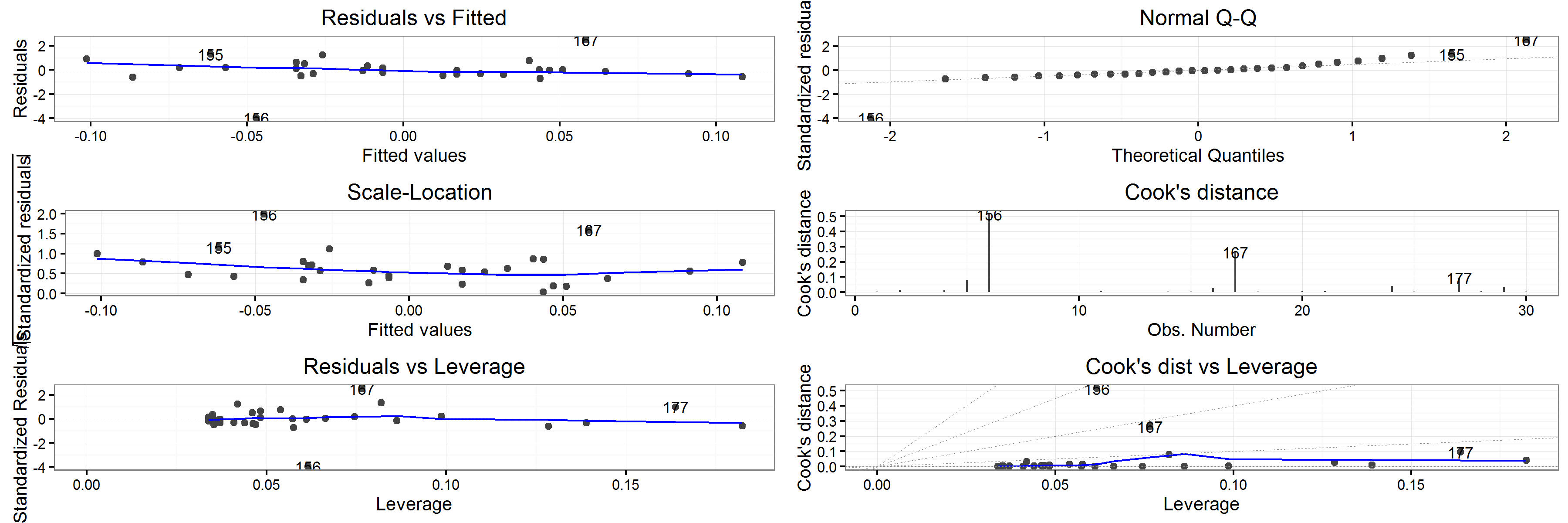


Table 1: Horizontal displacement: Correlation of artefact volume with Cook's distance values for each model

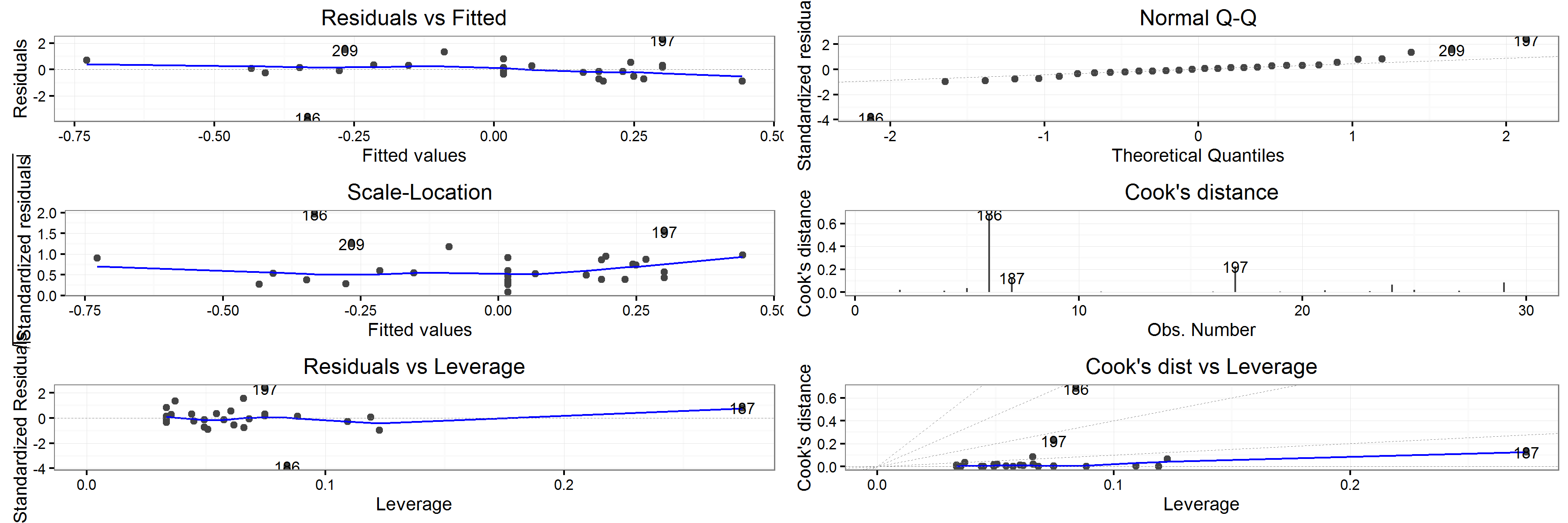
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| estimate | statistic | p.value | parameter | conf.low | conf.high | variable |
| -0.2111567 | -1.1431108 | 0.2626726 | 28 | -0.5310288 | 0.1613900 | elongation |
| -0.1779000 | -0.9566179 | 0.3469474 | 28 | -0.5057545 | 0.1948580 | flatness |
| 0.4514844 | 2.6774481 | 0.0122668 | 28 | 0.1089339 | 0.6981890 | length(mm) |
| -0.1899700 | -1.0238716 | 0.3146609 | 28 | -0.5149929 | 0.1828109 | shape |
| -0.2392109 | -1.3036325 | 0.2029761 | 28 | -0.5519158 | 0.1324751 | sphericity |
| 0.4195870 | 2.4459722 | 0.0209879 | 28 | 0.0698814 | 0.6774500 | thickness(mm) |
| -0.0550927 | -0.2919666 | 0.7724645 | 28 | -0.4072782 | 0.3113564 | volume |
| 0.0026988 | 0.0142807 | 0.9887073 | 28 | -0.3579184 | 0.3626154 | width(mm) |

## Diagnostic plots for linear models of horizontal direction of artefacts

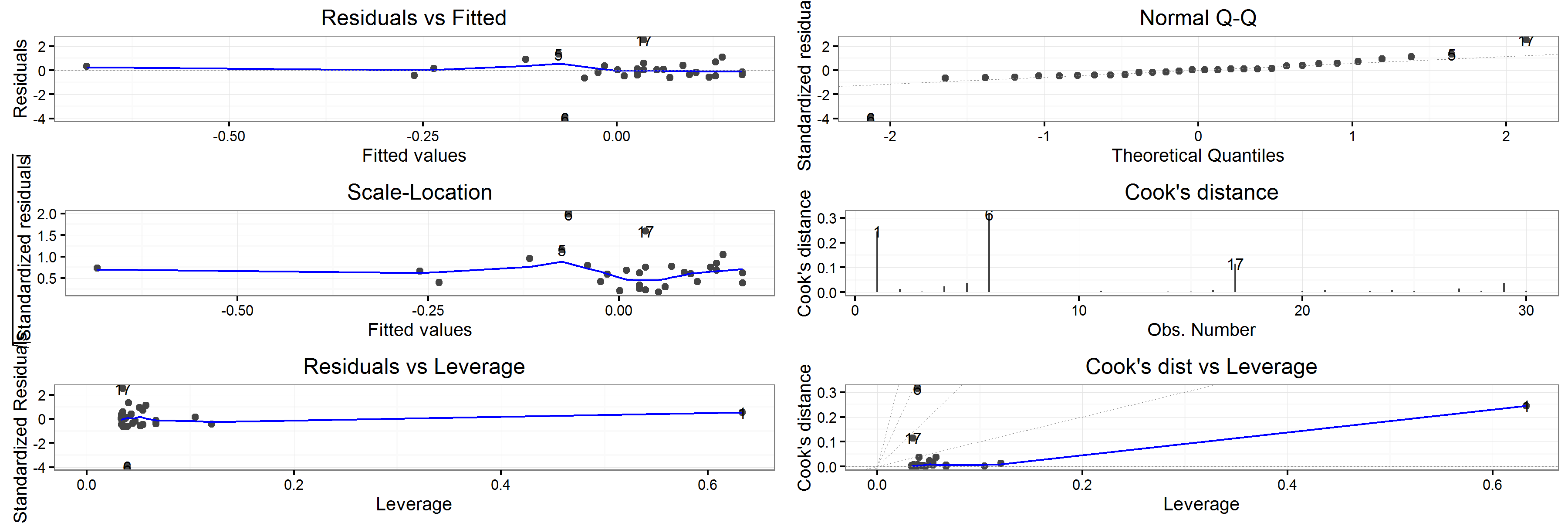
## $elongation



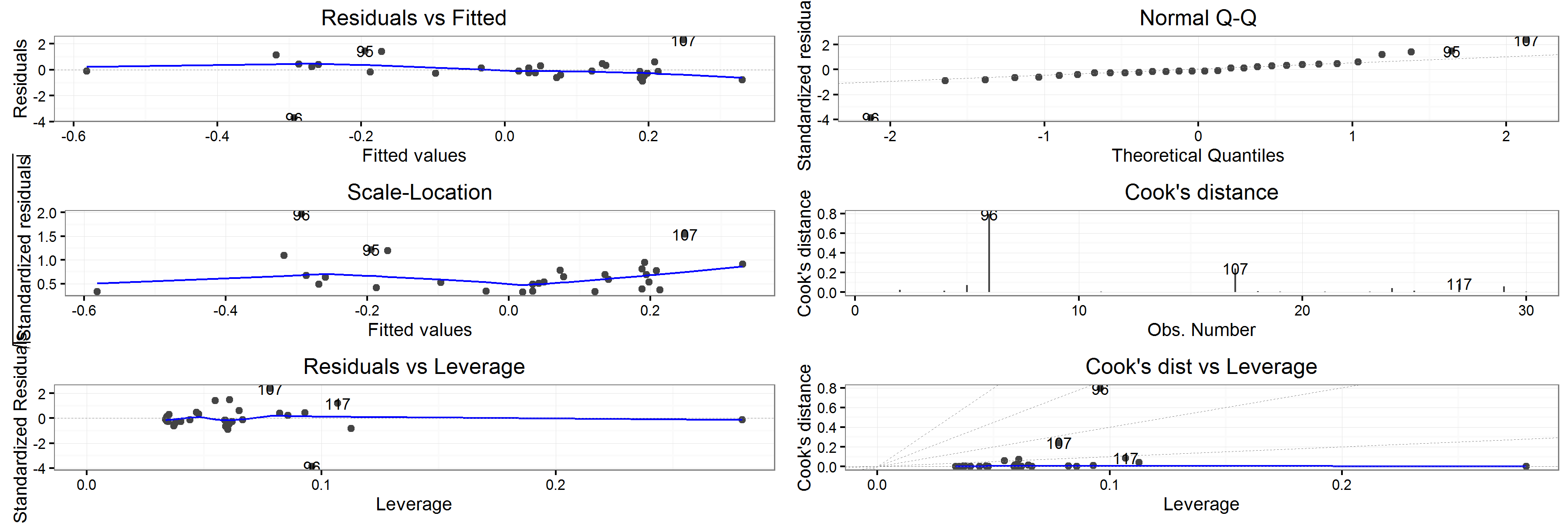
##   
## $flatness



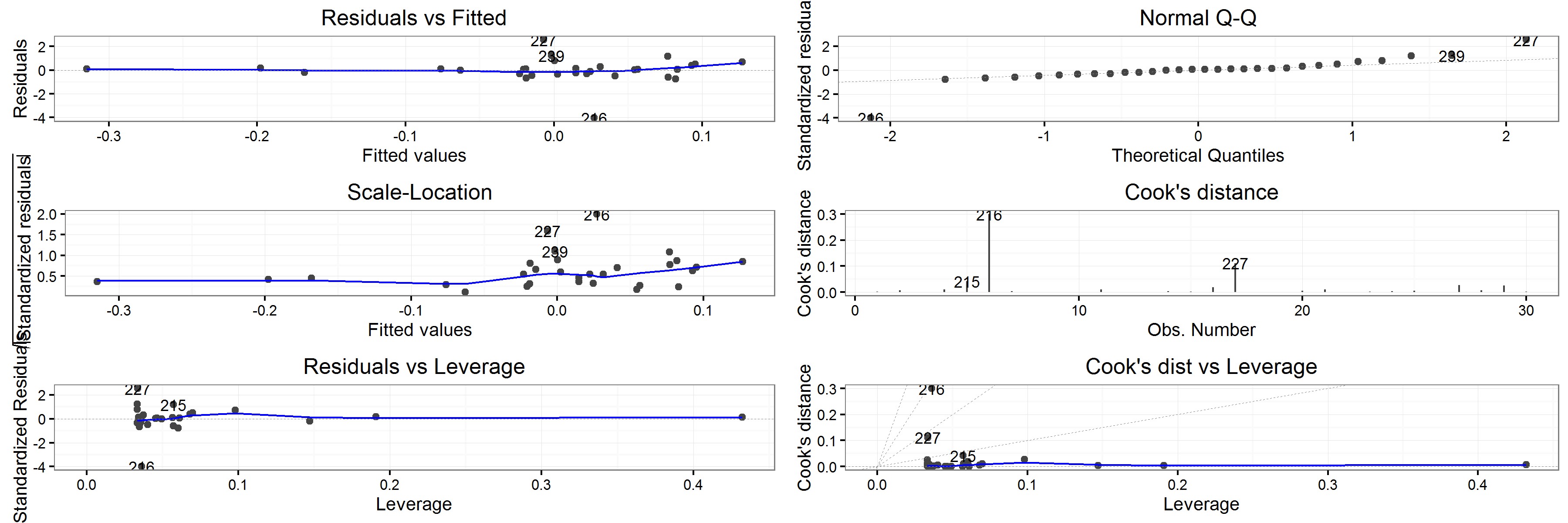
##   
## $`length(mm)`



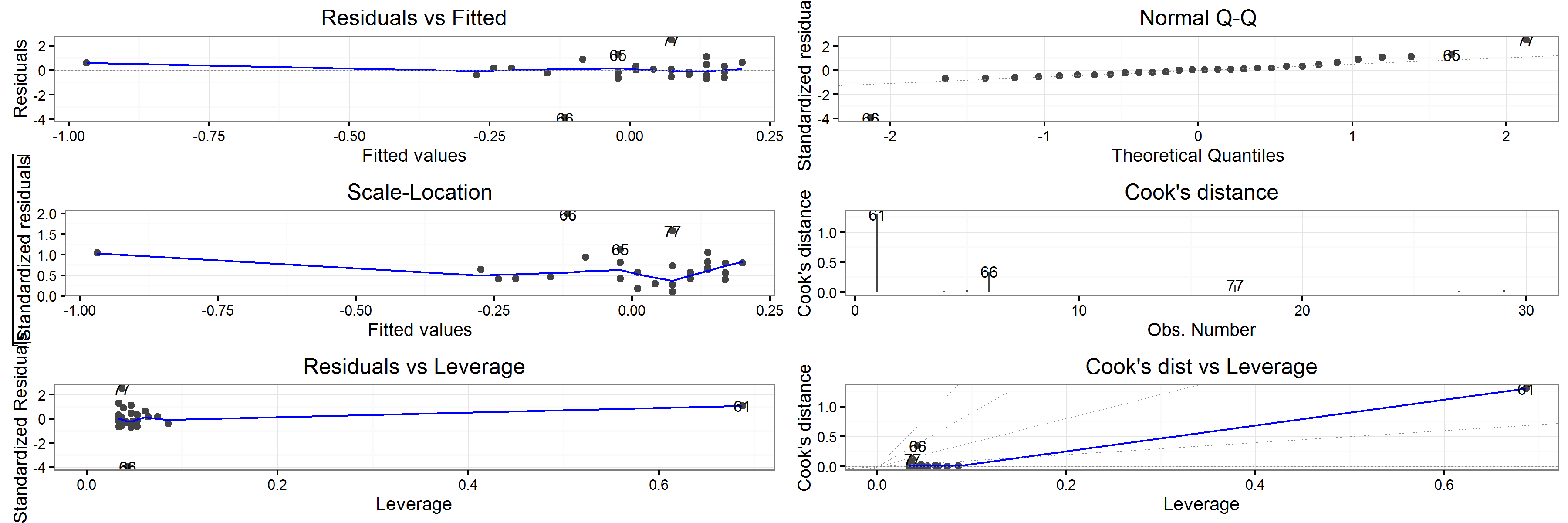
##   
## $shape



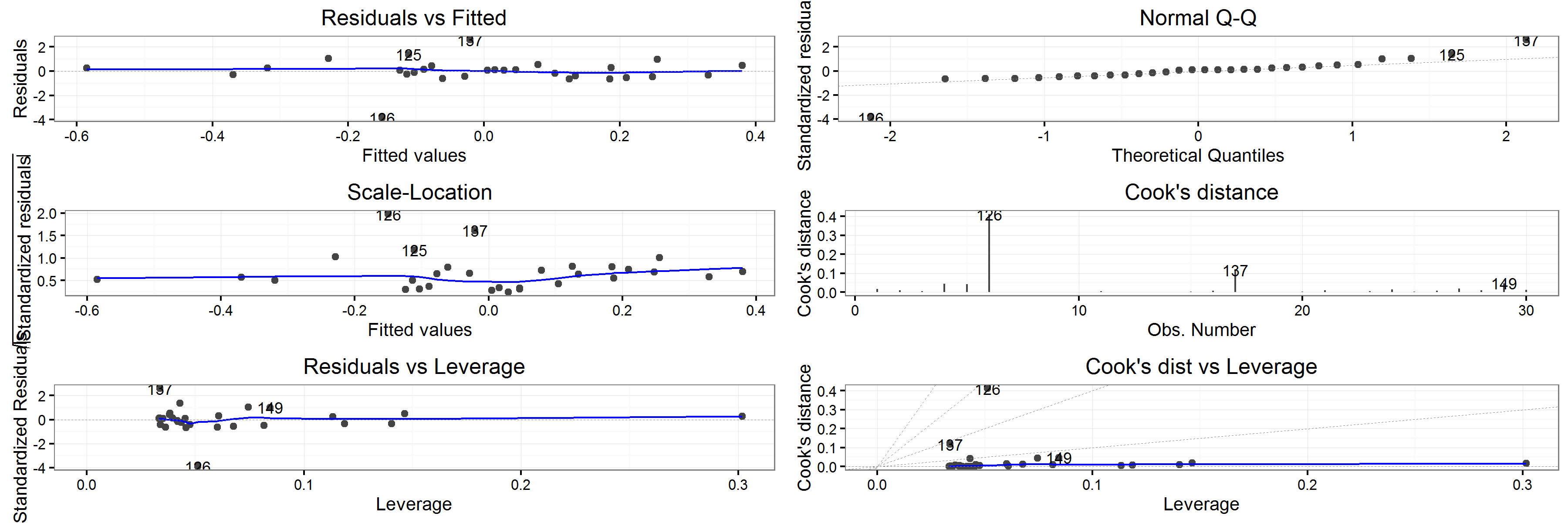
##   
## $sphericity



##   
## $`thickness(mm)`



##   
## $volume



##   
## $`width(mm)`

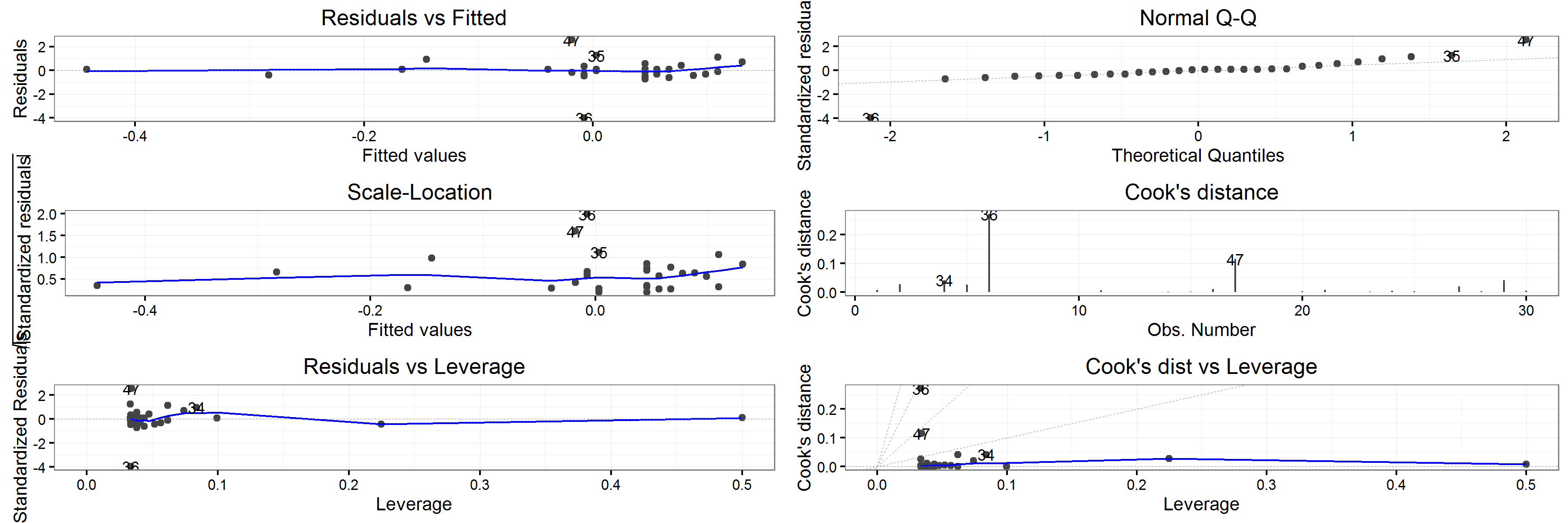
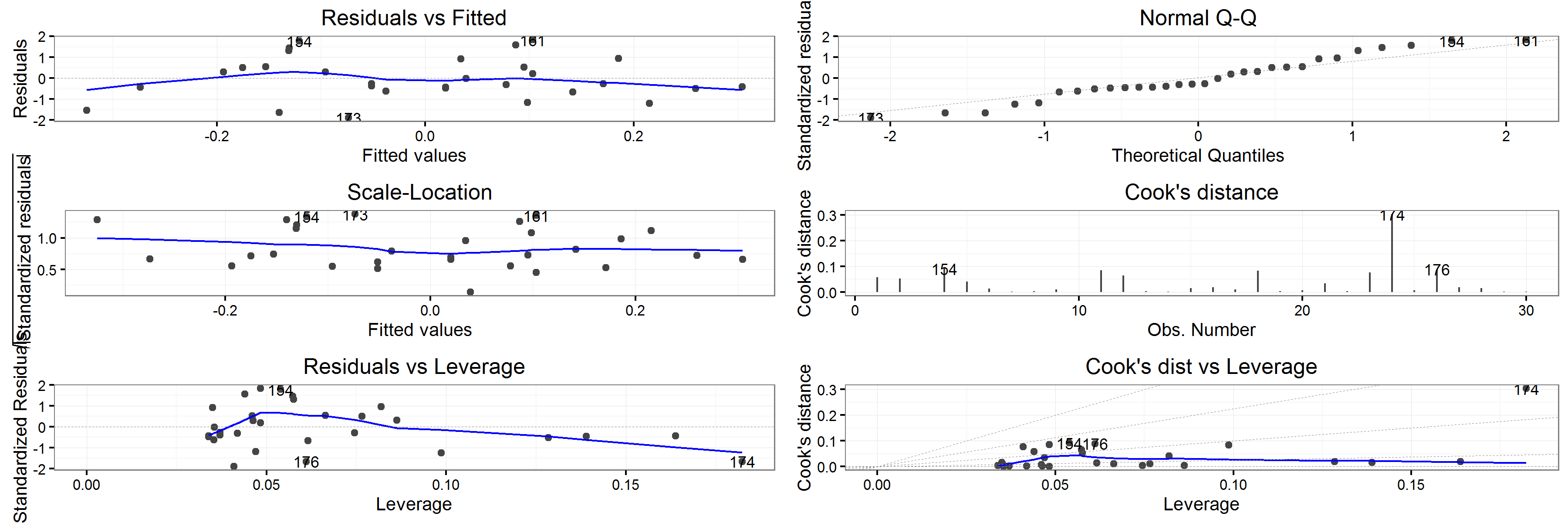


Table 2: Horizontal direction: Correlation of artefact volume with Cook's distance values for each model

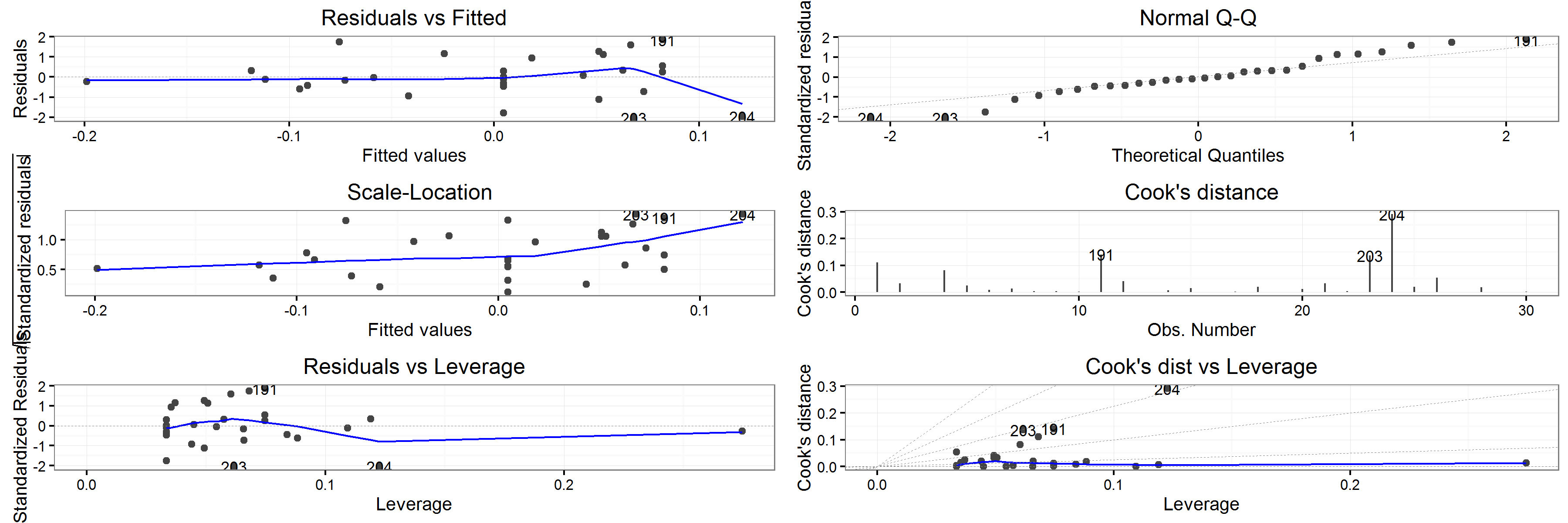
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| estimate | statistic | p.value | parameter | conf.low | conf.high | variable |
| 0.0753160 | 0.3996698 | 0.6924323 | 28 | -0.2929008 | 0.4240782 | elongation |
| 0.1217348 | 0.6489867 | 0.5216351 | 28 | -0.2494758 | 0.4617528 | flatness |
| 0.4254928 | 2.4879458 | 0.0190709 | 28 | 0.0770319 | 0.6813210 | length(mm) |
| 0.0949662 | 0.5047955 | 0.6176525 | 28 | -0.2747015 | 0.4401756 | shape |
| 0.1089994 | 0.5802276 | 0.5664017 | 28 | -0.2615403 | 0.4515371 | sphericity |
| 0.5473642 | 3.4608652 | 0.0017451 | 28 | 0.2330526 | 0.7581311 | thickness(mm) |
| 0.1339560 | 0.7152751 | 0.4803613 | 28 | -0.2377890 | 0.4714719 | volume |
| 0.1503118 | 0.8045156 | 0.4278778 | 28 | -0.2219782 | 0.4843520 | width(mm) |

## Diagnostic plots for linear models of vertical displacement of artefacts

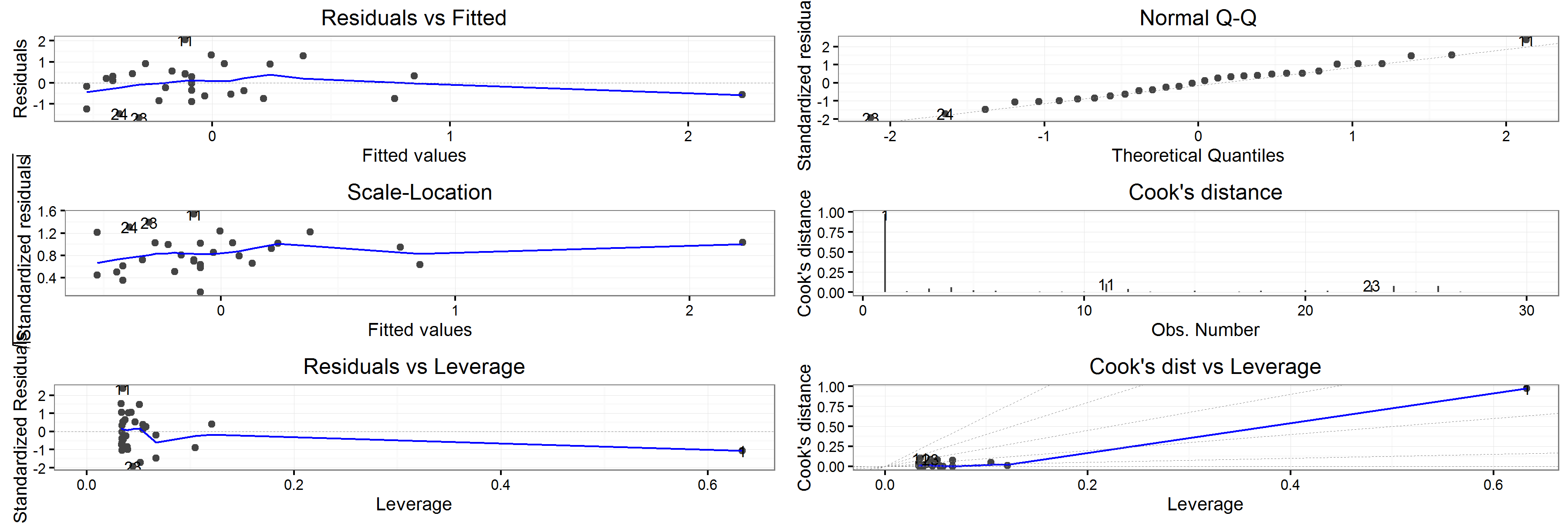
## $elongation



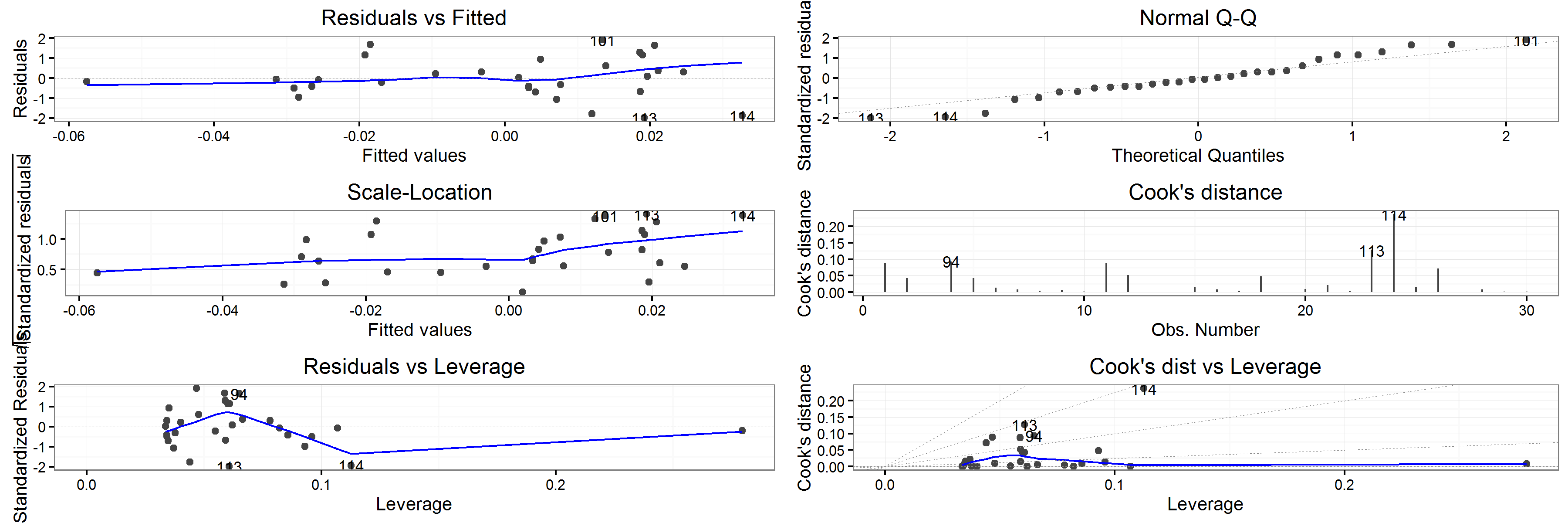
##   
## $flatness



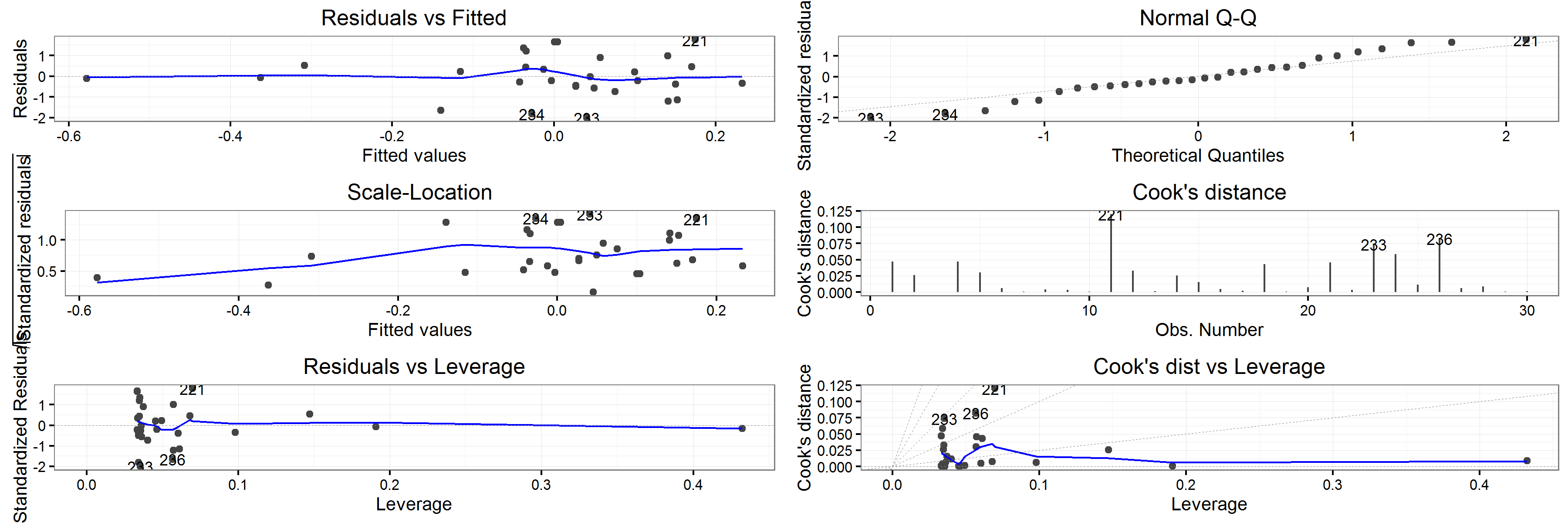
##   
## $`length(mm)`



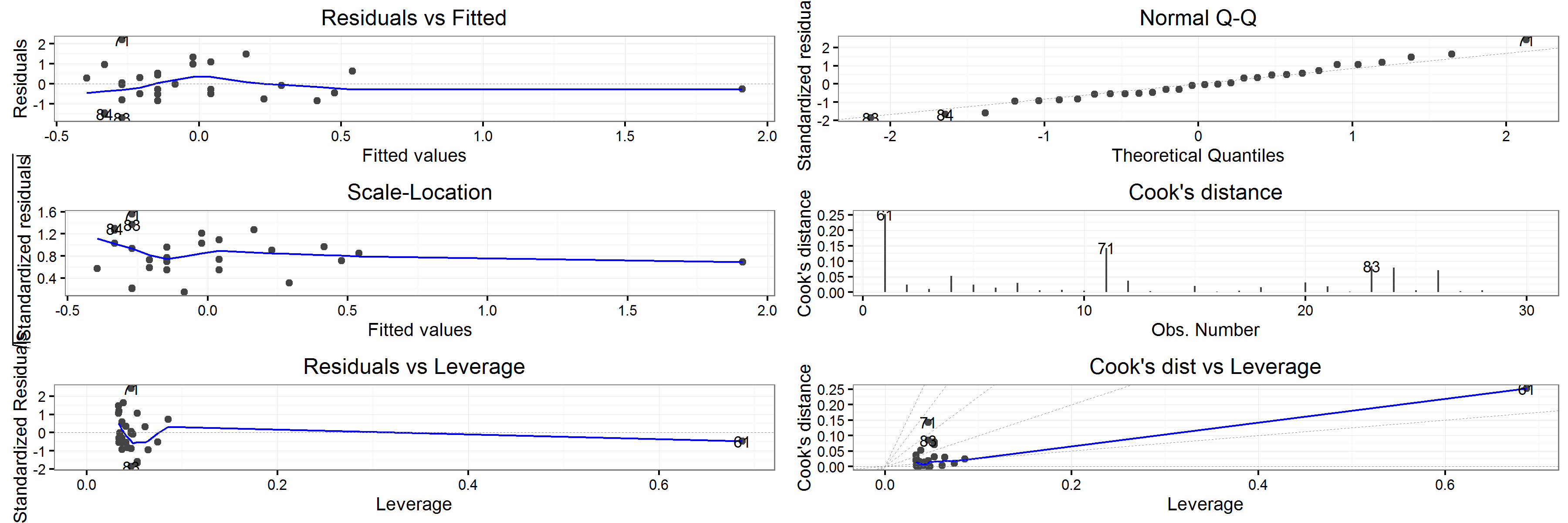
##   
## $shape



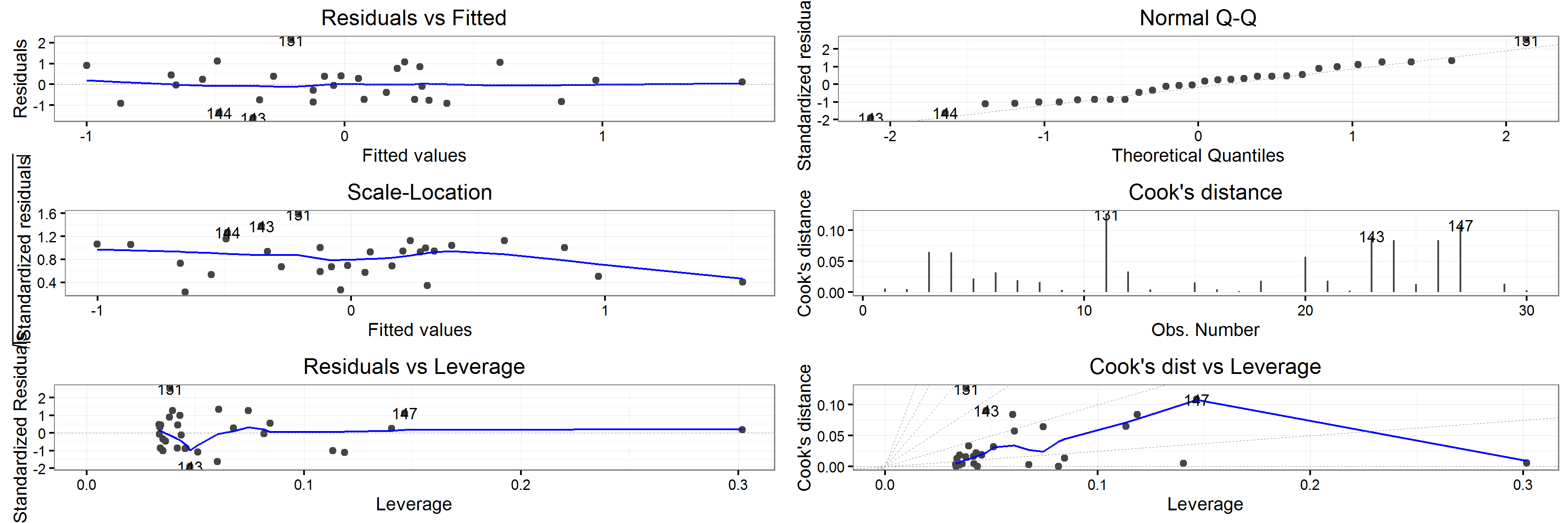
##   
## $sphericity



##   
## $`thickness(mm)`



##   
## $volume



##   
## $`width(mm)`

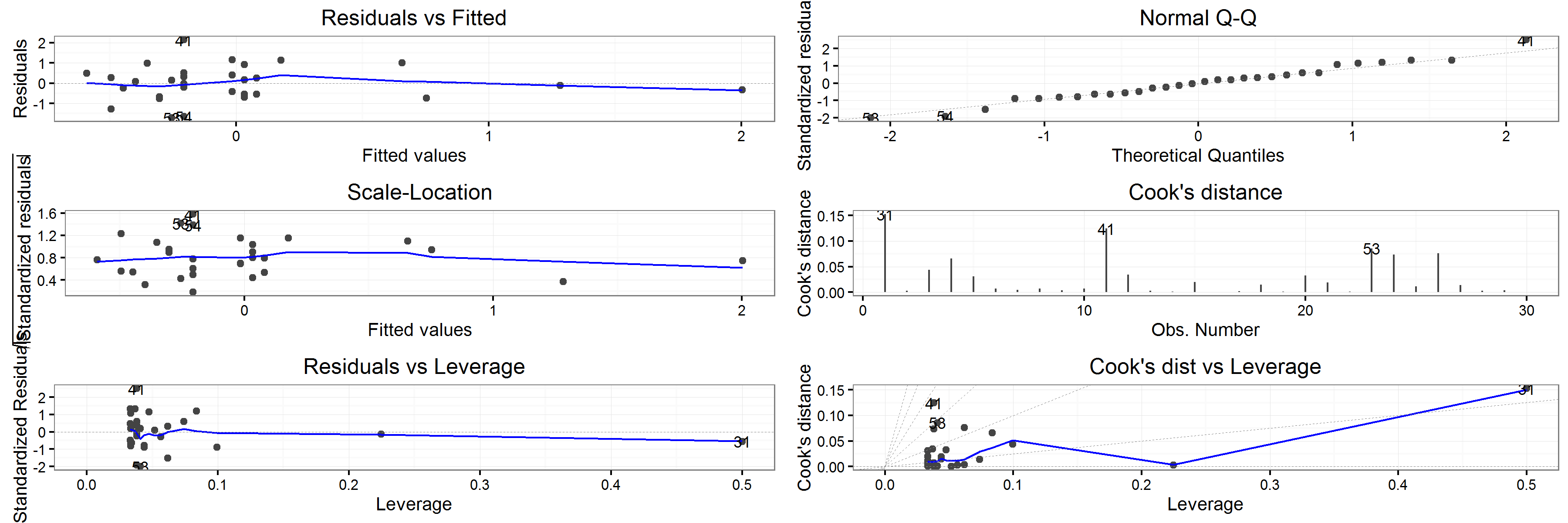
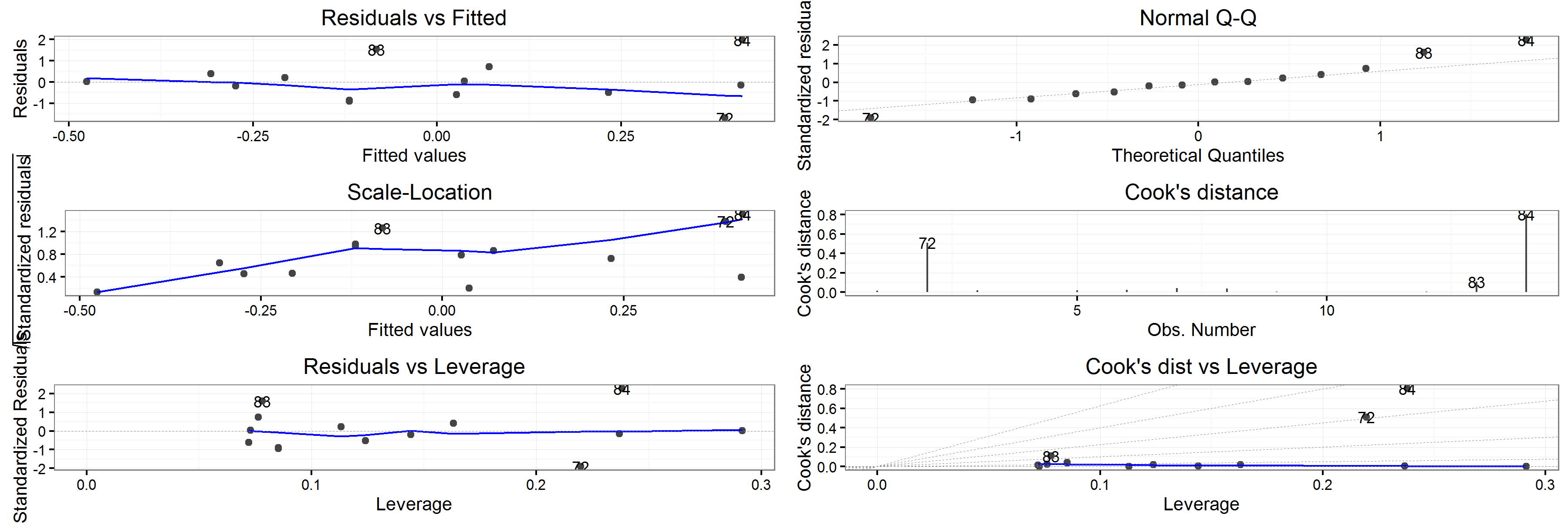


Table 3: Vertical displacement: Correlation of artefact volume with Cook's distance values for each model

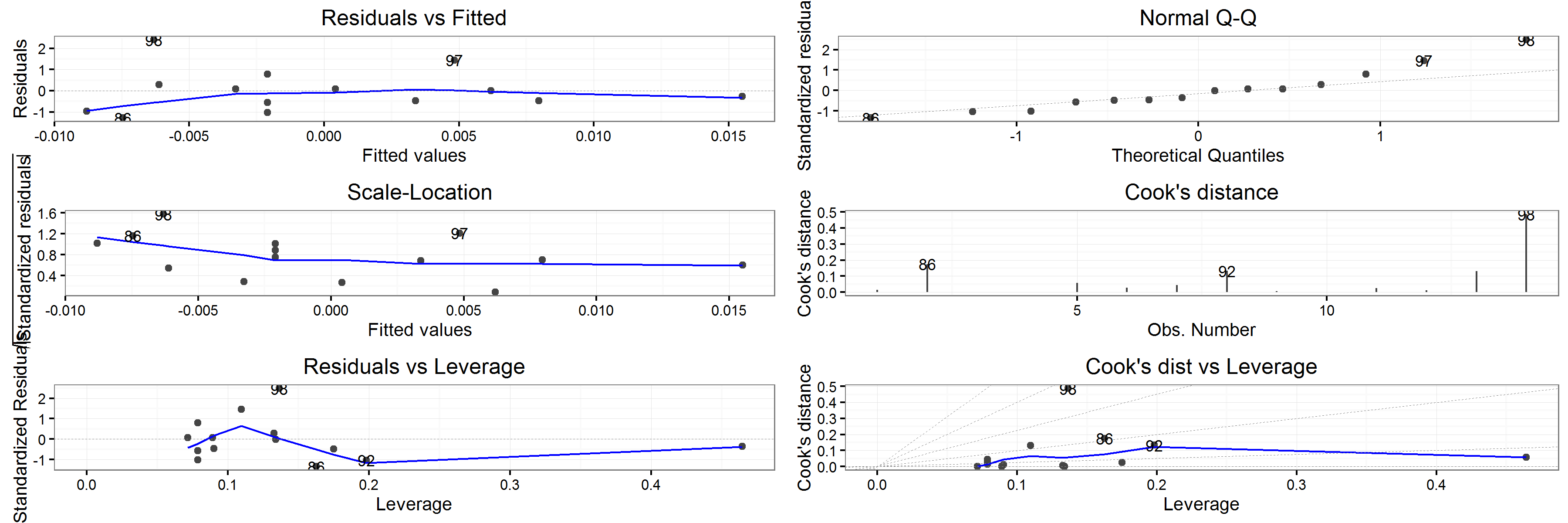
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| estimate | statistic | p.value | parameter | conf.low | conf.high | variable |
| -0.0949093 | -0.5044902 | 0.6178643 | 28 | -0.4401292 | 0.2747546 | elongation |
| -0.0128097 | -0.0677883 | 0.9464359 | 28 | -0.3713651 | 0.3490704 | flatness |
| 0.5070683 | 3.1130466 | 0.0042389 | 28 | 0.1796105 | 0.7333655 | length(mm) |
| 0.0056905 | 0.0301119 | 0.9761914 | 28 | -0.3553071 | 0.3652110 | shape |
| -0.0637791 | -0.3381759 | 0.7377533 | 28 | -0.4145236 | 0.3034630 | sphericity |
| 0.3525586 | 1.9935728 | 0.0560227 | 28 | -0.0088324 | 0.6324913 | thickness(mm) |
| -0.3094460 | -1.7219532 | 0.0961094 | 28 | -0.6025416 | 0.0572001 | volume |
| 0.2242869 | 1.2178418 | 0.2334477 | 28 | -0.1479361 | 0.5408532 | width(mm) |

## Diagnostic plots for linear models of artefact orientation

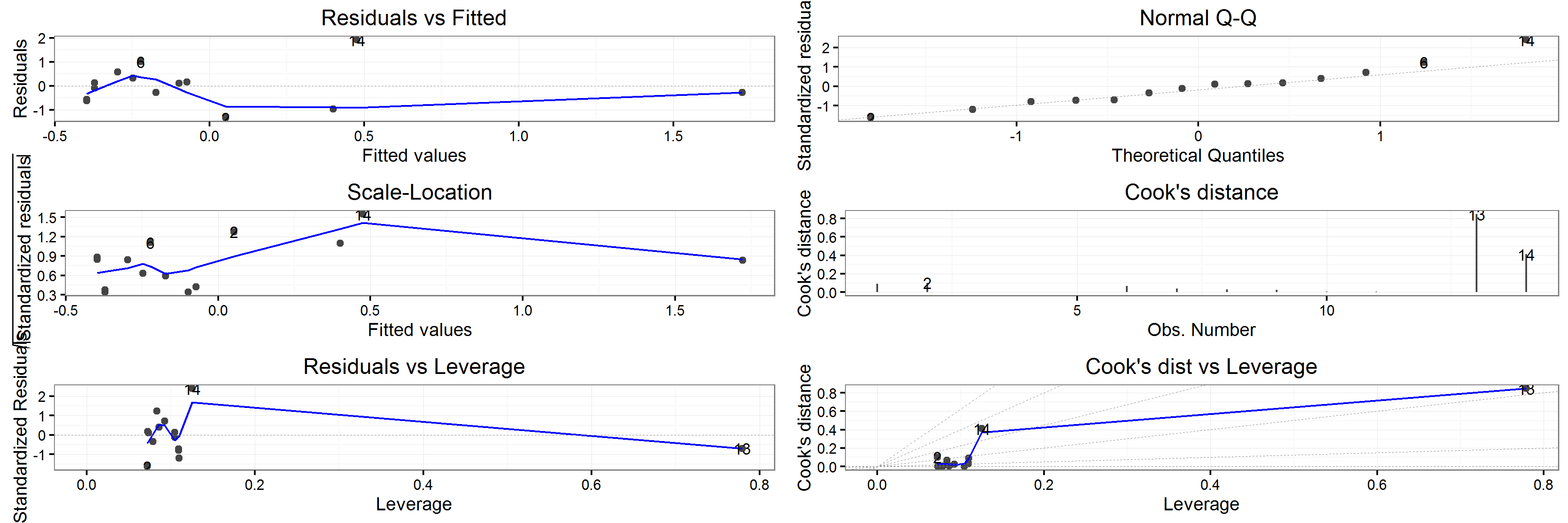
## $elongation



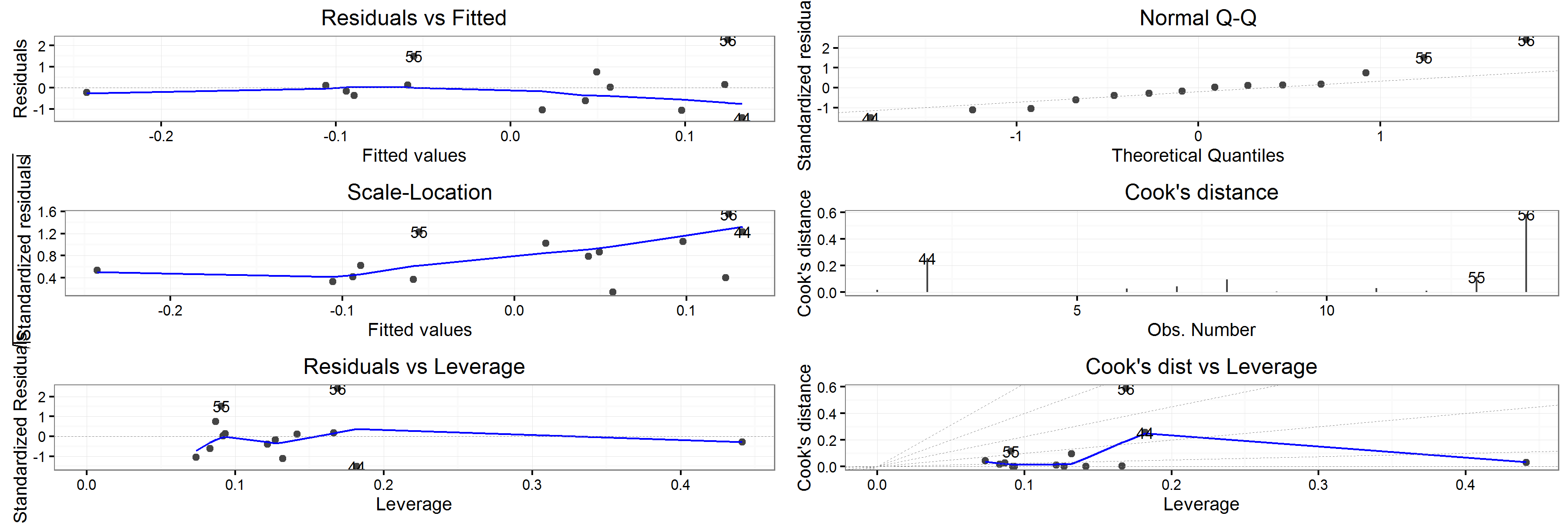
##   
## $flatness



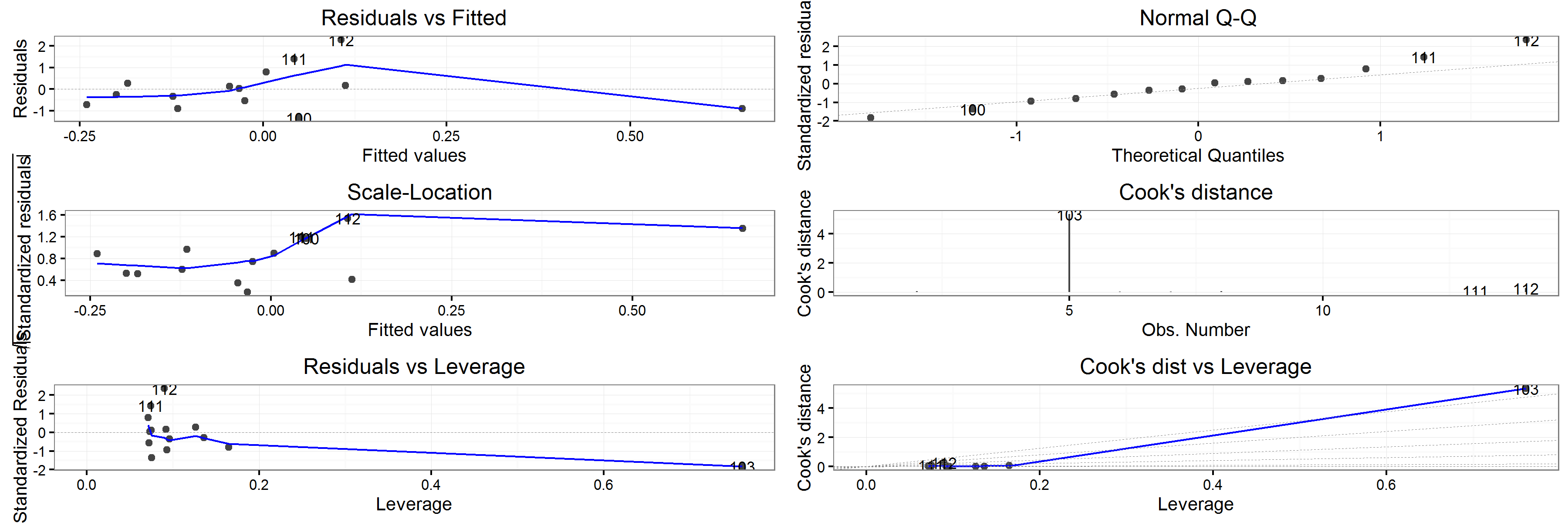
##   
## $`length(mm)`



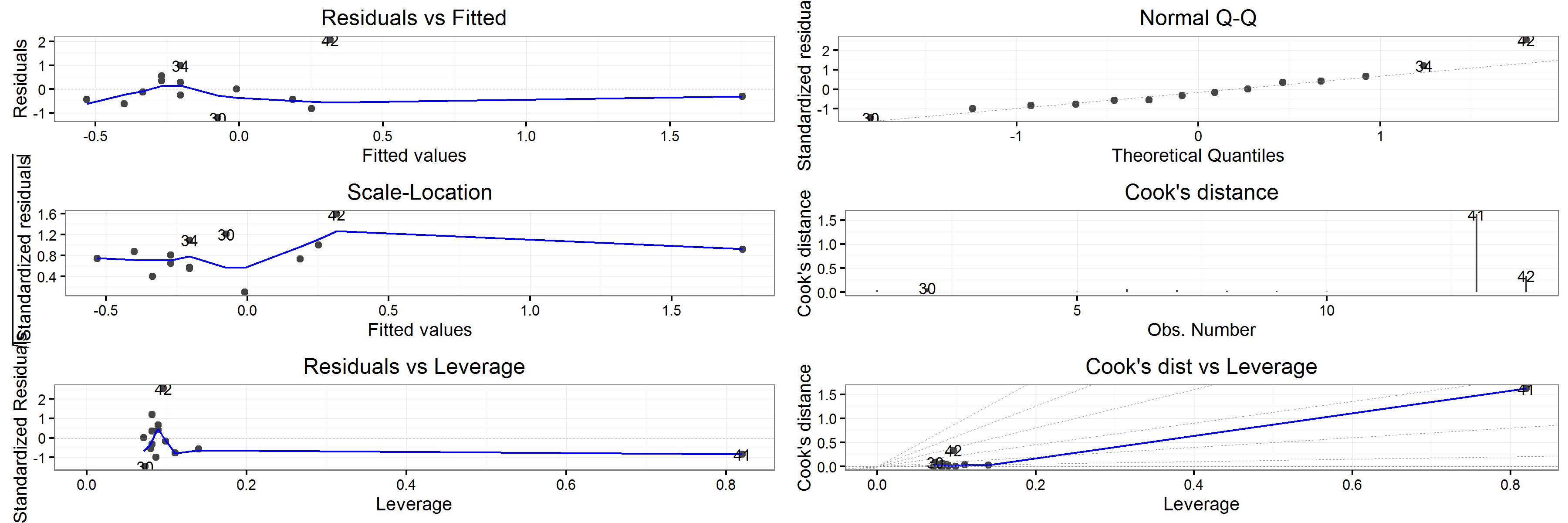
##   
## $shape



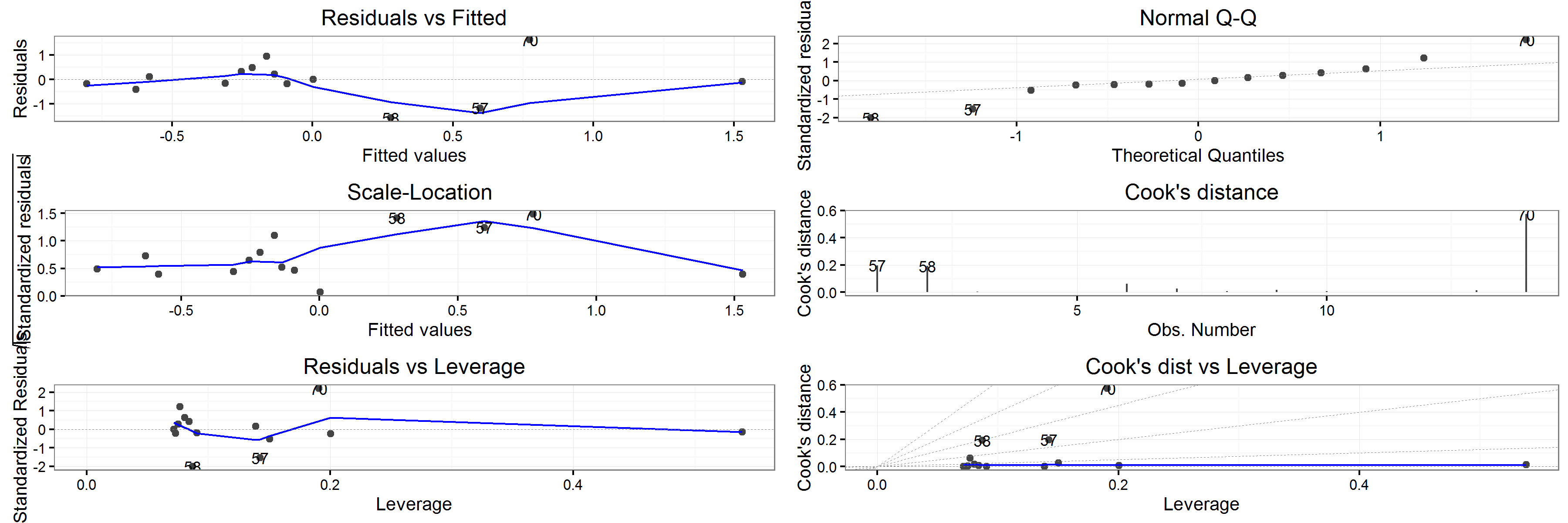
##   
## $sphericity



##   
## $`thickness(mm)`



##   
## $volume



##   
## $`width(mm)`

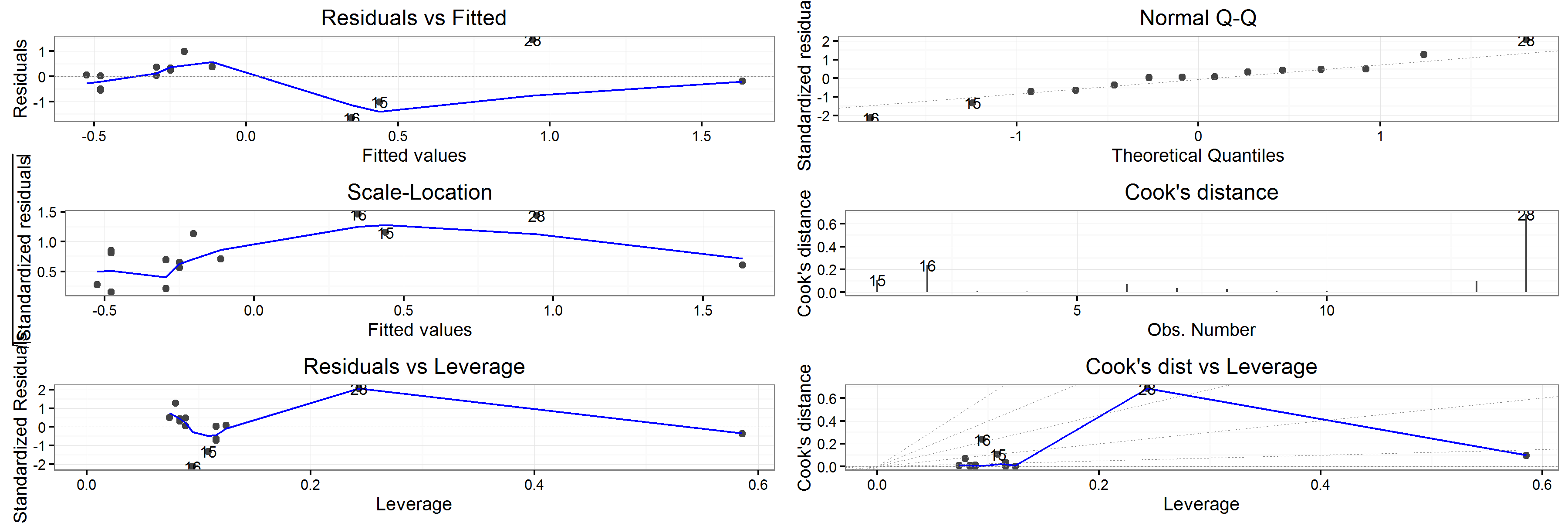
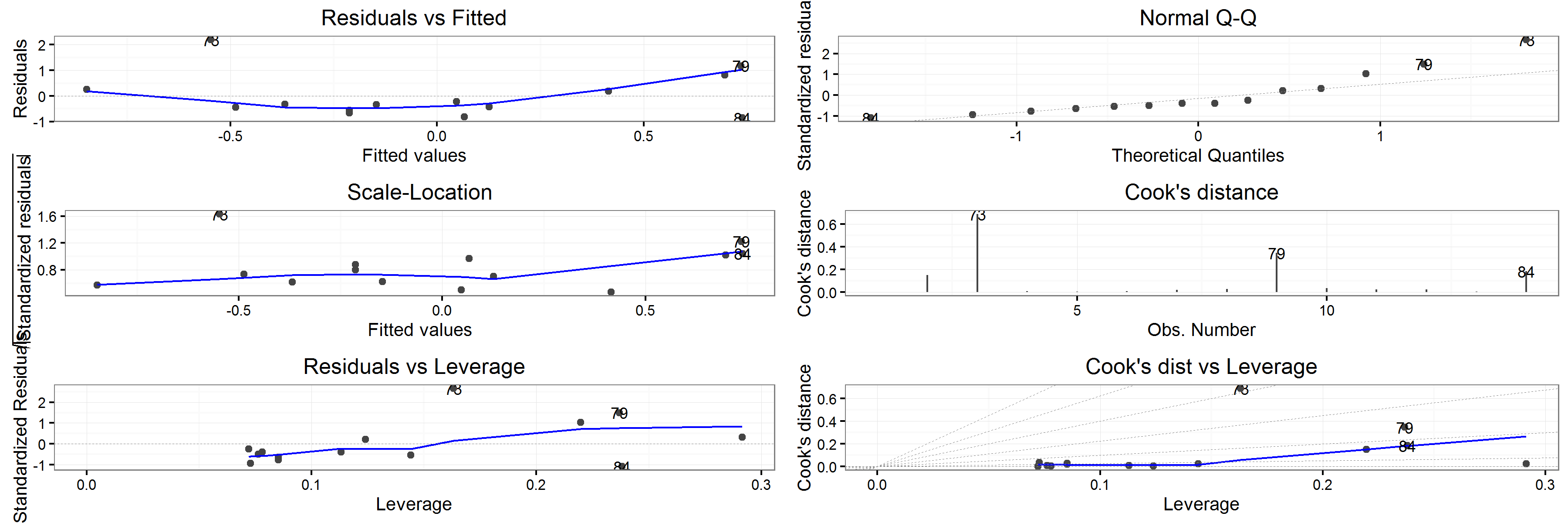


Table 4: Artefact orientation: Correlation of artefact volume with Cook's distance values for each model

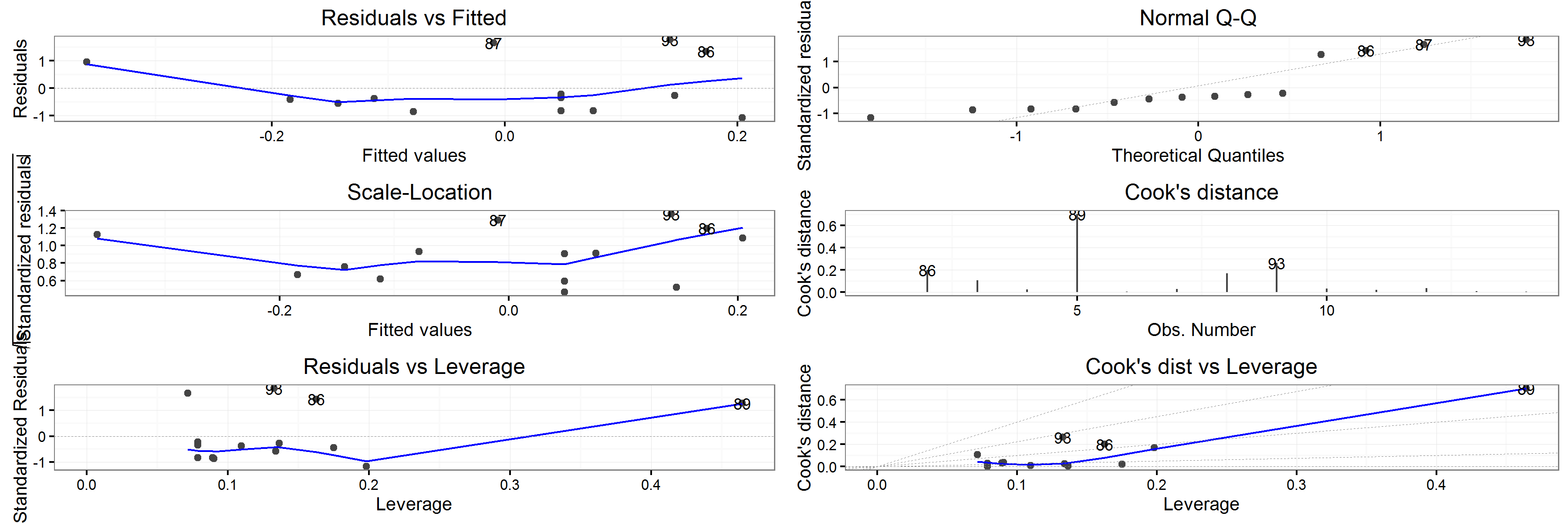
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| estimate | statistic | p.value | parameter | conf.low | conf.high | variable |
| 0.4503559 | 1.7473039 | 0.1061013 | 12 | -0.1054116 | 0.7917476 | elongation |
| 0.4507625 | 1.7492831 | 0.1057475 | 12 | -0.1049072 | 0.7919379 | flatness |
| 0.8538854 | 5.6833176 | 0.0001018 | 12 | 0.5911112 | 0.9527965 | length(mm) |
| 0.4527153 | 1.7588098 | 0.1040588 | 12 | -0.1024801 | 0.7928509 | shape |
| -0.0177394 | -0.0614606 | 0.9520042 | 12 | -0.5432061 | 0.5177129 | sphericity |
| 0.7887669 | 4.4450997 | 0.0007997 | 12 | 0.4440052 | 0.9300974 | thickness(mm) |
| 0.4695951 | 1.8425172 | 0.0902315 | 12 | -0.0812213 | 0.8006793 | volume |
| 0.5036017 | 2.0192788 | 0.0663690 | 12 | -0.0368147 | 0.8161147 | width(mm) |

## Diagnostic plots for linear models of artefact plunge

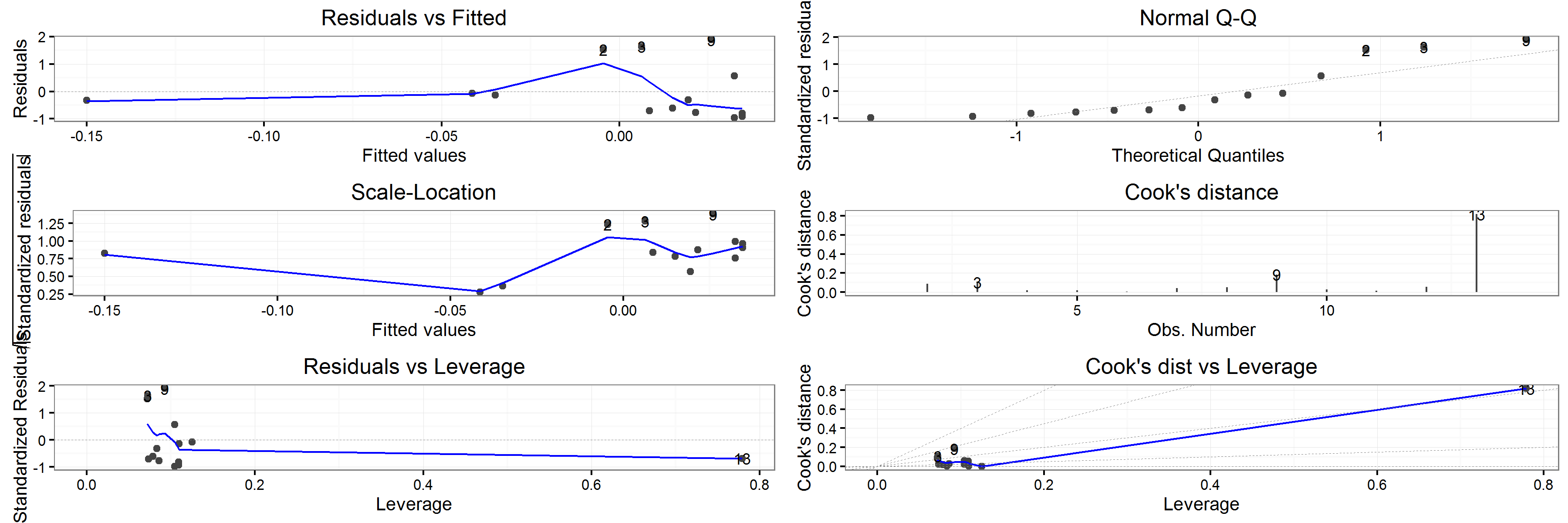
## $elongation



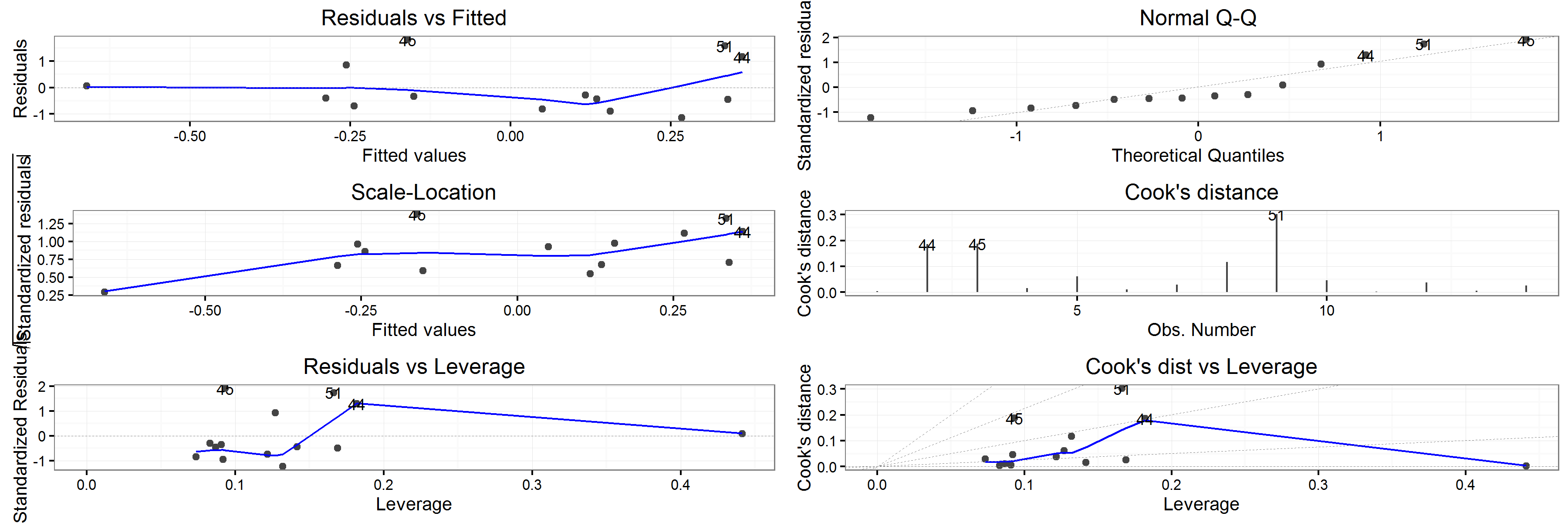
##   
## $flatness



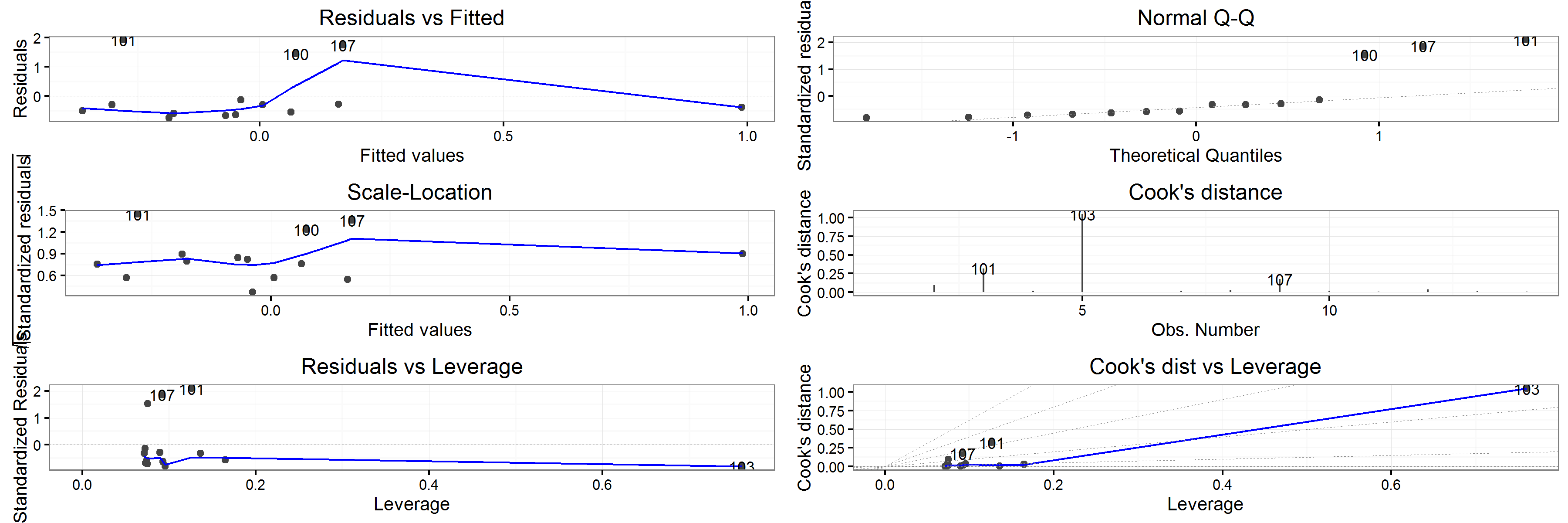
##   
## $`length(mm)`



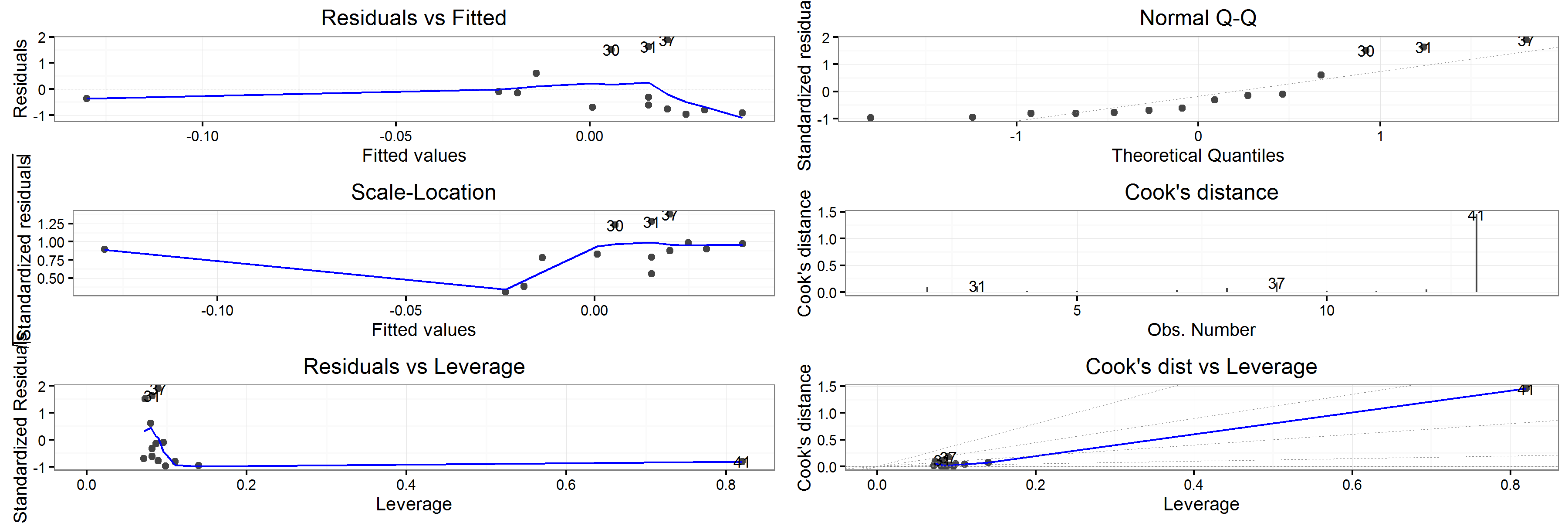
##   
## $shape



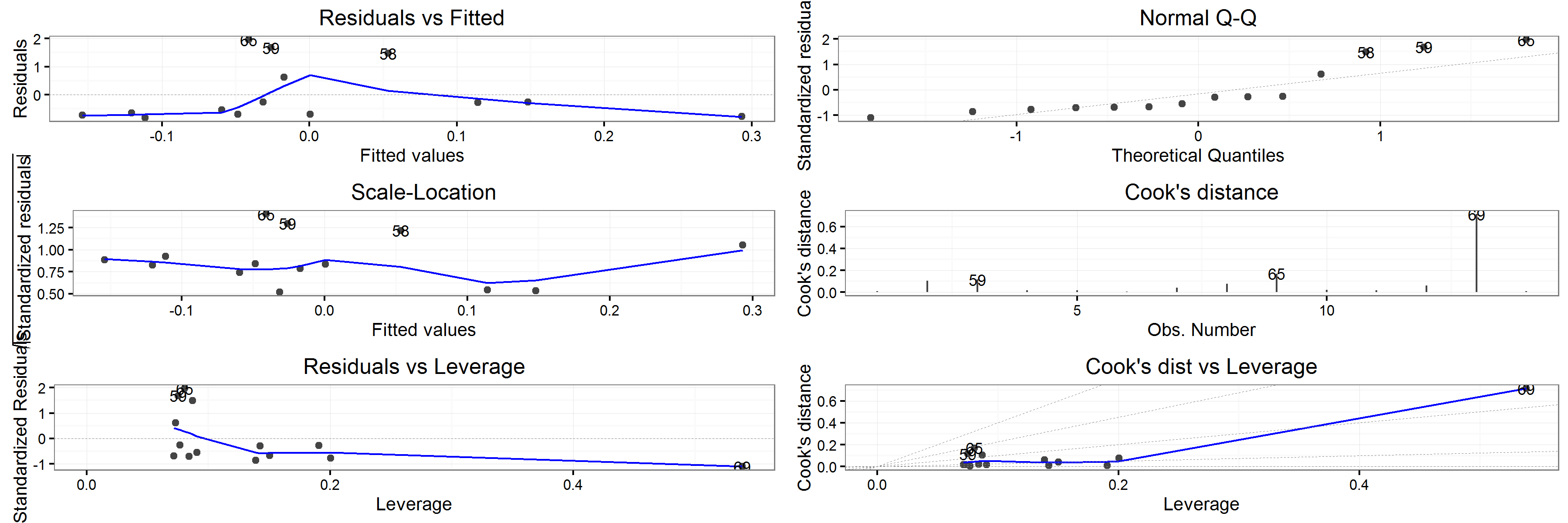
##   
## $sphericity



##   
## $`thickness(mm)`



##   
## $volume



##   
## $`width(mm)`

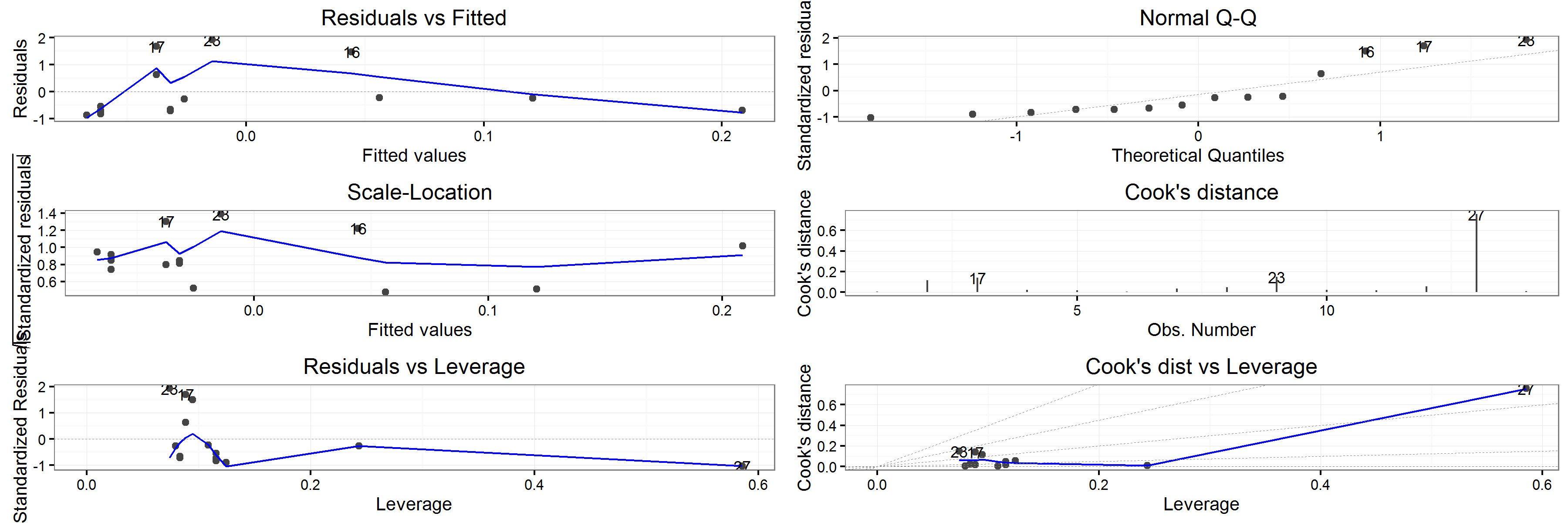


Table 5: Artefact plunge: Correlation of artefact volume with Cook's distance values for each model

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| estimate | statistic | p.value | parameter | conf.low | conf.high | variable |
| -0.0240870 | -0.0834639 | 0.9348588 | 12 | -0.5476673 | 0.5130493 | elongation |
| -0.1656422 | -0.5818390 | 0.5714441 | 12 | -0.6399765 | 0.4001007 | flatness |
| 0.6479169 | 2.9465902 | 0.0122214 | 12 | 0.1788058 | 0.8770066 | length(mm) |
| -0.2179372 | -0.7735506 | 0.4541661 | 12 | -0.6709345 | 0.3535210 | shape |
| -0.0845445 | -0.2939231 | 0.7738386 | 12 | -0.5887156 | 0.4669827 | sphericity |
| 0.6759241 | 3.1771524 | 0.0079632 | 12 | 0.2266164 | 0.8880286 | thickness(mm) |
| 0.6373726 | 2.8653648 | 0.0142116 | 12 | 0.1613619 | 0.8727935 | volume |
| 0.6625271 | 3.0640122 | 0.0098255 | 12 | 0.2034732 | 0.8827866 | width(mm) |

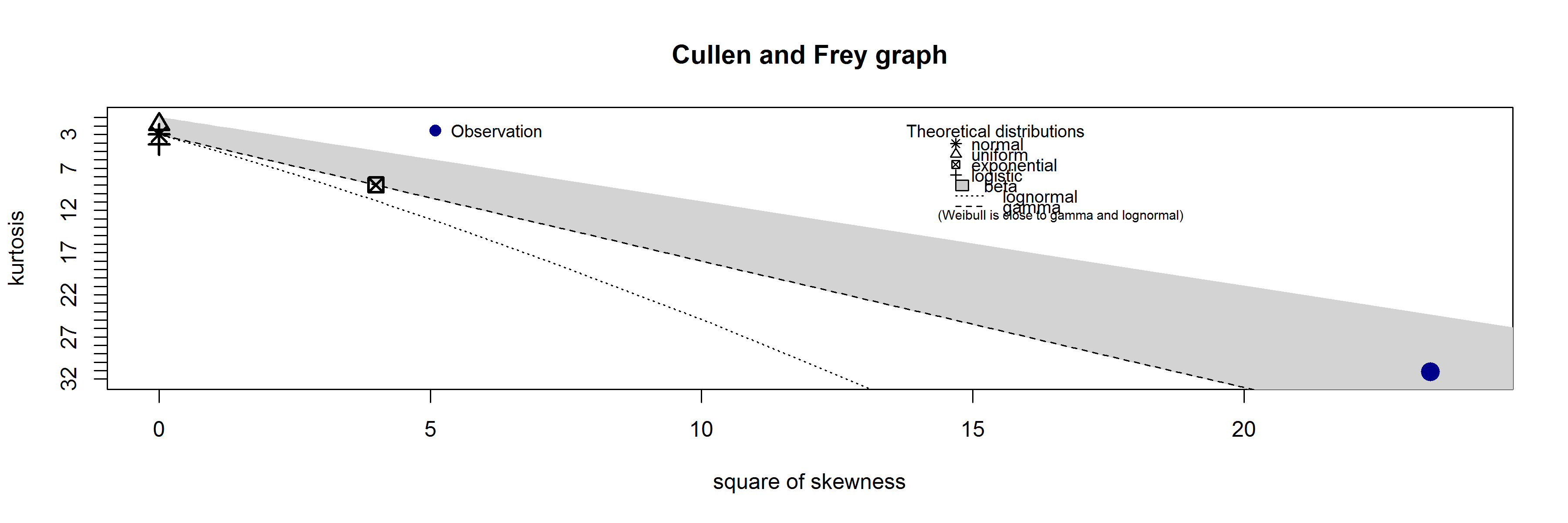
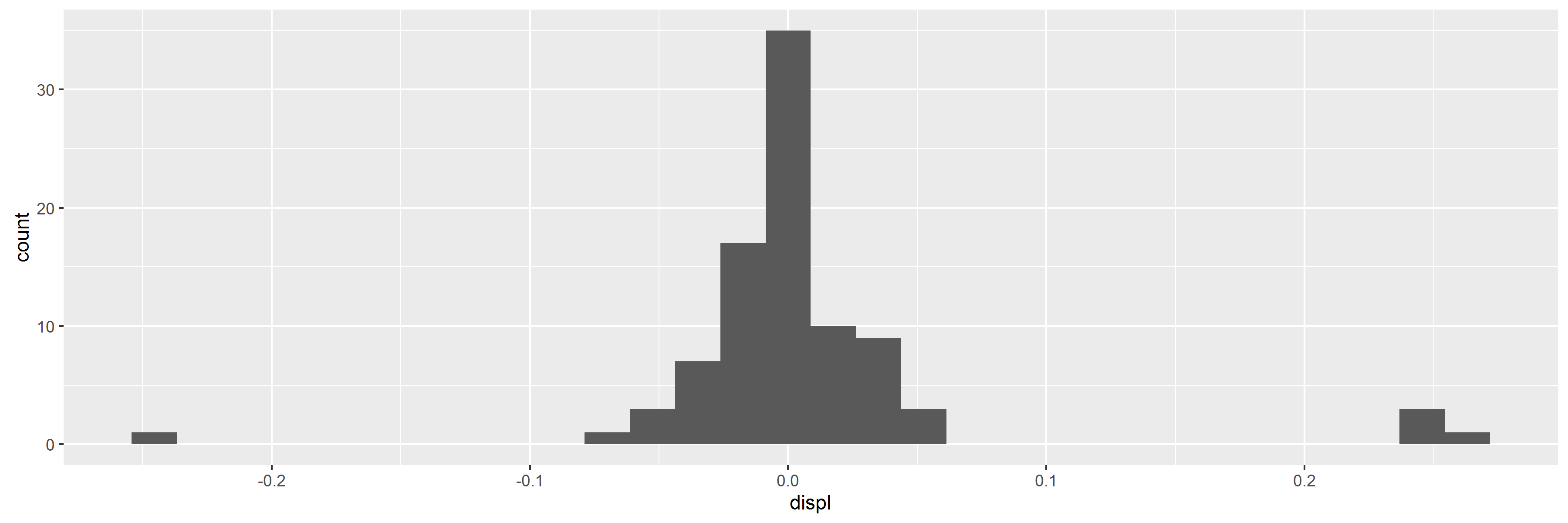
## Power analysis for the linear models

With thirty artefacts, the linear models in this paper have a power value of 0.889. This is a high power value, and indicates that our sample size is adequete for our regression models.

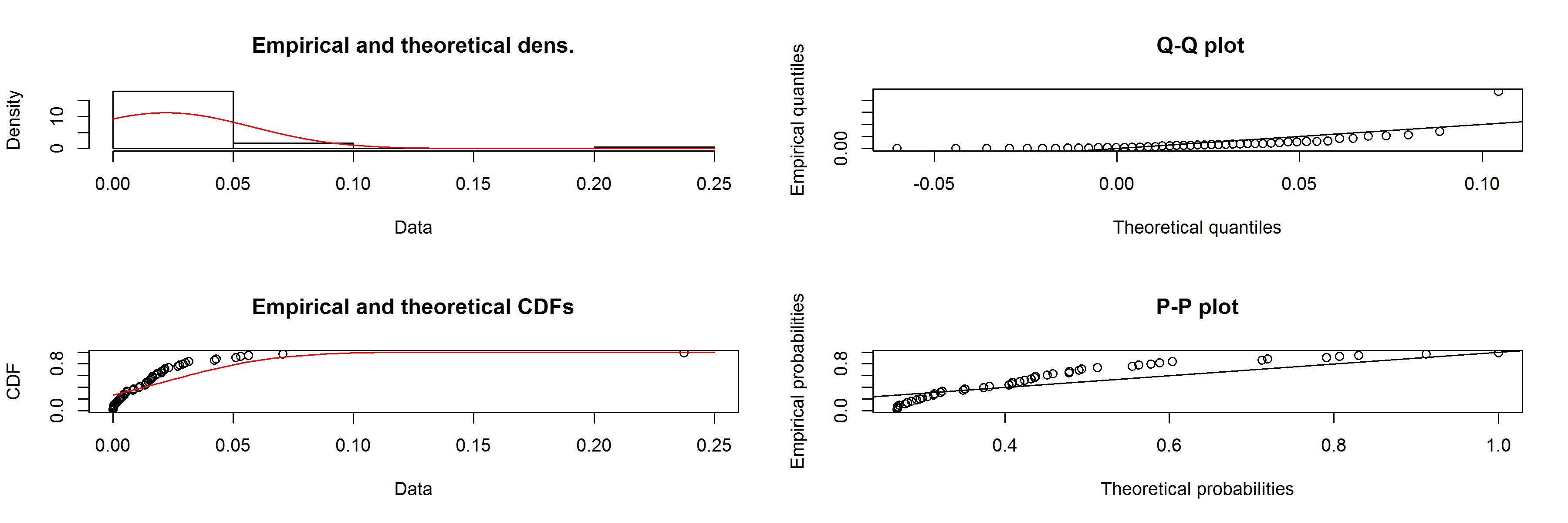
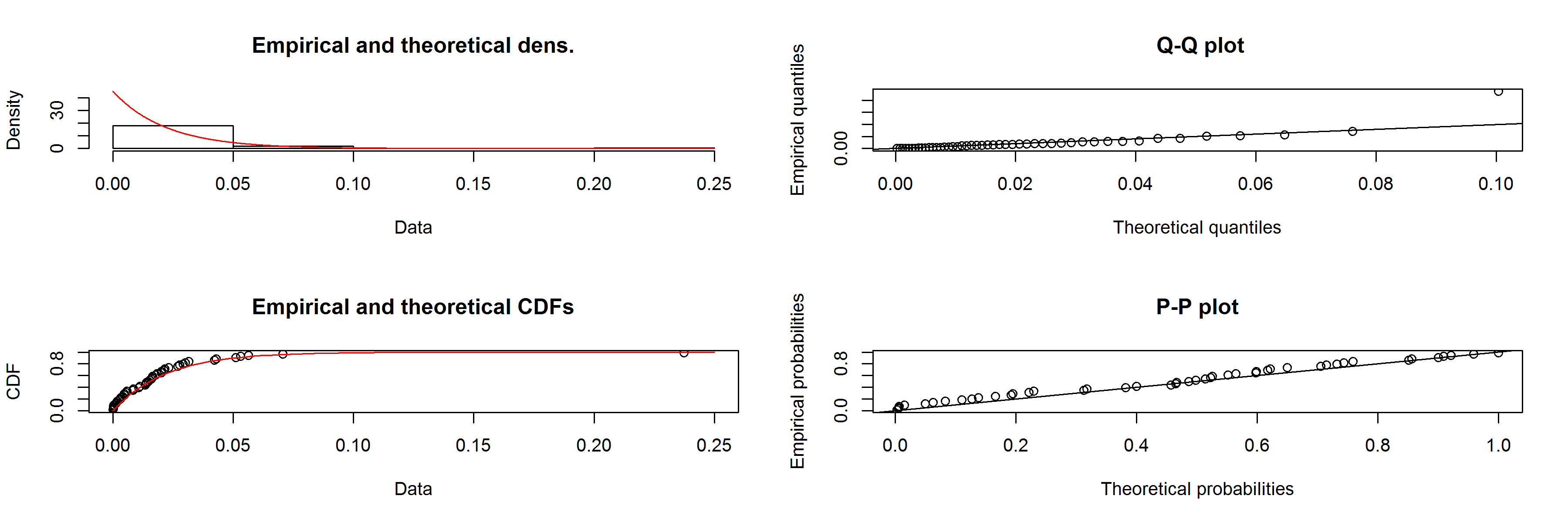
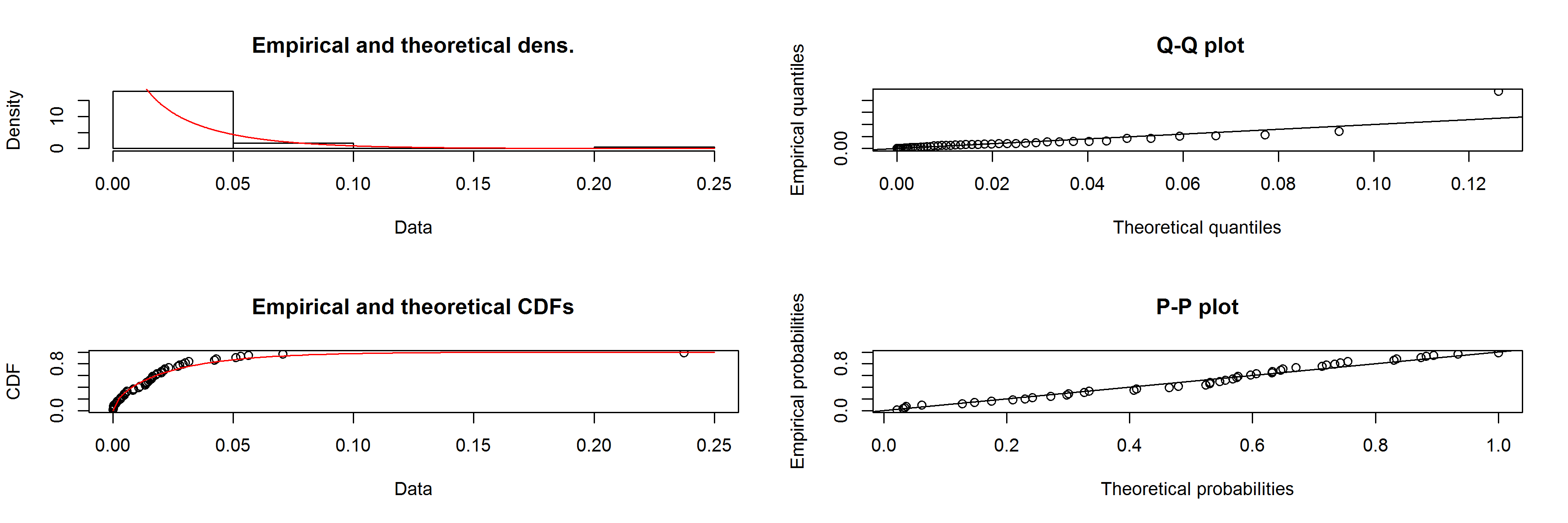
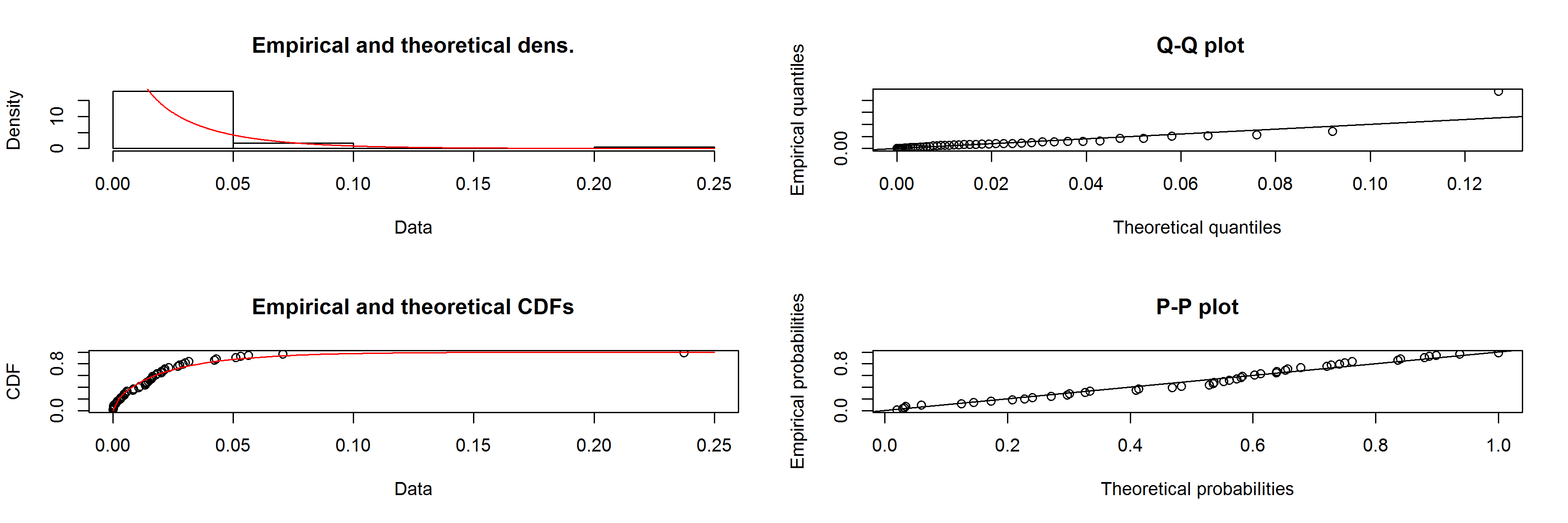
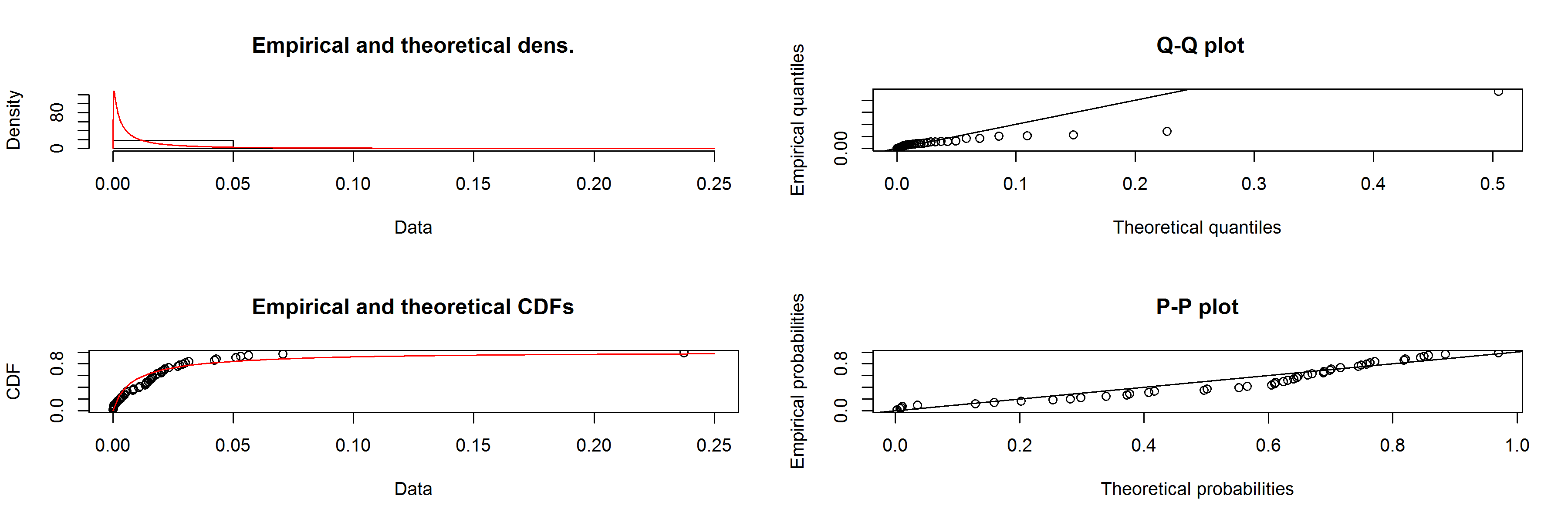
## Simulation of many trampling events

How to determine the how deep is too deep for trampling to have an effect on an artefact? Here we look at the distribution of distances below the starting point, and assess which theoretical distribution is the best fit.

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

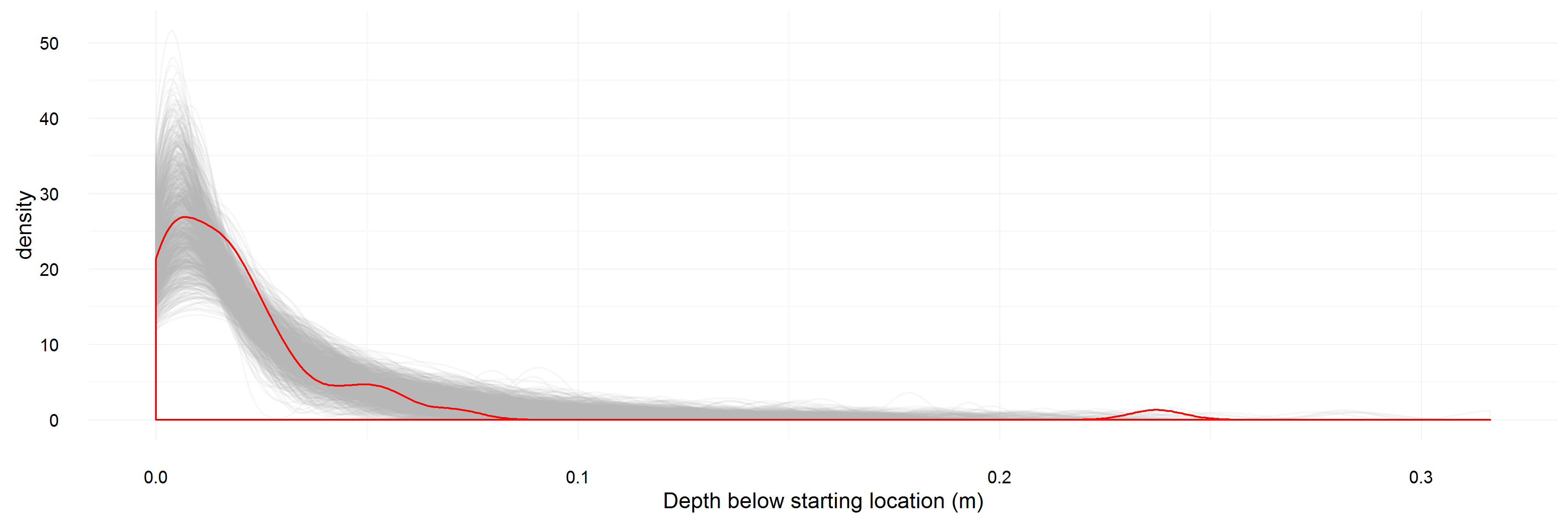


## summary statistics  
## ------  
## min: -0.237297 max: -6.05e-05   
## median: -0.014722   
## mean: -0.02209166   
## estimated sd: 0.03613033   
## estimated skewness: -4.841596   
## estimated kurtosis: 31.13909



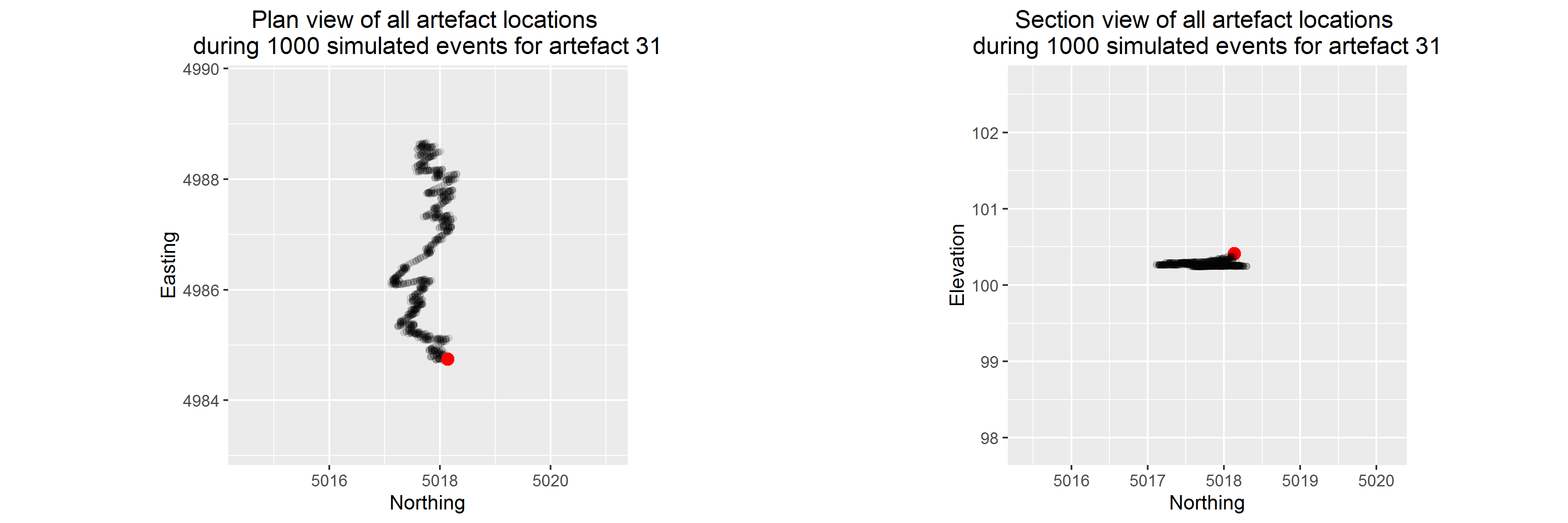
## [[1]]  
## NULL  
##   
## [[2]]  
## NULL  
##   
## [[3]]  
## NULL  
##   
## [[4]]  
## NULL  
##   
## [[5]]  
## NULL

We evaluated 5 possible distributions to find the best fit with our observed values of artefact displacement below the point of origin. Using visual inspection of diagnostic plots and the AIC valuess, we have determined that a gamma distribution is the best fit for the observed vertical displacenent values.



The plot above shows 1000 simulated gamma distributions (in grey) and our observed distrubtion in red.

The plot below shows the movement of one artefact during 1000 simulated trampling events.



# Colophon

This report was generated on 2016-10-20 15:09:05 using the following computational environment and dependencies:

## Session info --------------------------------------------------------------

## setting value   
## version R version 3.3.1 (2016-06-21)  
## system x86\_64, mingw32   
## ui RTerm   
## language (EN)   
## collate English\_Australia.1252   
## tz America/Los\_Angeles   
## date 2016-10-20

## Packages ------------------------------------------------------------------

## package \* version date source   
## assertthat 0.1 2013-12-06 CRAN (R 3.3.1)   
## bookdown 0.1.1 2016-08-03 Github (rstudio/bookdown@902a670)   
## colorspace 1.2-6 2015-03-11 CRAN (R 3.3.1)   
## DBI 0.5-1 2016-09-10 CRAN (R 3.3.1)   
## devtools 1.12.0 2016-06-24 CRAN (R 3.3.1)   
## digest 0.6.10 2016-08-02 CRAN (R 3.3.1)   
## dplyr \* 0.5.0.9000 2016-08-03 Github (hadley/dplyr@8b28b0b)   
## evaluate 0.9 2016-04-29 CRAN (R 3.3.1)   
## formatR 1.4 2016-05-09 CRAN (R 3.3.1)   
## ggfortify \* 0.2.0 2016-06-02 CRAN (R 3.3.1)   
## ggplot2 \* 2.1.0 2016-03-01 CRAN (R 3.3.1)   
## gridExtra \* 2.2.1 2016-08-03 Github (baptiste/gridextra@478a7d2)  
## gtable 0.2.0 2016-02-26 CRAN (R 3.3.1)   
## htmltools 0.3.5 2016-03-21 CRAN (R 3.3.1)   
## httpuv 1.3.3 2015-08-04 CRAN (R 3.3.1)   
## knitr \* 1.14 2016-08-13 CRAN (R 3.3.1)   
## magrittr 1.5 2014-11-22 CRAN (R 3.3.1)   
## memoise 1.0.0 2016-01-29 CRAN (R 3.3.1)   
## mime 0.5 2016-07-07 CRAN (R 3.3.1)   
## miniUI 0.1.1 2016-01-15 CRAN (R 3.3.1)   
## munsell 0.4.3 2016-02-13 CRAN (R 3.3.1)   
## plyr 1.8.4 2016-06-08 CRAN (R 3.3.1)   
## R6 2.1.3 2016-08-19 CRAN (R 3.3.1)   
## Rcpp 0.12.7 2016-09-05 CRAN (R 3.3.1)   
## readxl \* 0.1.1 2016-03-28 CRAN (R 3.3.1)   
## rmarkdown 1.0.9001 2016-08-03 Github (rstudio/rmarkdown@7769342)   
## scales 0.4.0 2016-02-26 CRAN (R 3.3.1)   
## shiny 0.14 2016-09-10 CRAN (R 3.3.1)   
## stringi 1.1.1 2016-05-27 CRAN (R 3.3.0)   
## stringr \* 1.1.0 2016-08-19 CRAN (R 3.3.1)   
## tibble 1.2 2016-08-26 CRAN (R 3.3.1)   
## tidyr 0.6.0.9000 2016-09-17 Github (hadley/tidyr@3c9335b)   
## withr 1.0.2 2016-06-20 CRAN (R 3.3.1)   
## xtable 1.8-2 2016-02-05 CRAN (R 3.3.1)   
## yaml 2.1.13 2014-06-12 CRAN (R 3.3.1)

The current git commit of this file is 459bd2eca6392d0ba0e5e20fea3a0f375eb6125a, which is on the master branch and was made by Ben Marwick on 2016-10-20 13:34:25. The current commit message is "add repro para".