

# Jimmy\_Carrillo\_Trabajo\_Final

Jimmy Carrillo Novoa

14/7/2021

## TITULO Y AUTOR.

Titulo del Proyecto: Evaluación del efecto de la densidad de cultivo y el tamaño del estanque sobre la intensidad de luz ( $\text{watt}/\text{sq2}$ ) mediante el análisis de modelos lineales de efectos fijos-aleatorios y modelos lineales generalizados.

## PROBLEMATICA A RESOLVER

El cultivo del Salmon en tierra en sistema RAS (Recirculation Aquaculture Systems) a aumentado notablemente en los últimos años como una solución a los problemas ambientales que el sistema tradicional de cultivo en jaulas de mar ha generado desde el inicio de la industria y como una alternativa de bioseguridad para la especie cultivada.

Debido a los altos niveles de peces maduros en los sistemas RAS que se han documentado en varios sistemas de RAS existe la necesidad de conocer la relación y el comportamiento de la intensidad de luz en el fondo del estanque de un sistema RAS de cultivo con respecto a la densidad de cultivo y el tamaño del estanque

## DESCRIPCION DETALLADA DE LOS DATOS ORIGINALES

El set de datos corresponde a una granja peces cultivados en un sistema de cultivo RAS con dos Módulos de estanques Mod A-Mod B. Cada modulo de cultivo tiene una configuración de estanques de 50m<sup>3</sup> y 200m<sup>3</sup>. El set de datos tiene un total de 272 mediciones de intensidad de luz asociado a diferentes densidades de cultivo y relacionado a los dos módulos de cultivo. La toma de muestras de intensidad de luz se realizó durante un año en donde cada mes se realizaron varias mediciones a diferentes estanques con distinto volumen y densidad.

Las muestras fueron tomadas mediante LI-1500 light sensor logger. El detalle de los datos recolectados se presenta en `light_intensity.xlsx`.

La intensidad de luz fue medida en  $\text{watt}/\text{m}^2$ , el volumen de los estanques esta relacionado a los metros cúbicos de agua y la densidad de cultivo fue medida por la biomasa total en el estanque al momento de la medición sobre el volumen de agua en el estanque en la unidad de medición correspondiente a  $\text{kg}/\text{m}^3$ .

## ANALISIS EXPLORATORIO DE LOS DATOS

### ANALISIS PARTE 1

ANALISIS DE MODELOS LINEALES MIXTOS DE EFECTOS FIJOS Y EFECTOS ALEATORIOS PARA EL ANALISIS DEL EFECTO DE LA DENSIDAD DE CULTIVO, EL VOLUMEN DEL ESTANQUE SOBRE LA INTENSIDAD DE LUZ

Exploratorio del Set de Datos “`Light_intensity`”. Medida de la intensidad de luz en  $\text{watt}/\text{m}^2$  de diferentes estanques a diferentes densidades de cultivo en dos Modulos de cultivo diferentes en un sistema RAS.

## Exploratorio de las avriables a analizar

```
## tibble [272 x 5] (S3: tbl_df/tbl/data.frame)
## $ Density : num [1:272] 35 42 55 55 55 55 55 68 79 70 ...
## $ Volume_m3: chr [1:272] "50m3" "50m3" "50m3" "50m3" ...
## $ Tank : chr [1:272] "1A1" "1A2" "1A4" "2A1" ...
## $ Module : chr [1:272] "A" "A" "A" "A" ...
## $ watt_sq2 : num [1:272] 0.155 0.087 0.0975 0.1595 0.105 ...
```

El set de datos presenta 5 variables, de las cuales 2 de ellas son numéricas: Density y watt\_sq2, dos variables como factor: Volume\_m3 y Module y finalmente una variable caracter: Tank.

Mas adelante la variable Tank sera tratada como factor.

#DEFINICION DE VARIABLES.

Density: Corresponde a la densidad de cultivo en cada estanque medida en Biomasa KG/Volumen del estanque m3.

Watt\_sq2: Medición de la intensidad de luz en el fondo del estanque, el valor mostrado en el set de datos corresponde al promedio de 5 mediciones realzadas en cada estanque.

Volume\_m3: Volumen de agua en cada estanque de cultivo.

Module: Dos diferentes módulos de cultivo utilizando con diferentes sistemas RAS de tratamiento de agua.

Tank: corresponde al los estanques de cultivo, para este análisis se midió la intensidad de luz en dos tipos de estanques 50m3 y 200 m3.

## VARIABLE RESPUESTA:

La variable watt\_sq2 corresponde a la variable regresora de tipo cuantitativa continua. Se clasifica como regresora debido a que el valor de medicion depende de otras variables.

## VARIABLES REGRESORAS:

Las variables regresoras para el analisis son: Density, Volume\_m3, Tank y Module.

## AJUSTE DE EVARIABLES

```
## tibble [272 x 5] (S3: tbl_df/tbl/data.frame)
## $ Density : num [1:272] 35 42 55 55 55 55 55 68 79 70 ...
## $ Volume_m3: Factor w/ 2 levels "200m3","50m3": 2 2 2 2 2 2 2 1 1 1 ...
## $ Tank : Factor w/ 18 levels "1A1","1A2","1A3",...: 1 2 4 6 7 8 9 11 13 15 ...
## $ Module : Factor w/ 2 levels "A","B": 1 1 1 1 1 1 1 1 1 1 ...
## $ watt_sq2 : num [1:272] 0.155 0.087 0.0975 0.1595 0.105 ...
```

Tabla 1. Ajuste de la variable Tank como factor.

## RESUMEN ESTADISTICO DE LAS VARIABLES

##	Density	Volume_m3	Tank	Module	watt_sq2
##	Min. : 17.00	200m3:154	1A4 : 16	A:181	Min. :0.00100
##	1st Qu.: 55.00	50m3 :118	2A2 : 16	B: 91	1st Qu.:0.05175
##	Median : 71.00		2A3 : 16		Median :0.11000
##	Mean : 70.47		2A4 : 16		Mean :0.13058
##	3rd Qu.: 84.12		2B : 16		3rd Qu.:0.17525

```
## Max.      :116.00          3A      : 16          Max.      :0.80000
##                               (Other):176
```

Tabla 2: Resumen estadístico de las variables. Se observa que existen 181 mediciones de intensidad de luz en el Módulo 1 y 91 mediciones de intensidad de luz en Módulo 2. Con un total de 272 mediciones.

De la tabla 2 se observa que las 272 mediciones se realizaron sobre estanques de 50m<sup>3</sup> y 200 m<sup>3</sup>, 154 observaciones en estanques de 200m<sup>3</sup> y 118 mediciones en estanques de 50m<sup>3</sup>. Lo anterior indica que hubieron mediciones en el mismo estanque de 200m<sup>3</sup> y 50m<sup>3</sup> a diferentes densidades.

La densidad de cultivo mínima observada fue de 17kg/3 y la densidad máxima de 116 kg/m<sup>3</sup>

La menor intensidad de luz observada fue de 0.00100 watt/sq2 y la máxima fue de 0,80 watt/sq2

El valor de threshold considerado para el bienestar animal es de 0.016 watt/sq2.

## ANÁLISIS ESTADÍSTICO Y EXPLORATORIO DE LAS VARIABLES

### BOX-PLOT INTENSIDAD DE LUZ watt/sq2 MODULO A y MODULO B.

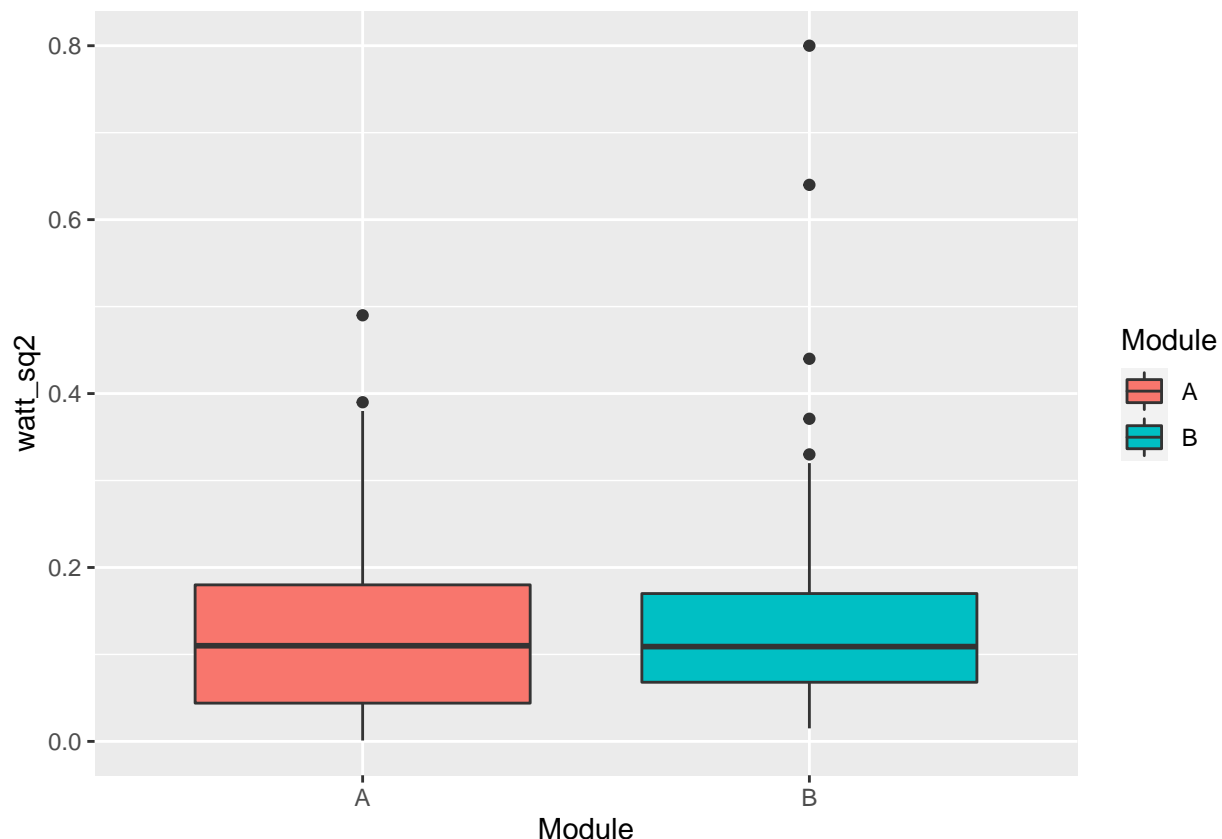


Figura 1. Box-Plot de la intensidad de luz en ambos módulos de cultivo. La medida de intensidad de luz es similar para ambos módulos sin diferenciar el volumen del estanque. Se observan algunos valores outlier los cuales no serán omitidos en el análisis.

### Media para la intensidad de luz de Module A y Module B

```
## # A tibble: 1 x 1
##   Mean_MA
##   <dbl>
## 1    0.127
```

Tabla 3. La media de intensidad de luz para el Module A es de 0.126 watt/m<sup>2</sup>

```
## # A tibble: 1 x 1
##   Mean_MB
##   <dbl>
## 1    0.138
```

Tabla 4. La media de intensidad de luz para el Module B es de 0.137 watt/m<sup>2</sup>

De las tablas anteriores se puede apreciar que las medias de intensidad de luz son similares para ambos Modulos.

## BOX-PLOT INTENSIDAD DE LUZ luz- watt/sqr 2 MODULO A y MODULO B

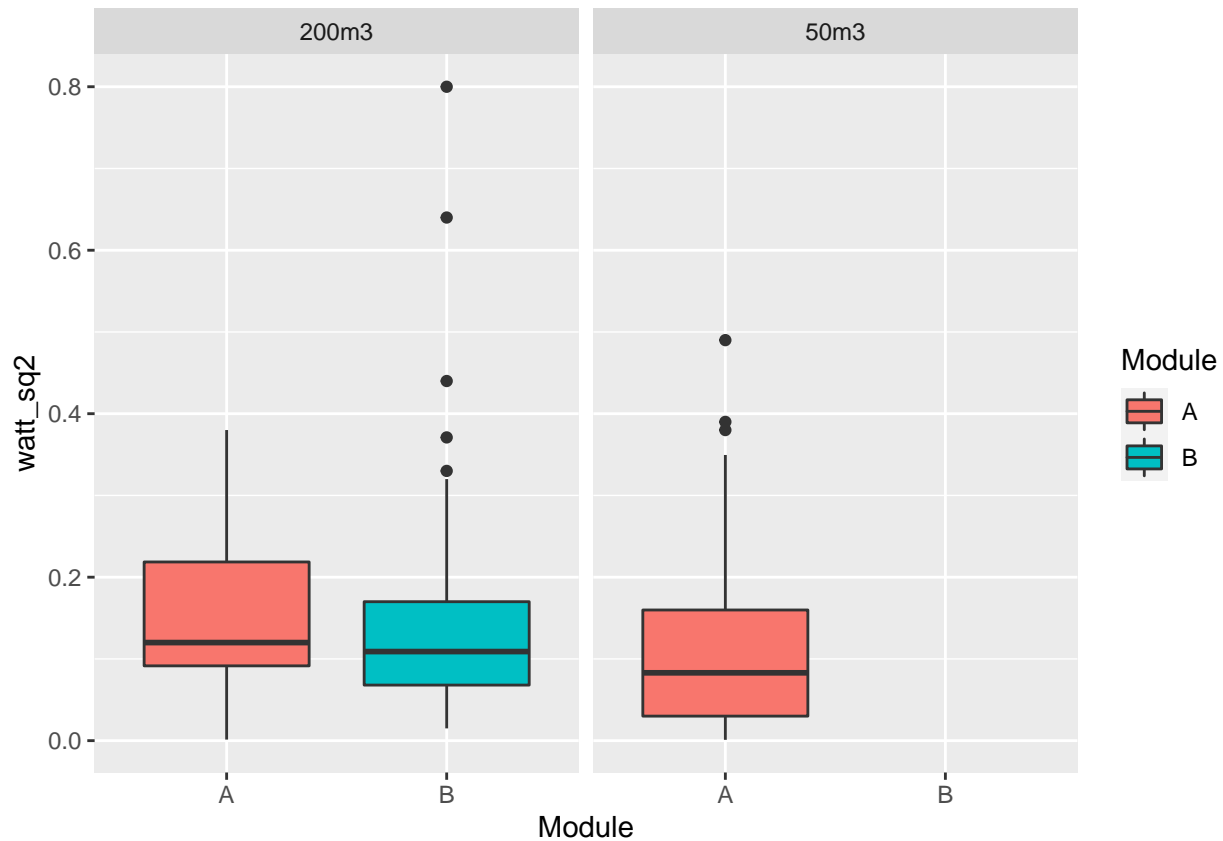


Figura 2. Box-Plot de la intensidad de luz en ambos modulos de cultivo y en estanques de 50m<sup>3</sup> y 200m<sup>3</sup>. La medida de intensidad de luz es similar para ambos modulos para estanques de 200m<sup>3</sup> y levemente menor en estanques de 50m<sup>3</sup> en el modulo A. No Hubieron mediciones de intensidad de luz en el Module B en estanques de 200m<sup>3</sup>

**EXPLORATORIO DE NORMALIDAD DE LA VARIABLE RESPUESTA light**  
**HISTOGRAMA DE LA VARIABLE RESPUESTA light watt/m2**

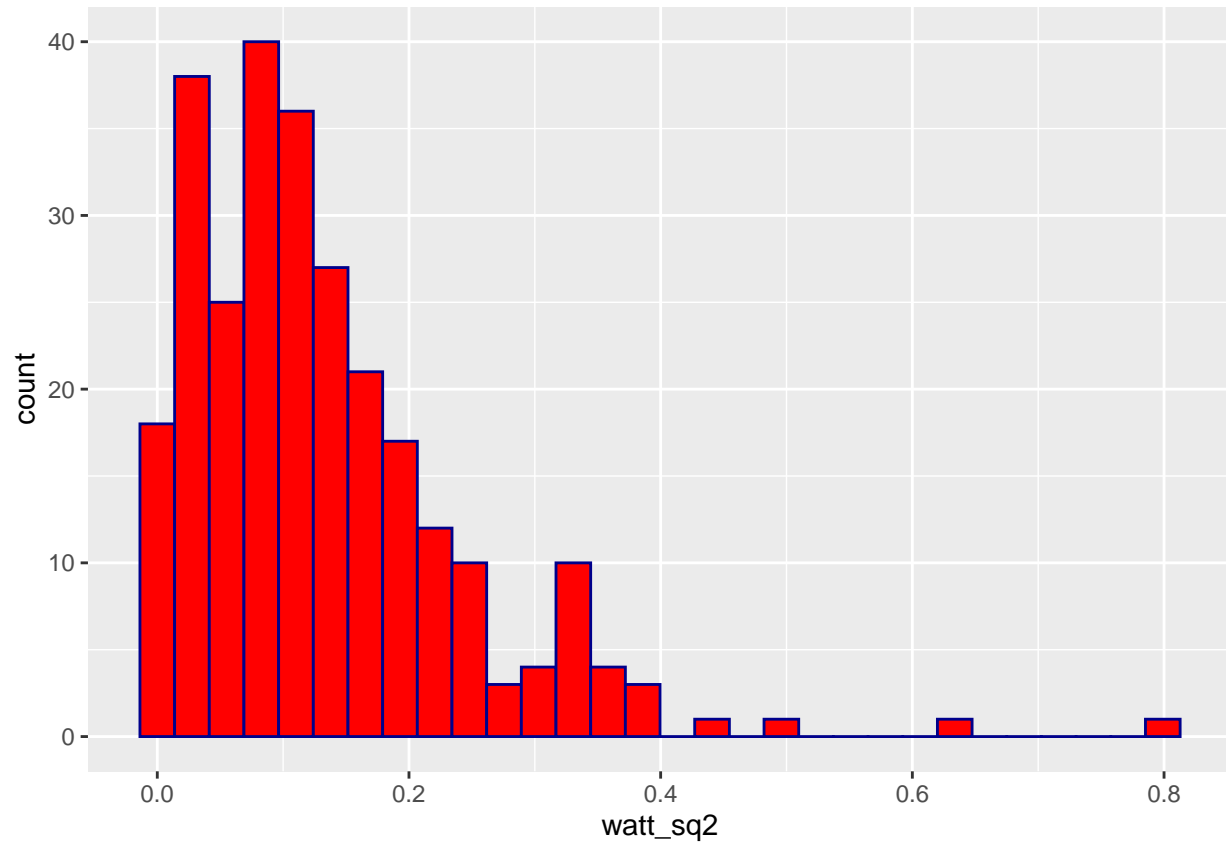


Figura 3. Histograma de la variable respuesta light. Curva con una distribucion positivamente sesgada hacia la derecha. De la figura 3 se puede observar que no presenta una tendencia de curva normal con una gran proporcion de datos concenardos desde los 0 watt/m2 hasta los 0.2 watt/m2

## HISTOGRAMA DE LA VARIABLE RESPUESTA watt/m2 CON LA CURVA NORMAL TEORICA

Histograma de light + Curva normal teórica

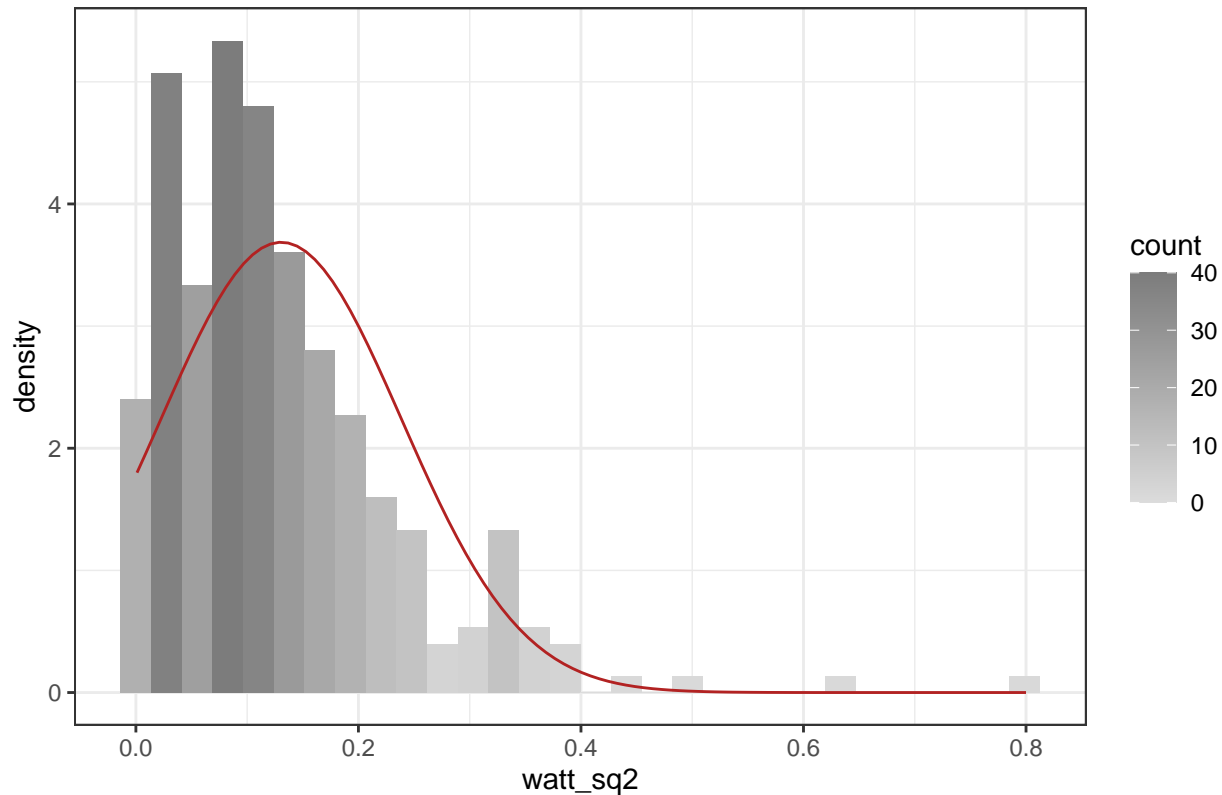


Figura 4. Histograma de la variable respuesta Light watt\_sq2 en donde se muestra la curva normal teorica. La figura 4 muestra como los datos de la variable respuesta se distribuyen en comparacion con la curva teorica normal, estas dos curvas muestran que no existe normalidad de la variable respuesta light.

## TEST DE NORMALIDAD PARA LA VARIABLE RESPUESTA light watt\_sq2

```
##  
## Shapiro-Wilk normality test  
##  
## data: light$watt_sq2  
## W = 0.86139, p-value = 6.459e-15
```

Tabla 5. Test de Normalidad para la variable respuesta light.

De acuerdo al analisis de Shapiro-Wilk mostrado en la tabla 5, test de normalidad, la variable respuesta light no presenta normalidad. Se rechaza la Hipotesis Nula.

## ANALISIS DE MODELOS LINEALES MIXTOS

### EVALUACION DE MODELOS DE EFECTOS FIJOS.

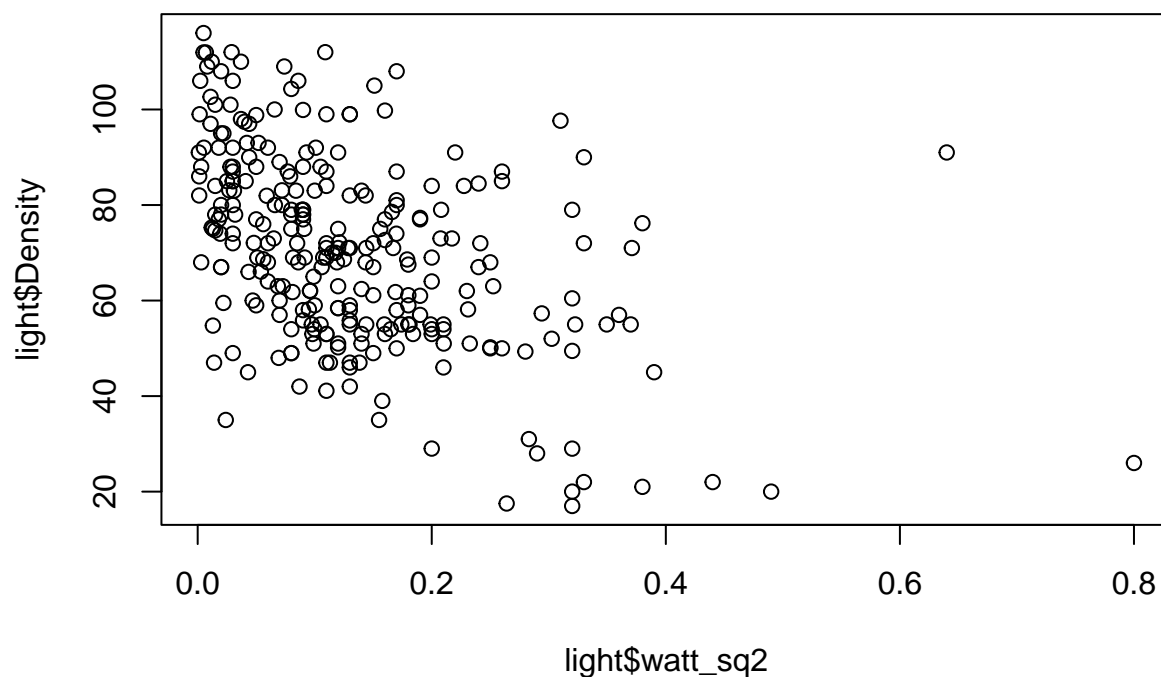
Ajuste de un modelo de efectos fijos donde la variable respuesta es light intensity y como efectos fijos del modelo las variables Density y Volumen de estanque. Density y Volumen como variables predictoras.

#### MODELO 1 DE EFECTOS FIJOS

```
##
## Call:
## lm(formula = watt_sq2 ~ Density, data = light)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.19582 -0.06262 -0.01919  0.04097  0.56107
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.3078758  0.0210796  14.605  < 2e-16 ***
## Density      -0.0025158  0.0002876  -8.749  2.37e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09569 on 270 degrees of freedom
## Multiple R-squared:  0.2209, Adjusted R-squared:  0.218
## F-statistic: 76.54 on 1 and 270 DF, p-value: 2.375e-16
```

Tabla 6: Modelo de efectos fijos 1. Intensidad de luz y densidad de cultivo. Los valores de Pr para las variables regresoras son menores al nivel de significancia del 5% lo cual son estadísticamente significativos. El ajuste del modelo es de 0.218. El p-value del modelo es menor al 5% por lo tanto tiene capacidad predictora.

## PLOT DEL MODELO 1 DE EFECTOS FIJOS



Figura

5. Plot del modelo 1 de efectos fija. Intensidad de luz y densidad de cultivo.

De la figura 5 se observa que no existe linealidad en el la variable intensidad y la densidad de cultivo.

## MODELO 2 DE EFECTOS FIJOS

```
##
## Call:
## lm(formula = watt_sq2 ~ Density + Volume_m3, data = light)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.17350 -0.05827 -0.01324  0.03965  0.54238
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.3481524   0.0219947   15.829 < 2e-16 ***
## Density       -0.0027531   0.0002813   -9.789 < 2e-16 ***
## Volume_m350m3 -0.0542924   0.0114508   -4.741 3.45e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0921 on 269 degrees of freedom
## Multiple R-squared:  0.281, Adjusted R-squared:  0.2756
## F-statistic: 52.56 on 2 and 269 DF, p-value: < 2.2e-16
```

Tabla 7 Modelo de efectos fijos 2. Intensidad de luz y densidad de cultivo. Los valores de Pr para las variables regresoras son menores al nivel de significancia del 5% lo cual son significativas. El ajuste del modelo es de 0.2756 y tiene capacidad predictora.



## PLOT DEL MODELO 2 AJUSTADO DE EFECTOS FIJOS

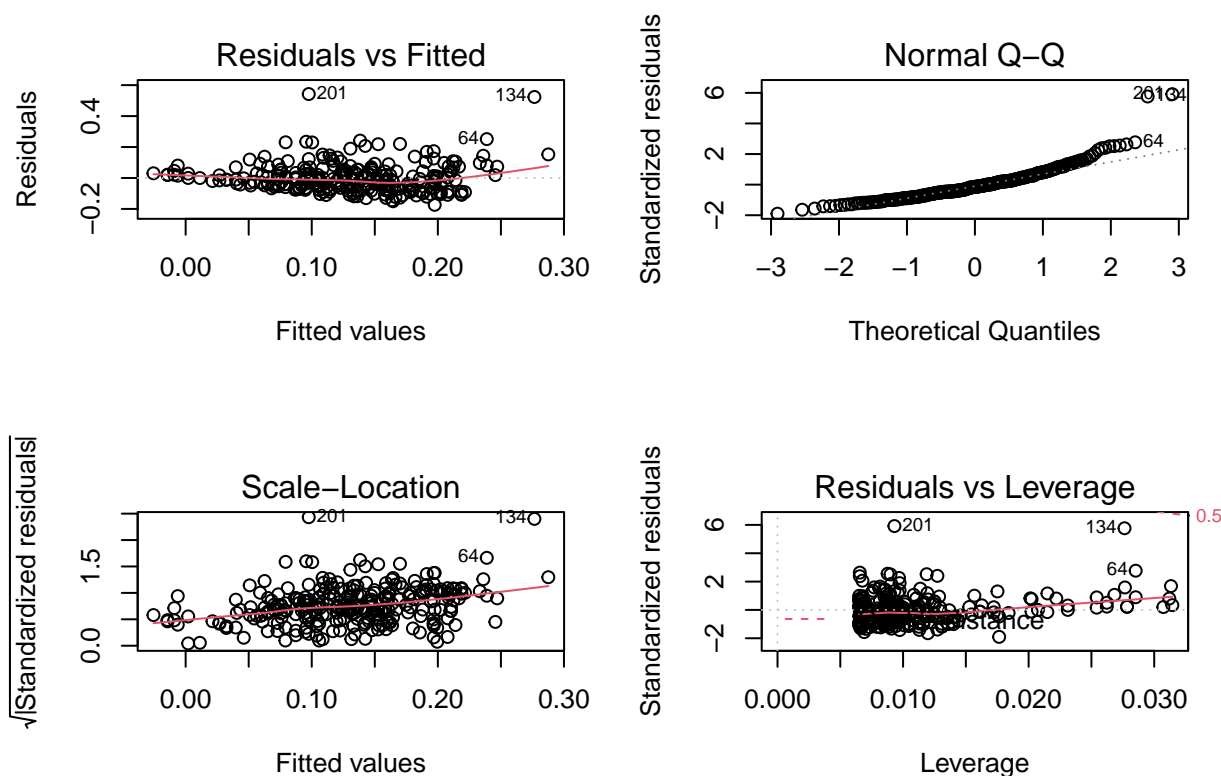


Figura 6. Plot del modelo 2 ajustado. Variable respuesta light y como efectos fijos densidad de cultivo y volumen del estanque.

La figura 6 muestra la curva NORMAL Q-Q en donde se aprecia que los valores de los residuales estandarizados no preentan linealidad, por lo cual el modelo no presenta normalidad.

## TEST DE NORMALIDAD MODELOS EFECTOS FIJOS

```
##
## Shapiro-Wilk normality test
##
## data: residuals(mod.1a)
## W = 0.87095, p-value = 2.356e-14

##
## Shapiro-Wilk normality test
##
## data: residuals(mod.1b)
## W = 0.87704, p-value = 5.568e-14
```

Tabla 8: Analisis Anova para los modelos de Efectos fijos.

La tabla 8 muestra el resultado del analisis de Shapiro-Wilk para Normalidad.El valor de p-value para ambos modelos es menor al nivel de significancia del 5% lo cual no presentan normalidad.

## MODELOS INTEGRADOS.

### MODELO LINEAL INTEGRADO 1

```
##
## Call:
```

```
## lm(formula = watt_sq2 ~ Density:Volume_m3, data = light)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.18439 -0.05824 -0.01363  0.04089  0.54419
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.316281   0.020326  15.560 < 2e-16 ***
## Density:Volume_m3 200m3 -0.002326   0.000279  -8.336 4.01e-15 ***
## Density:Volume_m3 350m3 -0.003083   0.000300 -10.276 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09194 on 269 degrees of freedom
## Multiple R-squared:  0.2835, Adjusted R-squared:  0.2782
## F-statistic: 53.22 on 2 and 269 DF,  p-value: < 2.2e-16
```

Tabla 9: Modelo 2 de efectos fijos integrado .Intensidad de luz, y el efecto integrado entre la densidad de cultivo y el volumen del estanque. Los valores de Pr del modelo es menor al nivel de significancia del 5% lo cual el modelo tiene capacidad predictora. El ajuste del modelo es de 0.2782.

## PLOT DE MODELO 1 DE EFECTOS FIJOS INTEGRADO

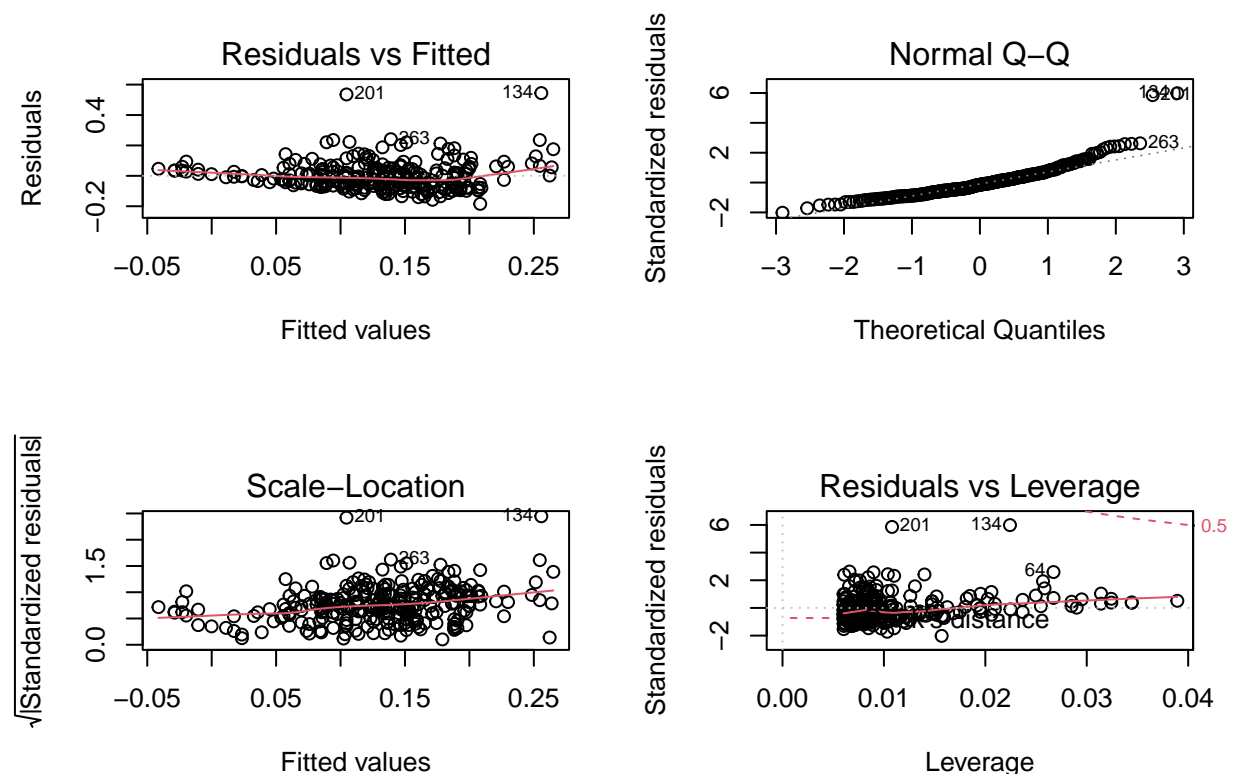


Figura 7. Plot del modelo 1 de efectos fijos integrado. Variable respuesta light y como efectos fijos densidad de cultivo y volumen del estanque.

De la Tabla 7 se observa que la curva Normal Q-Q no presneta normalidad.

## MODELO LINEAL INTEGRADO 2

```
##
## Call:
## lm(formula = watt_sq2 ~ Density + Volume_m3 + Density:Volume_m3,
##     data = light)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.18157 -0.05829 -0.01257  0.04063  0.53894
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.3241790   0.0320363   10.119 < 2e-16 ***
## Density        -0.0024275   0.0004233   -5.734 2.64e-08 ***
## Volume_m350m3  -0.0132485   0.0414936   -0.319  0.750
## Density:Volume_m350m3 -0.0005829  0.0005664   -1.029  0.304
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09209 on 268 degrees of freedom
## Multiple R-squared:  0.2838, Adjusted R-squared:  0.2758
## F-statistic: 35.4 on 3 and 268 DF, p-value: < 2.2e-16
```

Tabla 10: Modelo 2 de efectos integrado .Intensidad de luz, y el efecto integrado entre la densidad de cultivo y el volumen del estanque mas el efecto del la densidad y el volumen del estqanke. Los valores de Pr para las variables explicativas Volume\_m350m3 y Density:Volume\_m350m3 son menores al nivel de significancia del 5% lo cual esas variables de modelo son estadisticamente significativas. El ajuste del modelo es de 0.2782.

## PLOT DEL MODELO

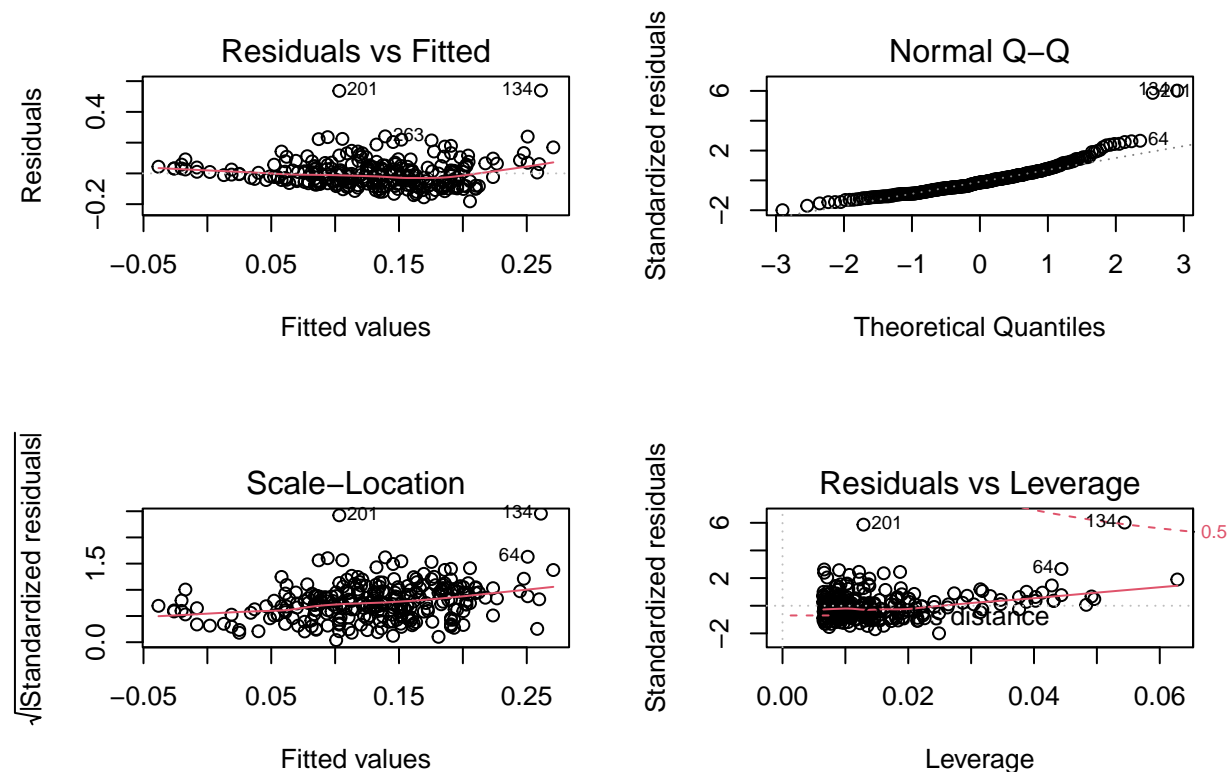


Figura 8. Plot del modelo 2 integrado . Variable respuesta light y como efectos fijos densidad de cultivo y volumen del estanque mas la interaccion entre la densidad de cultivo y el volumen del estanque.

De la Tabla 8 se observa que la curva Normal Q-Q no presneta normalidad. De la figura Fitted Values se aprecia una gran dispercion de los datos con una tendencia mas bien curva que lineal.

## TEST DE NORMALIDAD PARA LOS MODELOS INTEGRADOS

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals(mod.1.int)
## W = 0.87391, p-value = 3.563e-14

##
##  Shapiro-Wilk normality test
##
## data:  residuals(mod.2.int)
## W = 0.87474, p-value = 4.01e-14
```

Tabla 11: Analisis Anova para los modelos de Efectos Integrados.

La tabla 11 muestra el resultado del analisis de Shapiro-Wilk para Normalidad.El valor de p-value para ambos modelos es menor al nivel de significancia del 5% lo cual no presentan normalidad.

## MODELOS CON EFECTOS ALEATORIOS

### MODELOS 1 CON EFECTOS ALEATORIOS

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: watt_sq2 ~ Density + Volume_m3 + (1 | Module)
##   Data: light
##
## REML criterion at convergence: -498.3
##
## Scaled residuals:
##   Min       1Q   Median       3Q      Max
## -1.8838 -0.6327 -0.1438  0.4305  5.8890
##
## Random effects:
##   Groups   Name                Variance Std.Dev.
##   Module   (Intercept)  0.000000  0.0000
##   Residual                  0.008483  0.0921
## Number of obs: 272, groups:  Module, 2
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   0.3481524  0.0219947  15.829
## Density       -0.0027531  0.0002813  -9.789
## Volume_m350m3 -0.0542924  0.0114508  -4.741
##
## Correlation of Fixed Effects:
##              (Intr) Densty
## Density      -0.941
## Volm_m350m3 -0.386  0.178
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

Tabla 12 Modelo 1 de efectos aleatorios. Variable respuesta intensidad de luz y como variables regresoras densidad del estanque mas el volumen y como efecto aleatorio el modulo de cultivo.

La tabla 12 muestra el analisis del modelo lineal integrado sobre la variable respuesta light, el efectos aleatorio de la variable Module con un intercepto de 0 en donde esta variable no tiene un efecto aleatorio como variable regresora. Tambien se muestran el efecto fijo de las variables regresoras Density y Volume 50m3 con una correlacion baja de 0.178 entre ellas.

## MODELO 2 DE EFECTO ALEATORIO

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: watt_sq2 ~ Density + Volume_m3 + Density:Volume_m3 + (1 | Module)
## Data: light
##
## REML criterion at convergence: -486.2
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.9716 -0.6329 -0.1365  0.4412  5.8522
##
## Random effects:
## Groups Name Variance Std.Dev.
## Module (Intercept) 0.000000 0.00000
## Residual 0.008481 0.09209
## Number of obs: 272, groups: Module, 2
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)    0.3241790  0.0320363  10.119
## Density        -0.0024275  0.0004233  -5.734
## Volume_m350m3  -0.0132485  0.0414936  -0.319
## Density:Volume_m350m3 -0.0005829  0.0005664  -1.029
##
## Correlation of Fixed Effects:
##              (Intr) Densty V_3503
## Density      -0.973
## Volm_m350m3  -0.772  0.751
## Dnst:V_3503  0.727 -0.747 -0.961
## optimizer (nloptwrap) convergence code: 0 (OK)
## boundary (singular) fit: see ?isSingular
```

Tabla 13: Modelo 2 de efectos aleatorios. Variable respuesta intensidad de luz y como variables regresoras densidad del estanque mas el volumen mas la interaccion entre la densidad y el volumen y como efecto aleatorio el modulo de cultivo.

La tabla 13 muestra el analisis del modelo lineal integrado sobre la variable respuesta light, el efectos aleatorio de la variable Module con un intercepto de 0 en donde esta variable no tiene un efecto aleatorio como variable regresora. Tambien se muestran el efecto fijo de las variables regresoras Density y Volume mas el efecto fijo entre las variables Density y Volume La correlacion del efecto fijo entre la relacion de la variable Density y los estanques de 50m3 es de 0.727.

## TEST DE NORMALIDAD MODELOS EFECTOS ALEATORIOS

```
##
## Shapiro-Wilk normality test
##
```

```
## data: residuals(mod.A.1)
## W = 0.87704, p-value = 5.568e-14

##
## Shapiro-Wilk normality test
##
## data: residuals(mod.A.2)
## W = 0.87474, p-value = 4.01e-14
```

Tabla 14: Tes de Normalidad para los modelos aleatorios.

La tabla 14 indica que desde el analisis de normalidad Shapiro-Wilk los modelos no presentan Normalidad.

## CONCLUSION ANALISIS 1:

Los modelos lineales modelados anteriormente asi como la variable respuesta Watt\_sq2 no presentan una distribucion normal, por lo tanto no se cumplen con los supuestos de los Modelos Lineales Mixtos.

La variable predictora Volume\_m3 presenta solo dos niveles (binomial) 50m3 y 200 m3 por lo tanto no pueden ser considerado como un efecto aleatorio.

De los analisis de normalidad de los modelos lineales fijos, efectos fijos integrados y efectos aleatorios se observa que no presentan normalidad por lo tanto el analisis no se puede llevar a cabo mediante los modelos lineales antes descritos.

No se puede determinar mediante un analisis de modelos lineales de efectos fijos y efectos aleatorios el efecto de la densidad y el volumen en la intensidad de la luz medida en los estanques. La variable respuesta intensidad de luz: watt\_sq2 No tiene distribucion normal y la variable predictora Volume\_m3 es categorica binomial.

De las tablas de 5, 8, 11 y 14 analisis de normalidad Shapiro-Wilk para la variable regresora light, Modelos de efectos fijos, modelos lineales integrados y Modelos aleatorios, respectivamente se evidencia que los modelos no presentan normalidad.

## ANALISIS PARTE 2

### ANALISIS DE MODELOS LINEALES GENERALIZADOS PARA EL ANALISIS DEL EFECTO DE LA DENSIDAD DE CULTIVO, EL VOLUMEN DEL ESTANQUE SOBRE LA INTENSIDAD DE LUZ

#### METODOLOGIA ANALISIS 2:

- 1- Se elegira a priori ua distribucion para la variable respuesta light basada en su relacion media varianza.
- 2- Se elegira una funcion de enlace para la funcion de distribucion de la variable respuesta light dentro del modelo lineal generalizado.
- 3- Se construiran modelos lineales generalizados con las variables predictoras Density y Volume.
- 4- Se calculara la devianza del modelo elegido y se comparara con un modelo saturado el cual incluire la mayor cantidad de variables predictoras.

### DISTRIBUCION DE LA VARIABLE RESPUESTA light watt\_sq2

De acuerdo a la figura 4 Histograma de la variable respuesta light se observa que la variable respuesta light se asemeja a una distribucion Gamma.

## FUNCION DE ENLACE PARA LA VARIABLE RESPUESTA light watt\_sq2

Como la variable light tiene un distribucion Gamma se usara el enlace canonico “inverse”

### MODELOS GENERALIZADOS (GAMMA).

#### MODELO 1 GAMMA

Modelo 1 nulo

```
##
## Call:
## glm(formula = watt_sq2 ~ 1, family = Gamma(link = "inverse"),
##      data = light)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7856  -0.8024  -0.1667   0.3094   2.5745
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   7.6582     0.3848   19.9   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.6867737)
##
##      Null deviance: 233.56  on 271  degrees of freedom
## Residual deviance: 233.56  on 271  degrees of freedom
## AIC: -568.45
##
## Number of Fisher Scoring iterations: 6
```

Tabla 15: GLM 1 Nulo. Representa solo el parametro de respuesta light representando la media global u para tof=dos los y.

La tabla 15 muestra el resultado del GLM nulo con Residual deviance de 233.56 con 271 degree of freedom y AIC -568.45 el valor de pr es menor al nivel de significancia del 5% por lo tanto el valor del intercepto es estadisticamente significativo.

#### MODELO 2 GAMMA

Modelo Saturado

```
##
## Call:
## glm(formula = watt_sq2 ~ Density + Density:Volume_m3 + Density:Module +
##      Module:Volume_m3, family = Gamma(link = "inverse"), data = light)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6495  -0.6479  -0.1885   0.3131   2.6266
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -1.113739    0.780500  -1.427    0.155
```

```
## Density                0.112445    0.015041    7.476 1.09e-12 ***
## Density:Volume_m350m3  0.079177    0.015599    5.076 7.22e-07 ***
## Density:ModuleB        0.007767    0.012502    0.621    0.535
## Volume_m350m3:ModuleB      NA          NA          NA          NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.5940503)
##
## Null deviance: 233.56 on 271 degrees of freedom
## Residual deviance: 176.80 on 268 degrees of freedom
## AIC: -647.2
##
## Number of Fisher Scoring iterations: 6
```

Tabla 16. GLM Saturado. El modelo GLM2 completo representa la interaccion de la variable respuesta light con la densidad de cultivo mas la interaccion entre la variable predictora Density y Volume mas la interaccion entre la variable predictora Density y Module mas la interaccion entre Module y Volume.

La tabla 16 muestra que el intercepto del modelo es significativamente distinto de 0 por lo cual es estadisticamente significativo al igual que el p-values de la variable Density y la interaccion entre Density y Volume.

El p-value de la interaccion volume y Module es mayor al nivel de significancia del 5% por lo cual no es estadisticamente significativo. El valor de AIC para el modeloSaturado es de -647.2

## MODELO 3 GAMMA

Modelo Completo.

```
##
## Call:
## glm(formula = watt_sq2 ~ Density:Volume_m3, family = Gamma(link = "inverse"),
## data = light)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.6332  -0.6442  -0.1634   0.2967   2.5639
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -1.11302    0.76381  -1.457    0.146
## Density:Volume_m3200m3  0.11687    0.01314   8.897 <2e-16 ***
## Density:Volume_m350m3  0.19161    0.01904  10.064 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Gamma family taken to be 0.5851563)
##
## Null deviance: 233.56 on 271 degrees of freedom
## Residual deviance: 177.03 on 269 degrees of freedom
## AIC: -648.82
##
## Number of Fisher Scoring iterations: 6
```

Tabla 17. GLM 3 Completo. El modelo saturado completo considera las variables regresoras que son



estadísticamente significativas, considerando el menor número de parámetros posibles.

La tabla 17 muestra que todos los valores p-value para las variables regresoras son estadísticamente significativas con un AIC de -648.82

## COMPARACION DE LOS MODELOS MEDIANTE ANALISIS ANOVA

```
## Analysis of Deviance Table
##
## Model 1: watt_sq2 ~ 1
## Model 2: watt_sq2 ~ Density + Density:Volume_m3 + Density:Module + Module:Volume_m3
## Model 3: watt_sq2 ~ Density:Volume_m3
##   Resid. Df Resid. Dev Df Deviance
## 1      271      233.56
## 2      268      176.80  3    56.759
## 3      269      177.03 -1    -0.228
```

Tabla 18. Valores de Df residuales para los tres GLM.

La tabla 18 muestra que el modelo 2 “Completo” tiene un menor valor de diferencia de residuos lo que podríamos suponer que es el mejor modelo que se ajusta a la interacción entre las variable respuesta y las regresoras.

## CALCULO DE LOS p-Value PARA LOS TRES MODELOS

```
## [1] 0.9515296
## [1] 0.9999962
## [1] 0.9999836
```

Tabla 19. Tabla de p-values para los tres modelos GLM.

La tabla 19 muestra que los tres modelos son adecuados para el set de datos. Ninguno de los p-values se rechaza la adecuación de los modelos al 5%.

#RESUMEN DE LOS AIC DE LOS TRES GLM

Table 1: Tabla 20.Resumen AIC Modelos Generales Linealizados

GLM_Nulo	AIC: -568.45
GLM_SAT	AIC: -647.2
GLM_COMPL	AIC: -648.82

Tabla 20. Resumen de los AIC para los tres modelos GLM.

La tabla 20 muestra que el menor de los AIC corresponde al modelo completo por lo tanto es el modelo que elegimos como el mejor que representa la interacción entre la variable respuesta light watt\_sq2 y las variables regresoras.

## PORCENTAJE DE VARIABILIDAD EXPLICADO POR EL MODELO COMPLETO.

## [1] 24.20363

Tabla 21. Devianza del modelo Completo.

La tabla 21 muestra la  $D^2$  del modelo Completo. El modelo GLM 3 Completo representa en 24.20% la interaccion entre la variable respuesta light y las variables regresoras Density, Volume y Module.

## PLOT DE GLM 3 COMPLETO.

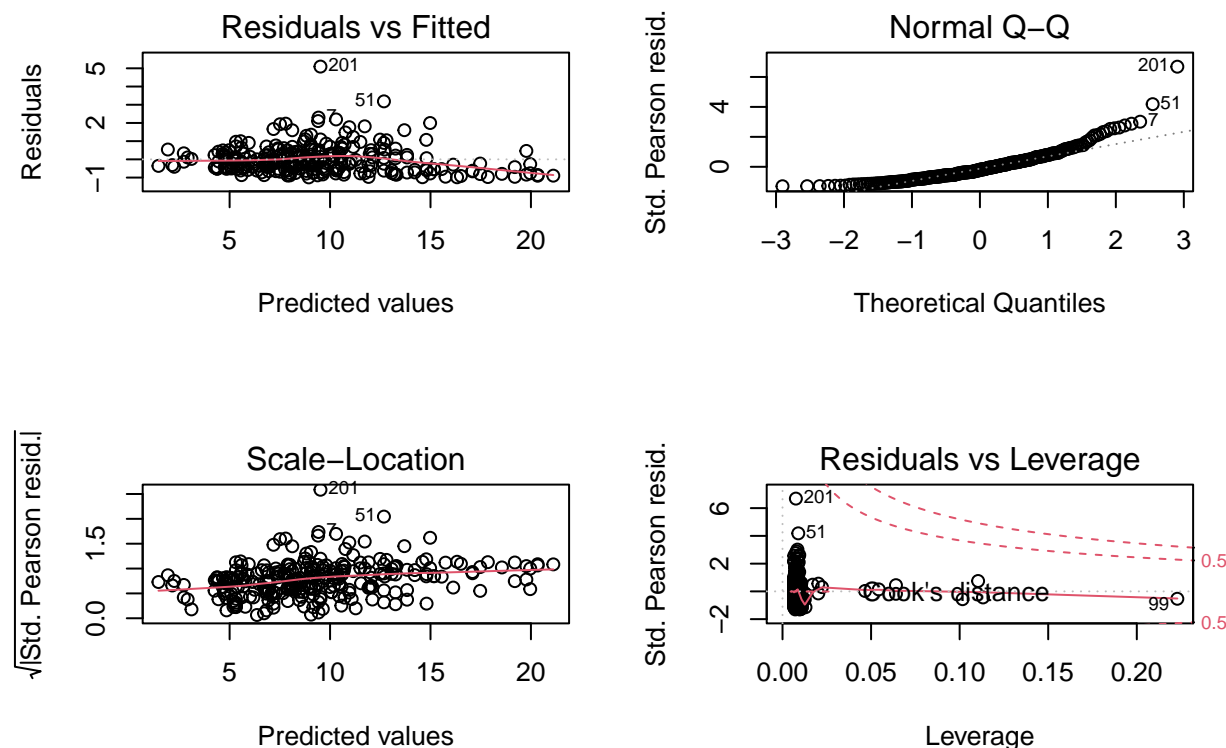


Figura 9. PLOT de GLM3.

La figura 9 muestra el la curva de Residual vs Fitted la correlacion de los residuales versus los ajustados, muestra la aproximacion y relacion de los errores con la media de la poblacional. Esto muestra la poca variabilidad de la variable respuesta con las variables regresoras explicado en la tabla 21 con el analisis de la Deviance.

De la figura Residual vs Leverage no se ven puntos sobre las lineas roja segmentadas por lo cual no se ven alto valores sobre la distancia Cook.

## CONCLUSION ANALISIS 2:

Mediante un analisis modelos generales linealizados se puede predecir el efecto de la densidad y del tamano del estanques sobre la intensidad de la luz medida.

De los tres GLM modelados el mejor modelo que se ajusta al set de datos es el GLM 3 Completo, pero solo tiene capacidad de explicar en un 24.20% la variabilidad de los datos analizados.

Si bien los GLM explican la variabilidad de los datos para este analisis, es necesario otro tipo de analisis mas complejo como el Analisis de componentes principales para explicar de mejor manera la interaccion de las variables analizadas.