R3. Monitoreo continuo de variables que influyen en la cosecha de piña para determinar los factores limitantes

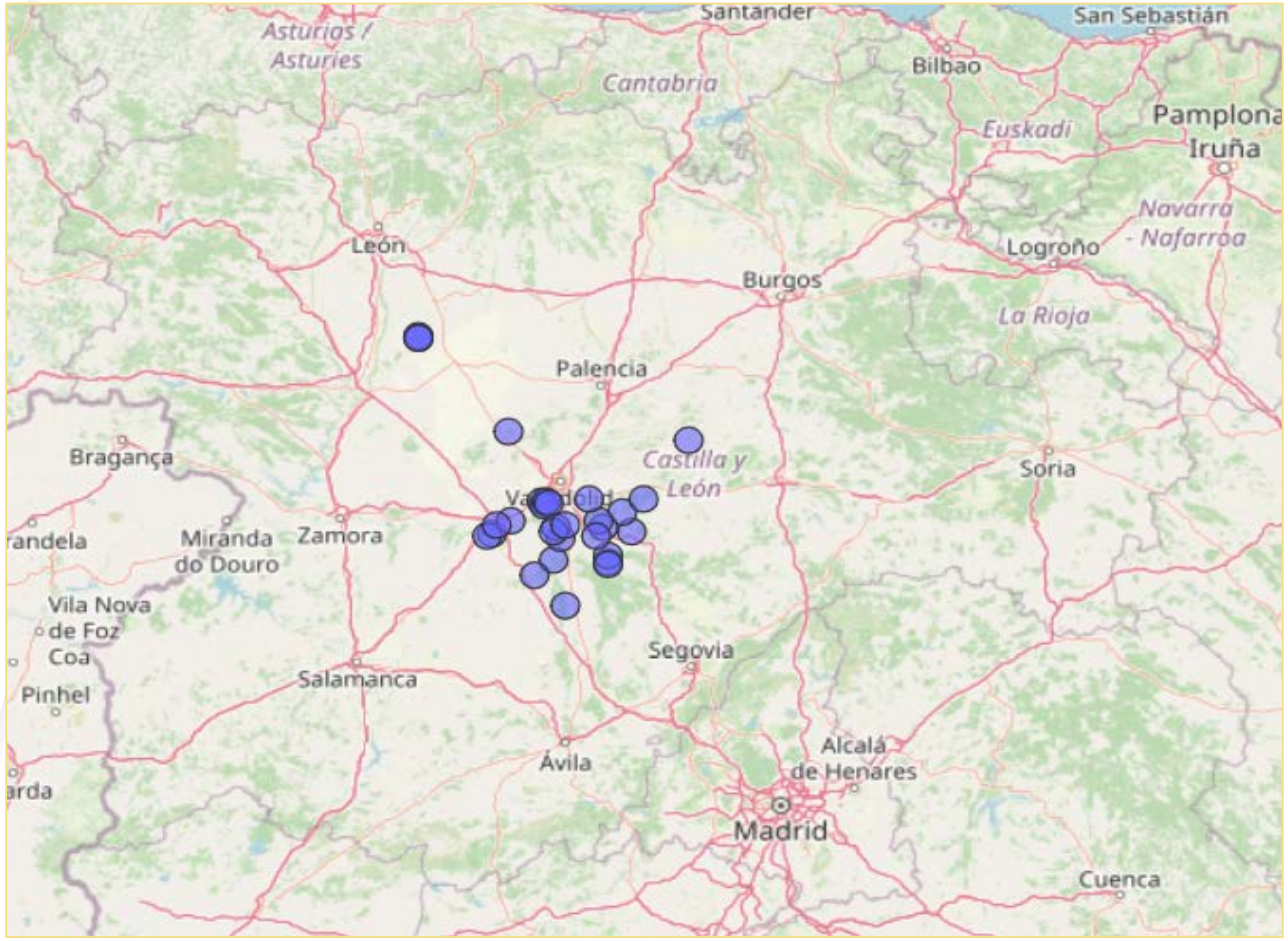
Resultados del trabajo de sensores remotos e índices de afección

Ángela Blázquez Casado
Coordinación GO PINEA

cesefor
Centro de Servicios y Promoción Forestal y de su Industria de Castilla y León

1. Localización de las parcelas

- 58 parcelas de la red de seguimiento de *Pinus pinea* de la Junta de Castilla y León y el INIA
- Años de estudio 2013-2020



2. Variables respuesta.

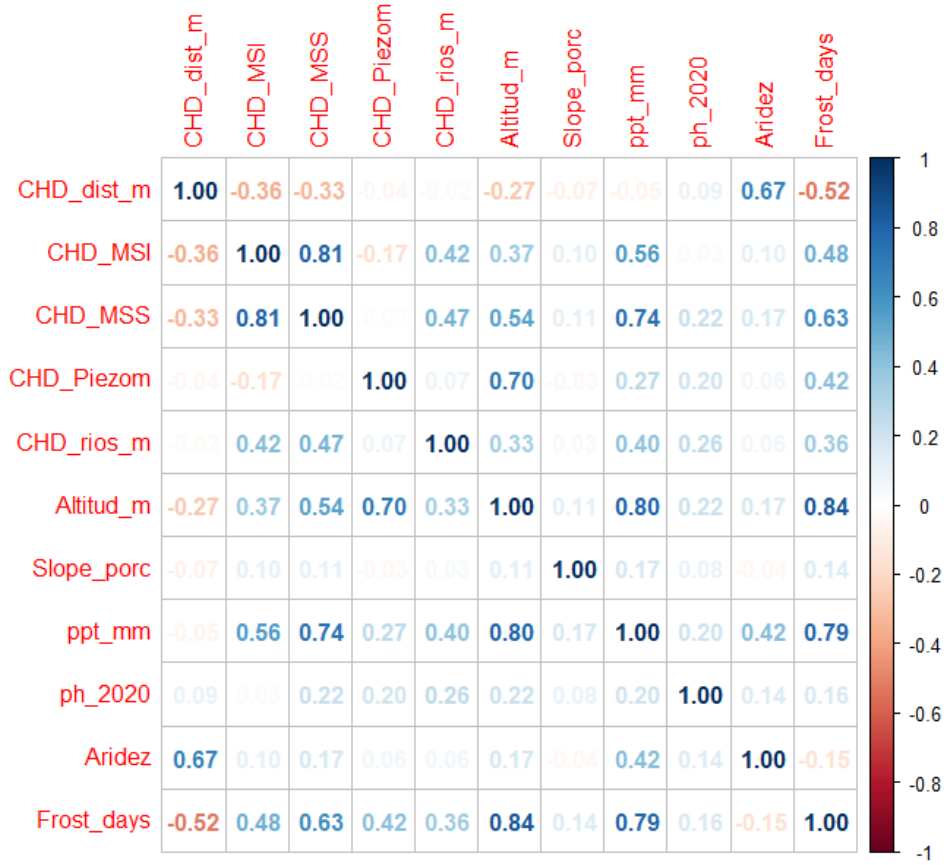
- PV – Peso medio de la piña en verde a una humedad estándar del 37% (g)
- Npp – Número de piñones por piña
- Pmpcc – Peso medio de un piñón con cáscara (g)
- %_VANOS – Porcentaje de piñón dañado
- REND_FINAL – Rendimiento final: Kg de piñón blanco / Kg de piña verde

3. Monitoreo de variables climáticas, sensores remotos y suelo.

- **VARIABLES DE SUELO:**
  - **Tipo de suelo: Arcillas, dunas, fondos de valle, etc..**
  - **Distancias a manantiales, profundidad de la capa freática, distancia a ríos**
  - **Altitud, pendiente, ph, textura, aridez, idas de heladas, profundidad del suelo**
- **VARIABLES SATELITALES**
  - Landsat: Bandas 4,5,7 + Índices NDVI, NBR, SAVI, TCGVI, TCWET, etc…
  - Sentinel 1: Bandas VH, VV + Índices VV_VH
  - Sentinel 2: Bandas 3,4,8A ,12 + Índices NDVI, NBR, SAVI, TCGVI, TCWET, etc…
- **VARIABLES CLIMÁTICAS**
  - Evapotranspiración potencial (pet)
  - Radiación de onda corta de superficie descendente (srad)
  - Déficit de presión de vapor (cpd)
  - Temperatura máxima del mes (tmmx)
  - Etc.

3. Metodología.

- **Análisis de correlación (>0,8)**
- **Eliminación de variables con mas de un 50% de los datos vacíos**
- **Completado de la base de datos.** Method: cart
  - Total inicial 486 variables
  - Total final 140 variables
- **Preparación de la muestra**
  - Train: 75%
  - Test: 25 %

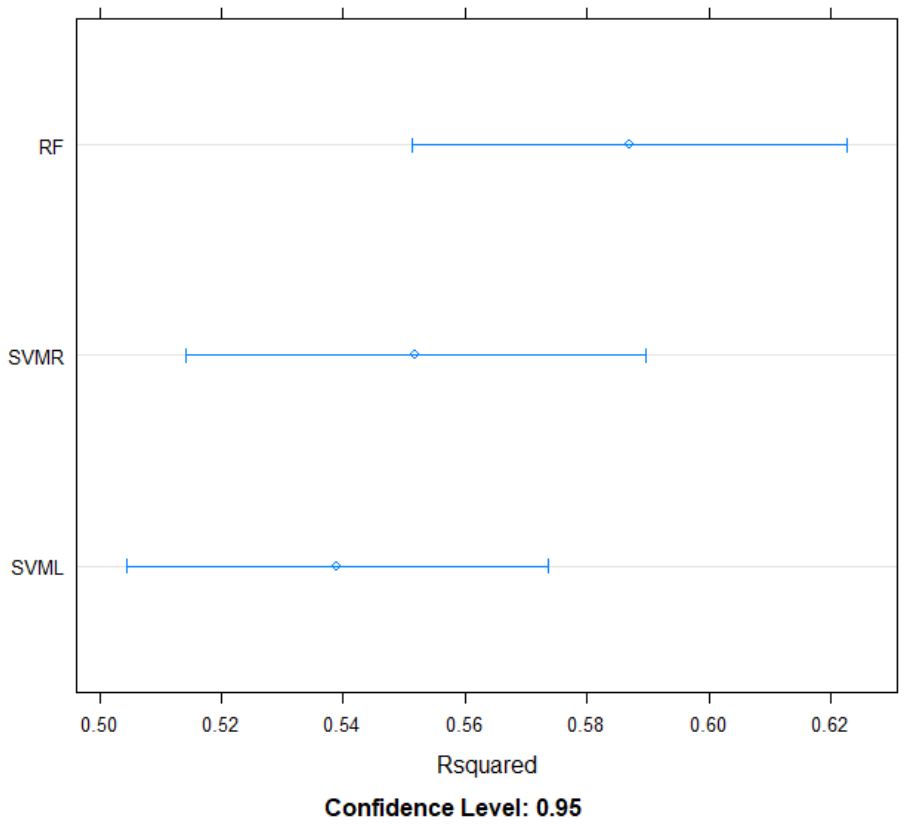


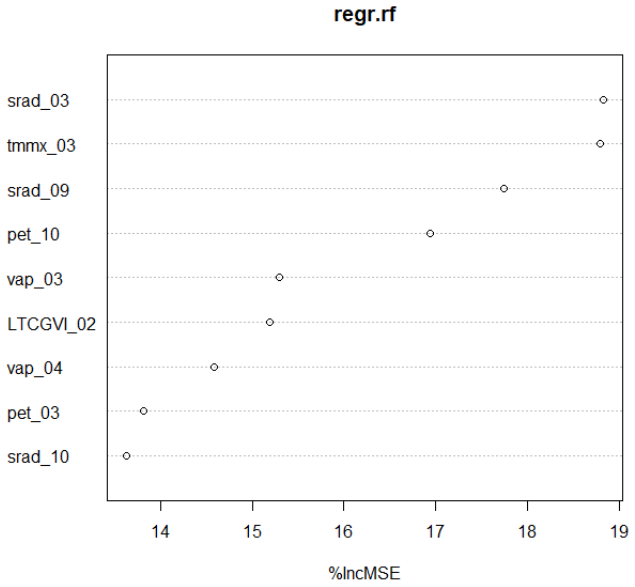
	CHD_dist_m	CHD_MSI	CHD_MSS	CHD_Piezom	CHD_rios_m	Altitud_m	Slope_porc	ppt_mm	ph_2020	Aridez	Frost_days
CHD_dist_m	1.00	-0.36	-0.33	-0.04	-0.02	-0.27	-0.07	-0.05	0.09	0.67	-0.52
CHD_MSI	-0.36	1.00	0.81	-0.17	0.42	0.37	0.10	0.56	0.03	0.10	0.48
CHD_MSS	-0.33	0.81	1.00	[illegible]	0.47	0.54	0.11	0.74	0.22	0.17	0.63
CHD_Piezom	-0.04	-0.17	[illegible]	1.00	0.07	0.70	-0.03	0.27	0.20	0.06	0.42
CHD_rios_m	-0.02	0.42	0.47	0.07	1.00	0.33	0.03	0.40	0.26	0.06	0.36
Altitud_m	-0.27	0.37	0.54	0.70	0.33	1.00	0.11	0.80	0.22	0.17	0.84
Slope_porc	-0.07	0.10	0.11	-0.03	0.03	0.11	1.00	0.17	0.08	-0.04	0.14
ppt_mm	-0.05	0.56	0.74	0.27	0.40	0.80	0.17	1.00	0.20	0.42	0.79
ph_2020	0.09	0.03	0.22	0.20	0.26	0.22	0.08	0.20	1.00	0.14	0.16
Aridez	0.67	0.10	0.17	0.06	0.06	0.17	-0.04	0.42	0.14	1.00	-0.15
Frost_days	-0.52	0.48	0.63	0.42	0.36	0.84	0.14	0.79	0.16	-0.15	1.00

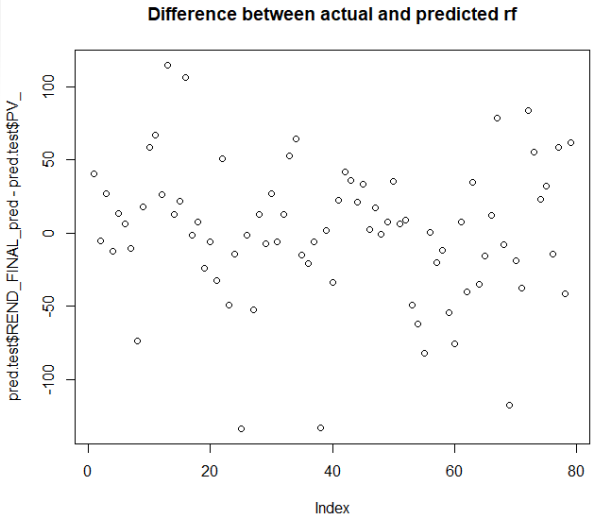
4. Modelos de regresión

Variable PV – Peso medio de la piña en verde con humedad corregida 37%

```
Call:
 randomForest(formula = PV_ ~ ., data = df.train, importance = T)
               Type of random forest: regression
                     Number of trees: 500
No. of variables tried at each split: 3

          Mean of squared residuals: 2382.137
                    % Var explained: 56.23
>
```



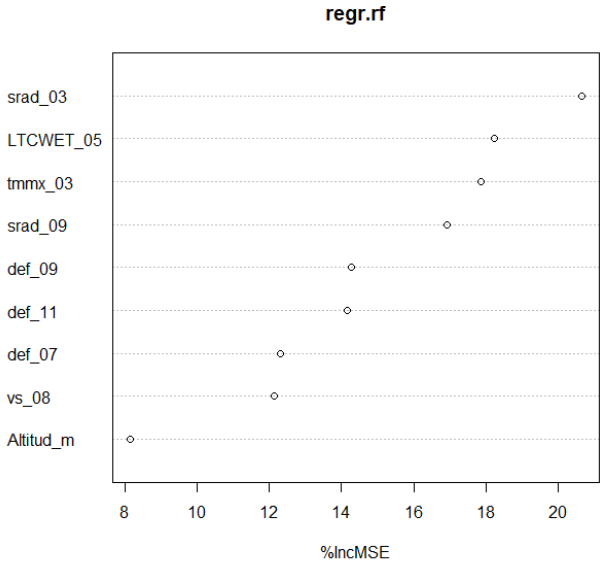


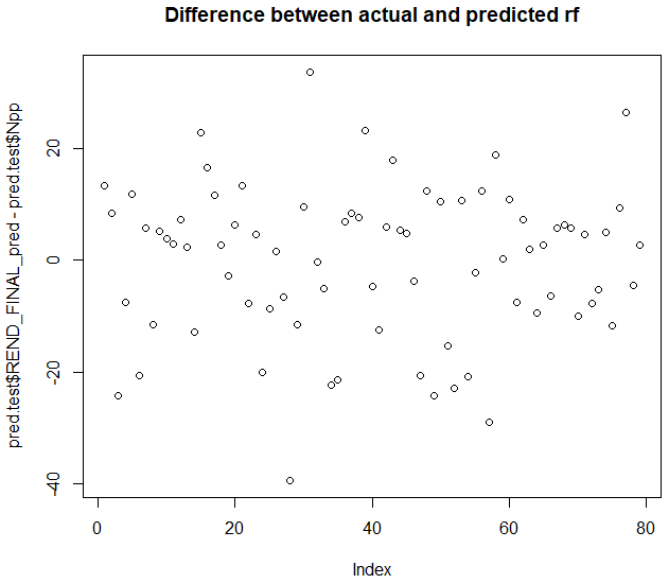


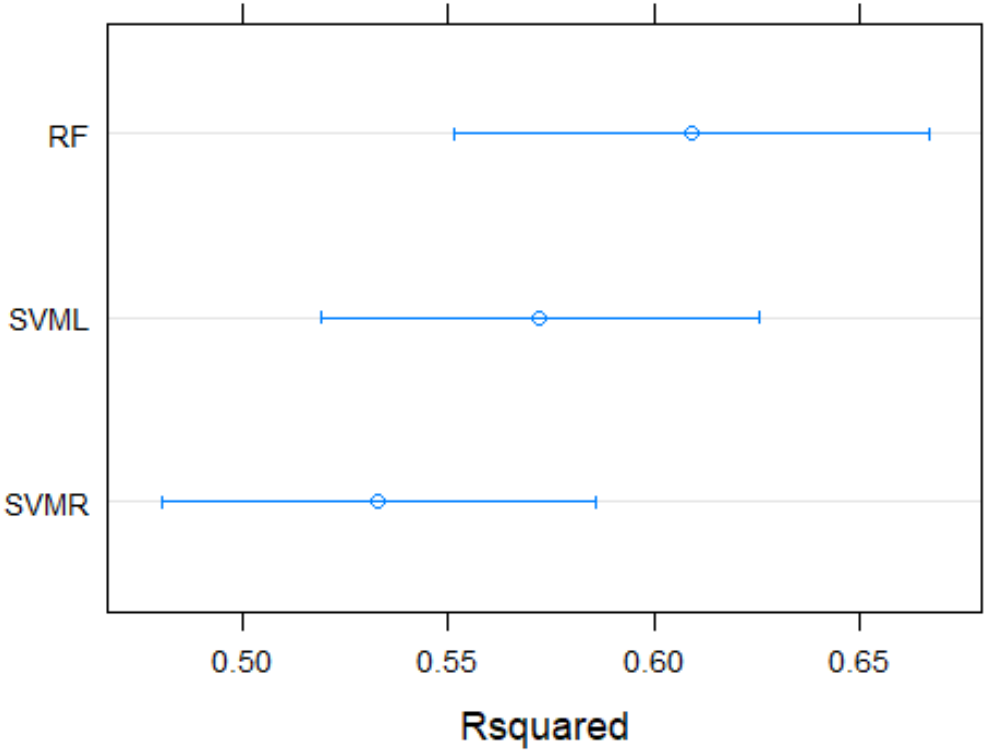
4. Modelos de regresión

Variable Npp - Número de piñones por piña

```
Call:
 randomForest(formula = Npp ~ ., data = df.train, ntree = 1500)
               Type of random forest: regression
                     Number of trees: 1500
No. of variables tried at each split: 3

          Mean of squared residuals: 253.0786
                    % Var explained: 56.62
```



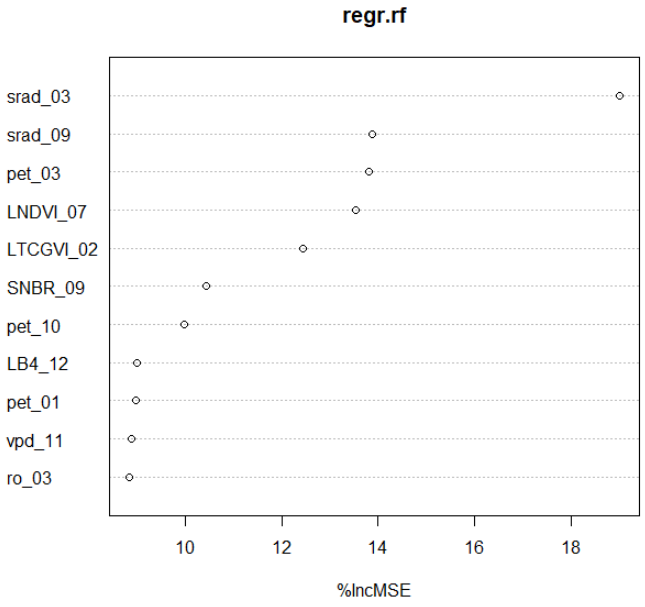


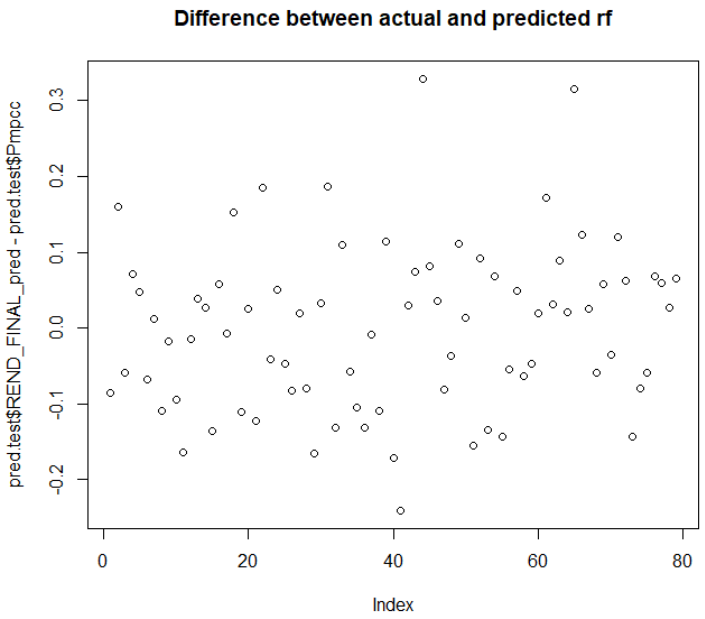


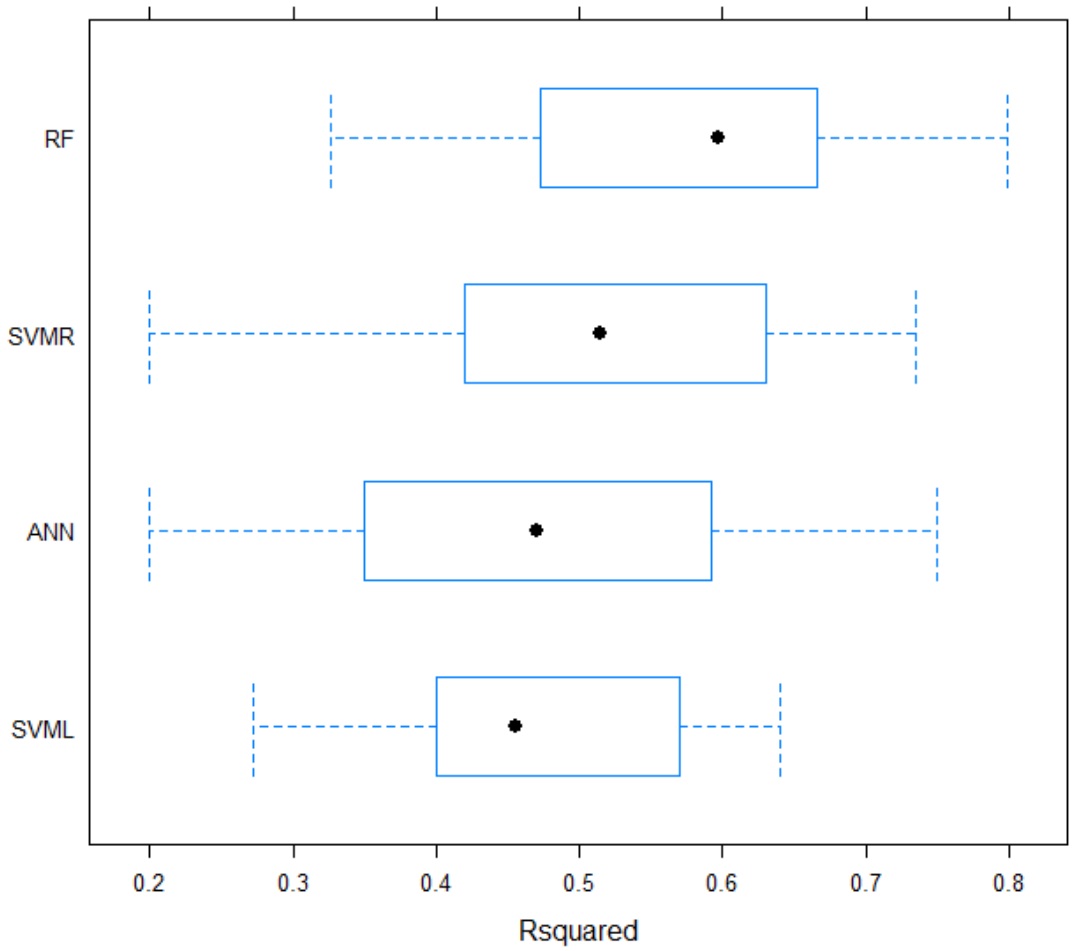
4. Modelos de regresión

Variable Pmpcc - Peso medio de un piñón con cáscara (g)

```
Call:
 randomForest(formula = Pmpcc ~ ., data = df.train, importance = T)
               Type of random forest: regression
                     Number of trees: 500
No. of variables tried at each split: 6

          Mean of squared residuals: 0.01005881
                    % Var explained: 57.61
```





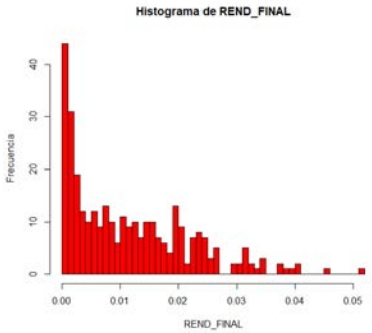


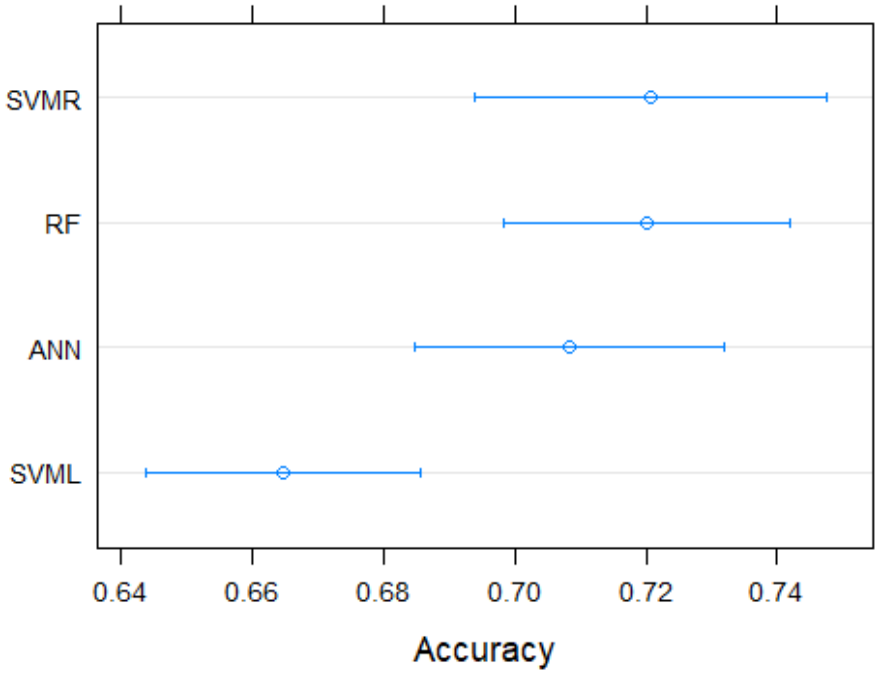
5. Modelos de clasificación

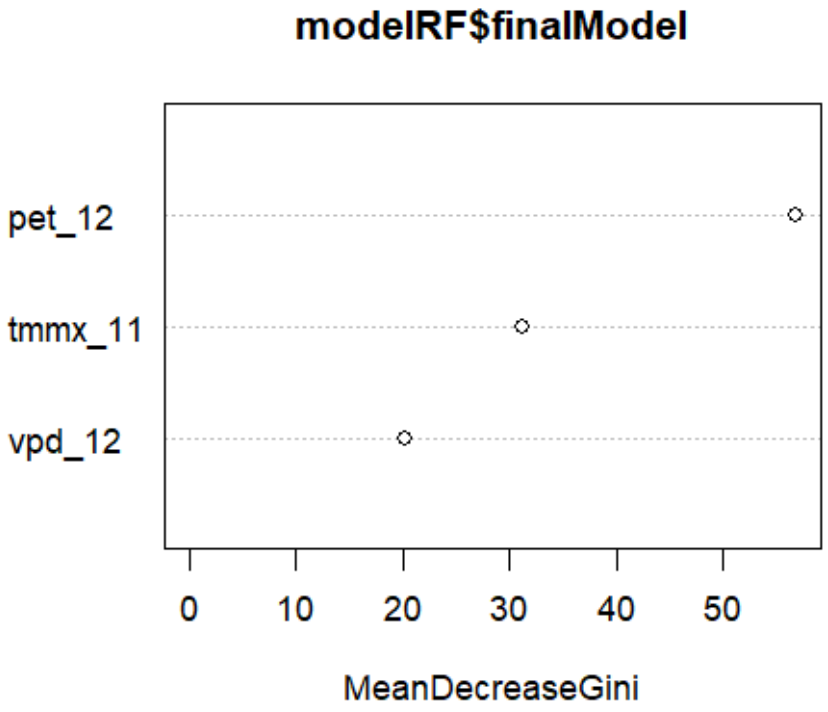
RENDIMIENTO FINAL

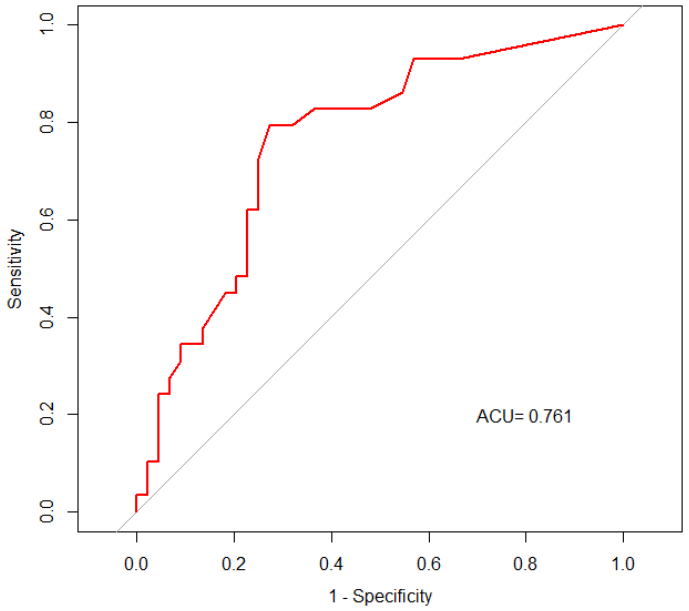
```
        OOB estimate of  error rate: 29.83%
Confusion matrix:
    1  2 3 4 class.error
1 129 23 0 0   0.1513158
2  22 34 6 2   0.4687500
3   4  9 4 0   0.7647059
4   1  3 1 0   1.0000000
```

- Rendimiento alto, medio, bajo y muy bajo







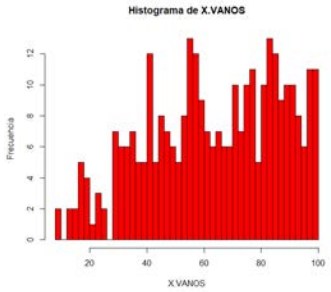


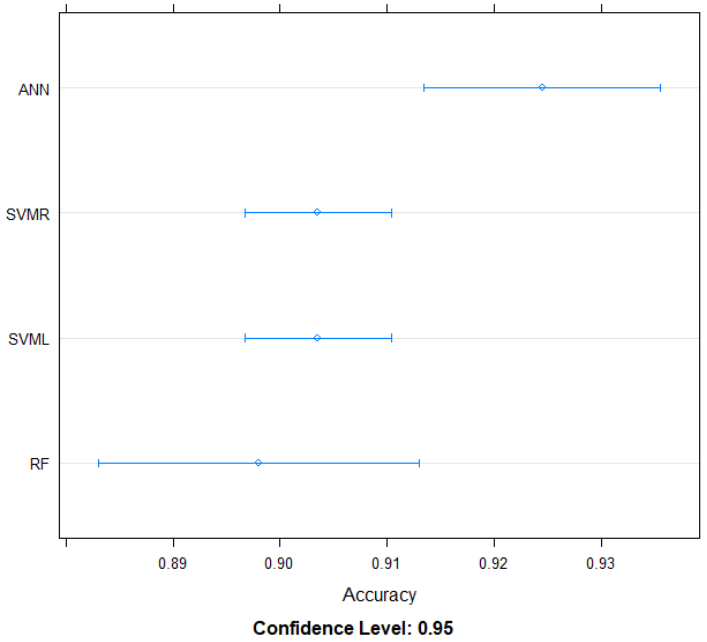
5. Modelos de clasificación

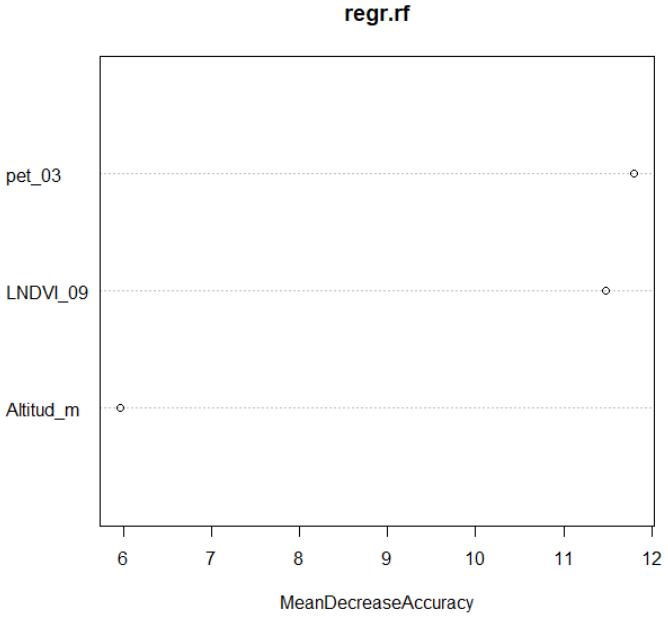
- **% VANOS**

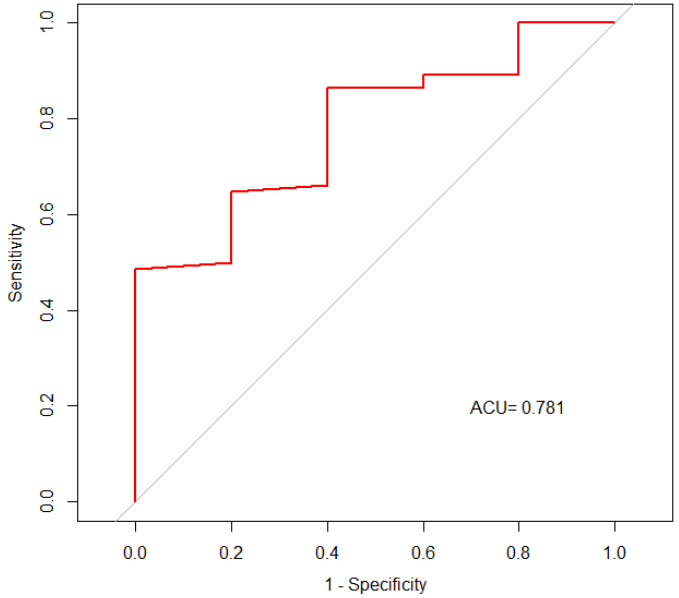
- % VANOS:
  - Alto>30%
  - Bajo <30%

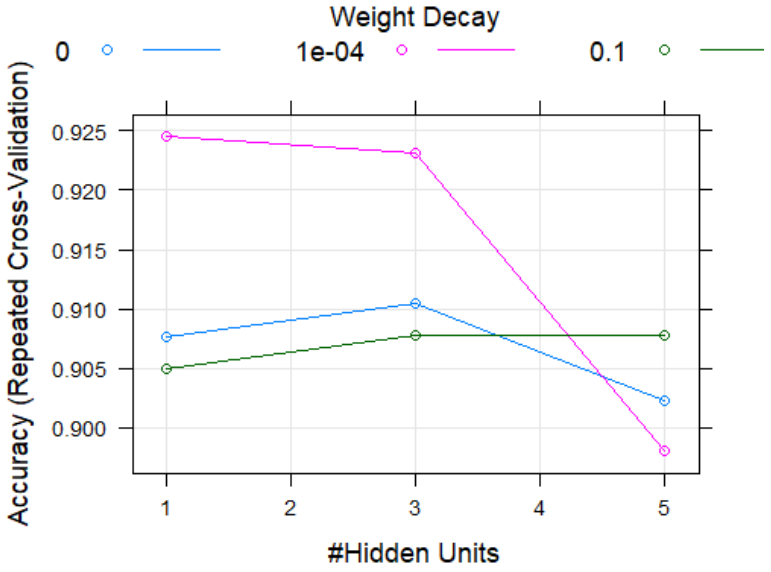
```
Call:
 randomForest(formula = X.VANOS ~ ., data = df.train, importance = T)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 1

        OOB estimate of  error rate: 9.24%
Confusion matrix:
  1   2 class.error
1 5  18  0.78260870
2 4 211  0.01860465
```











6. Conclusiones.

- **Habitualmente aparecen variables climáticas como variables mas importantes**
- **Las mas importantes son:**
  - Radiación solar meses de marzo y septiembre
  - Evapotranspiración potencial

↓

  - **Variables climáticas estimadas a partir de satélite**

↓

  - **monitorización mas local → Tree Talkers**

- **Parece que las condiciones climáticas de marzo y septiembre son las críticas**

↓

- **Mapas de afección en función de estas variables restrictivas**

Muchas gracias

+INFORMACIÓN:
angela.blazquez@cesefor.com

cesefor

CORAZÓN FORESTAL, espíritu investigador

PINEA
gopinea.org