



Universidad de Oviedo

**UNIVERSIDAD DE OVIEDO**

**MODELO DE PREDICCIÓN DE VOLÚMENES DE  
PRODUCCIÓN Y DISTRIBUCIÓN DE TAMAÑO DE FRUTOS  
A COSECHA EN MANDARINO 'TANGOR MURCOTT'**

**GRISELDA RITA ROMINA BÓBEDA**

**PROGRAMA DE DOCTORADO EN INFORMÁTICA**



## RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

1.- Título de la Tesis	
Español/Otro Idioma: Modelo de predicción de volúmenes de producción y distribución de tamaño de frutos a cosecha en mandarina 'tangor Murcott'	Inglés: Prediction model of production volumes and fruit size distribution at harvest in 'tangor Murcott' mandarin trees.
2.- Autor	
Nombre: <b>Griselda Rita Romina Bóbeda</b>	
Programa de Doctorado: Informática	
Órgano responsable: Centro Internacional de Postgrado	

### RESUMEN (en español)

Los objetivos de esta tesis fueron definir modelos de estimación para anticipar los volúmenes de producción y la distribución de tamaño de frutos a cosecha en mandarina tangor Murcott, en la provincia de Corrientes, Argentina. Complementariamente también se definieron los modelos para naranjo dulce Valencia late y Salustiana.

Se trabajó con el modelo propuesto: *Producción (tn)=numero árboles x peso medio de frutos x carga*, para el cual se determinaron sus componentes a partir de la información obtenida en diferentes temporadas en huertos comerciales de tangor Murcott (*Citrus reticulata* x *C. sinensis*) y naranjo dulce (*C. sinensis*). Si bien la propuesta original del plan de tesis se centraba en el pronóstico de producción de tangor Murcott, se pudieron obtener datos de campo e imágenes satelitales para huertos de naranjo dulce de las variedades Valencia late y Salustiana. Por ello, se incluye como objetivo específico la definición de la metodología para estimación de carga en cítricos, que contempla además de tangor Murcott, a estas dos variedades de naranjo dulce.

El *peso medio de frutos* (y) quedó definido por las curvas de crecimiento obtenidas en función del diámetro ecuatorial correspondiente a los días después de plena floración (x). La variabilidad de huerto y temporada, fueron considerados mediante modelos no lineales mixtos y modelos sigmoidales. Las curvas de crecimiento de frutos se definieron por la quinta parametrización del modelo Logístico  $y=10,0140+e^{(-2,5866)*0,9831x}$ . Con la curva de crecimiento para cada tamaño comercial se construyeron tablas de crecimiento. La relación entre el diámetro ecuatorial (y) y el peso de los frutos (z) se estudió mediante regresión no lineal, y se ajustó al modelo  $z=0,00186*y^{2,66435}$ , que permitió convertir los diámetros medidos a pesos estimados.

La estimación de carga de frutos se realizó sobre una muestra del 2 a 3% de árboles por huerto en tangor Murcott y naranjo dulce Valencia late y Salustiana. En los árboles seleccionados se midió altura, diámetro mayor, menor y radio de copa, número de frutos por



árbol en abril y junio mediante recuento total, corregido por factor de fruto oculto, y mediante recuentos en secciones de copa y extrapolado al total considerando xvii volúmenes de cuerpos geométricos. Al momento de la cosecha se registró la carga por árbol, tomada como valor de comparación.

Para determinar la variabilidad de las cargas estimada y real a cosecha se utilizó el modelo  $y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \varepsilon_{ijk}$ , el mismo se estudió con efectos aleatorios de temporada, huerto y árbol.

La precisión de cada método se evaluó calculando el error porcentual (% *Error*), los momentos de estimación se compararon con la prueba de t-Student para muestras pareadas y la relación entre la carga estimada y la real se estableció por coeficientes de correlación de Pearson. Se logró estimar la carga mediante recuento total de frutos en junio, corregido por coeficiente de fruto oculto, con un error de 9,85%.

Buscando procedimientos accesibles y de bajo costo para estimar la carga se combinaron diferentes conjuntos de variables y técnicas de aprendizaje automático.

Se comparó la precisión obtenida solamente con datos de campo mediante árboles de regresión, usando el algoritmo M5-Prime y con el agregado de información climática y de sensores remotos utilizando árboles de regresión, máquinas de vectores soporte, redes neuronales artificiales, métodos vagos, bosques aleatorios y para la selección de los caracteres relevantes se utilizó el método Borda. Cada conjunto de datos se dividió en subconjuntos de entrenamiento y prueba (75% y 25%), se repitió 50 veces el proceso de entrenamiento con el modelo de validación cruzada con 10 cajas. La precisión se evaluó calculando el error cuadrático medio (RMSE), el error absoluto medio (MAE) y el error porcentual (% *Error*).

Se logró estimar la carga por árbol en tanger Murcott, y naranjas dulces tardías Salustiana y Valencia con un 3,99% de error, utilizando una máquina de vectores soporte basado en la especie, la edad de los árboles, las condiciones de riego, la precipitación en el periodo de maduración de los frutos, la humedad en el periodo de crecimiento de los frutos, la reflectancia roja e infrarroja cercana en febrero, y el índice normalizado de vegetación, la reflectancia infrarroja cercana y roja en diciembre.

Finalmente se estimaron los *volúmenes de producción* mediante informantes calificados con errores porcentuales entre 6,22 y 9,57% y con los métodos objetivos aplicando el modelo propuesto con las componentes estimadas previamente, se obtuvieron errores medios de 5,96% cuando se efectuaron estimaciones por huerto basadas en información parcelaria, y de 7,58% cuando se realizaron pronósticos globales para todos los huertos empleando información parcelaria, meteorológica y satelital.

Se estimó la distribución de calibres comerciales esperada a la cosecha en base a las tablas de crecimiento definidas. La correspondencia entre la distribución estimada y real se evaluó mediante una matriz de confusión multiclase, y se obtuvo una precisión global de 0,784 y un  $k = 0,721$ , lo que indica una buena estimación del modelo.



## RESUMEN (en Inglés)

The objectives of this thesis were to define an estimation model to anticipate the production volumes and the fruit size distribution at harvest for tangor Murcott, in Corrientes province, Argentina. Furthermore, the models for Valencia late and Salustiana sweet oranges were also defined.

The proposed model was:  $Production(tn) = \text{number of trees} \times \text{average weight of fruits} \times \text{load}$ , its components were determined by the information obtained from different seasons in commercial orchards of tangor Murcott (*Citrus reticulata* × *C. sinensis*) and sweet orange (*C. sinensis*). Although the original proposal in the thesis plan was focused on prediction of production for tangor Murcott, field data and satellite images were obtained from sweet orange orchards of Valencia late and Salustiana varieties. Because of it, the definition of a methodology for estimating load in citrus is included as a specific objective, embracing, in addition to tangor Murcott, to these two varieties of sweet orange.

The average weight of fruits ( $y$ ) was defined by growth curves obtained as a function from the equatorial diameter corresponding to days after the full flowering ( $x$ ). The variability of orchard and season variability were considered through mixed nonlinear models and sigmoidal models. The fruit growth curves were defined by the fifth parameterization of Logistic model  $y = 10,0140 + e^{(-2,5866) \cdot 0,9831x}$ . Growth tables were developed from the growth curve of each commercial size. The relation between equatorial diameter ( $y$ ) and weight of fruits ( $z$ ) was studied by a nonlinear regression, and it fit within the potential model  $z = 0,00186 \cdot y^{2,66435}$ , this, allowed to convert the measured diameters into estimated weights.

The fruit load estimation was done from a sample of 2 to 3% trees for each orchard in tangor Murcott and sweet orange Valencia late and Salustiana. In the selected trees, height, larger and smaller diameter of crown and crown radius, number of fruits per tree, in April and June were measured by total count, it was corrected with the hidden fruit factor, and by counts in cup sections, and it was extrapolated for the total considering volumes of geometric bodies. During harvesting, the load per tree was recorded, which was taken as a comparison value. To determine the variability in estimated and real loads in the harvest, the final model was used  $y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \epsilon_{ijk}$  it was studied with random effects for season, orchard and tree. The precision of each method was assessed by calculating the percentage error (% Error), the estimation moments were compared with the t-Student test for paired samples and the relationship between the estimated and the real load was established by the correlation coefficient of Pearson. The load has been estimated by total fruit count in June, and it was corrected by the coefficient of hidden fruit, with an error of 9.85%.

Looking for accessible and low-cost procedures to estimate the load, different sets of variables and machine learning techniques were combined.

The accuracy obtained was compared only with field data using regression trees, using the M5-Prime algorithm, and with the addition of climate and remote sensing information using regression trees, support vector machines, artificial neural networks, vague methods, random forests and for the selection of the relevant characters the Borda method was used. Each dataset was divided into training and testing subsets (75% and 25%), and the training process was repeated 50 times with the 10-box cross-validation model. Accuracy was evaluated by calculating Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and percentage error (% Error).

It was possible to estimate the load per tree in tangor Murcott, and Salustiana and Valencia late sweet oranges with an error of 3.99%, using a support vector machine based on the species, the age of the trees, the irrigation conditions, the precipitation in the ripening period of the fruit, humidity in the fruit growth period, red and near-infrared reflectance in February, and normalized vegetation index, near-infrared and red reflectance in December.

Finally, the production volumes were estimated by qualified informants with percentage errors between 6.22 and 9.57% and with the objective methods applying the proposed model with its components previously estimated, they were obtained average errors of 5.96% when



Universidad de Oviedo  
*Universidá d'Uviéu*  
University of Oviedo

species, the age of the trees, the irrigation conditions, the precipitation in the ripening period of the fruit, humidity in the fruit growth period, red and near-infrared reflectance in February, and normalized vegetation index, near-infrared and red reflectance in December. Finally, the production volumes were estimated by qualified informants with percentage errors between 6.22 and 9.57% and with the objective methods applying the proposed model with its components previously estimated, they were obtained average errors of 5.96% when estimates were made per orchard based on parcel information, and 7.58% when global forecasts were made for all orchards using orchards meteorological and satellite information. The distribution of commercial sizes expected at harvest was estimated in based on the defined growth tables. Correspondence between both estimated and real distribution was evaluated using a multiclass confusion matrix, it was obtained an overall precision of 0.784 and a  $k = 0.721$ , which indicates a good estimation of the model.

**SR. PRESIDENTE DE LA COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO  
EN \_\_\_\_\_**



Universidad de Oviedo  
*Universidá d'Uviéu*  
University of Oviedo

## Justificación

<lugar> <fecha>

Presidente de la Comisión Académica del Programa de Doctorado

Fdo.: .

Contra la presente Resolución, podrá interponer recurso de alzada ante el Rectorado, en el plazo de un mes, a partir del día siguiente al de la presente notificación, de conformidad con el art. 122 de la Ley 39/2015, de 1 de octubre, de Procedimiento Administrativo Común de las Administraciones Públicas

SR.

---

(Doctorando, Director de la Tesis Doctoral y Sr. Director del Centro Internacional de Postgrado)



Universidad de Oviedo

**UNIVERSIDAD DE OVIEDO**

**MODELO DE PREDICCIÓN DE VOLÚMENES DE  
PRODUCCIÓN Y DISTRIBUCIÓN DE TAMAÑO DE FRUTOS  
A COSECHA EN MANDARINO 'TANGOR MURCOTT'**

**GRISELDA RITA ROMINA BÓBEDA**

**PROGRAMA DE DOCTORADO EN INFORMÁTICA**

## **TRABAJO PUBLICADO**

- Bóbeda G. R. R.; Combarro E. F.; Mazza S. M.; Giménez L. I.; Díaz I. 2018. Estimation of citrus load with machine learning techniques and minimum number of features. *International Journal of Computational Intelligence Systems*. Atlantis Press. (12): 79-89.

## **TRABAJOS ACEPTADOS PARA PUBLICAR**

- Bóbeda, G. R. R. y Gaiad, J. E. Eds. 2021. Manual de Producción Citrícola. Instrumentos para una actividad sustentable. Convocatoria Publicar 2021. EUDENE. Bóbeda, G. R. R.; Mazza, S. M. Sección II-Pronóstico de producción: Capítulos III, IV, V. Aceptado en prensa.
- Bóbeda, G.R.R.; Mazza, S.M.; Rico, N.; Brénes Pérez; C.F.; Gaiad, J.E.; Díaz; S.I. On the identification of factors affecting citrus harvest estimation. Revista de la Facultad de Ciencias Agrarias de la UNCuyo. Aceptado con correcciones (07-04-2022).



# INDICE GENERAL

## CAPÍTULO I

<b>PLANTEAMIENTO DEL PROBLEMA, ANTECEDENTES.....</b>	<b>1</b>
1.1. ORIGEN DE LOS CÍTRICOS .....	1
1.2. PRODUCCIÓN DE CÍTRICOS.....	2
1.2.1. En el mundo .....	2
1.2.2. En Argentina .....	3
1.3. PLANTEAMIENTO Y DESCRIPCIÓN DEL PROBLEMA .....	5
1.3.1. Importancia de los pronósticos anticipados de producción .....	5
1.3.2. Métodos para la estimación de producción .....	7
1.4. OBJETIVOS .....	8
1.4.1. General.....	9
1.4.2. Particulares .....	9
1.5. ANTECEDENTES .....	10
1.5.1. Número de árboles .....	10
1.5.2. Peso de fruto medio.....	11
1.5.3. Carga .....	14
1.5.3.1. Informantes calificados .....	14
1.5.3.2. Información a nivel parcelario.....	16
1.5.3.3. Información de sensores remotos.....	19
1.6. HIPÓTESIS .....	21
1.7. BIBLIOGRAFÍA .....	22

## CAPÍTULO II

<b>DESCRIPCIÓN DEL ÁREA Y MATERIAL DE ESTUDIO .....</b>	<b>27</b>
2.1. ÁREA DE ESTUDIO.....	27
2.1.1. Provincia de Corrientes.....	28
2.1.2. Provincia de Entre Ríos .....	29
2.2. SUELO Y CLIMA .....	29
2.2.1. Provincia de Corrientes.....	29
2.2.2. Provincia de Entre Ríos .....	31
2.3. MATERIAL DE ESTUDIO.....	32
2.3.1. Mandarinas.....	32
2.3.2. Naranjas.....	33
2.3.2.1. Variedad Valencia late.....	33
2.3.2.2. Variedad Salustiana.....	34
2.3.3. Los huertos cítricos .....	34
2.3.3.1. Caracterización general de los huertos .....	34
2.4.2. Huertos bajo estudio.....	35
2.5. BIBLIOGRAFÍA .....	39

## CAPÍTULO III

<b>MODELIZACIÓN DEL CRECIMIENTO DE LOS FRUTOS .....</b>	<b>40</b>
3.1. INTRODUCCIÓN .....	40
3.1.2. Crecimiento y desarrollo de los frutos .....	40
3.1.3. Maduración de los frutos.....	42
3.1.4. Estimación del tamaño de los frutos a la cosecha .....	44
3.1.4.1. Regresión no lineal .....	45

3.1.4.1.1. Modelos de crecimiento sigmoidal.....	47
3.1.4.1.2. Estimación de parámetros en modelos de regresión no-lineal .....	52
3.1.4.1.3. Obtención de valores iniciales de los parámetros .....	53
3.1.4.1.4. Criterios de selección de modelos .....	56
3.1.4.2. Modelos no lineales mixtos.....	58
3.3. MATERIALES Y MÉTODOS .....	63
3.3.1. Datos Experimentales .....	63
3.3.1.1. Variables estudiadas .....	65
3.3.1.2. Factores.....	66
3.3.2. Análisis de datos .....	67
3.3.2.1. Modelización del crecimiento de frutos de tangor Murcott .....	67
3.3.2.2. Análisis de modelos no lineales mixtos .....	69
3.4. RESULTADOS Y DISCUSIÓN .....	69
3.4.1. Selección de modelos sigmoidales.....	70
3.4.2. Modelización del crecimiento de frutos con modelos no lineales mixtos ....	77
3.5. RESUMEN DE RESULTADOS .....	83
3.6. BIBLIOGRAFÍA .....	84

## **CAPÍTULO IV**

<b>TABLAS DE CRECIMIENTO DE FRUTOS .....</b>	<b>90</b>
4.1. INTRODUCCIÓN .....	90
4.1.1. Raleo de frutos.....	94
4.1.2. Tablas de crecimiento .....	97
4.1.3. Estimación del peso de los frutos .....	99
4.2. MATERIALES Y MÉTODOS .....	100

4.2.1. Datos experimentales .....	100
4.2.2. Análisis Estadístico .....	102
4.3. RESULTADOS Y DISCUSIÓN .....	104
4.3.1. Tablas de Crecimiento.....	104
4.3.2. Relación diámetro-peso .....	110
4.3.3. Uso de tablas de crecimiento para raleo de frutos.....	115
4.4. RESUMEN DE RESULTADOS .....	117
4.5. BIBLIOGRAFÍA .....	119

## CAPÍTULO V

<b>ESTIMACIÓN DE CARGA BASADA EN INFORMACIÓN PARCELARIA.....</b>	<b>122</b>
5.1. INTRODUCCIÓN .....	122
5.1.1. Mediciones parcelarias .....	123
5.1.1.1. Recuento total.....	124
5.1.1.2. Recuento en secciones predefinidas de copa .....	125
5.1.1.3. Otras mediciones.....	128
5.1.2. Momento de realizar las mediciones .....	129
5.1.3. Estudio de la variabilidad de la carga.....	129
5.2. MATERIALES Y MÉTODOS .....	132
5.2.1. Datos Experimentales .....	132
5.2.1.1. Dimensiones de los árboles.....	135
5.2.1.2. Estimaciones de carga .....	136
5.2.1.3. Valor de comparación .....	138
5.2.2. Análisis de datos .....	139
5.2.2.1 Error de estimación (de acá en adelante corregido).....	139

5.2.2.2. Variabilidad de la carga estimada y la cosecha real .....	139
5.3. RESULTADOS Y DISCUSIÓN .....	140
5.3.1. Porcentajes de error .....	140
5.3.2. Variabilidad de la carga estimada y real a cosecha.....	144
5.5. RESUMEN DE RESULTADOS .....	147
5.6. BIBLIOGRAFÍA .....	148

## **CAPÍTULO VI**

<b>ESTIMACIÓN DE CARGA BASADA EN INFORMACIÓN PARCELARIA, CLIMÁTICA Y SATELITAL .....</b>	<b>152</b>
6.1. INTRODUCCIÓN .....	152
6.1.1. Estimación con información de campo.....	154
6.1.2. Uso de sensores remotos.....	155
6.1.2.1. Sensores remotos.....	155
6.1.2.1. Índices de vegetación .....	157
6.1.3. Aprendizaje automático.....	159
6.1.3.1. Métodos basados en arboles de decisión .....	160
6.1.3.1.1. Árboles de Regresión .....	161
6.1.3.1.2. Arboles de inferencia condicional (CTREE) .....	164
6.1.3.1.3. Métodos de ensamble: Bosques Aleatorios .....	165
6.1.3.2. Métodos vagos, K-vecinos más cercanos .....	166
6.1.3.3. Redes Neuronales Artificiales.....	166
6.1.3.4. Máquinas de Vectores Soporte .....	168
6.1.4. Selección de características .....	172
6.1.5. Antecedentes.....	174

6.2. MATERIALES Y MÉTODOS .....	175
6.2.1. Datos Experimentales .....	176
6.2.1.1. Huertos.....	176
6.2.1.2. Variables.....	178
6.2.2. Análisis de datos .....	185
6.2.2.1. Estimación de carga basada en datos de campo (características de los huertos e información de los árboles) .....	187
6.2.2.2. Estimación de carga basada en características de los huertos, información de los árboles, variables climáticas e información satelital .....	188
6.3. RESULTADOS Y DISCUSIÓN .....	190
6.3.1. Estimación de carga basada en datos de campo (características de los huertos e información de los árboles) .....	190
6.3.2. Estimación de carga basada en características de los huertos, información de los árboles, variables climáticas e información satelital .....	194
6.4. RESUMEN DE RESULTADOS .....	203
6.5. BIBLIOGRAFÍA .....	206

## **CAPÍTULO VII**

### **ESTIMACIÓN ANTICIPADA DE VOLUMENES DE PRODUCCIÓN Y DISTRIBUCIÓN DE TAMAÑOS COMERCIALES A NIVEL PARCELARIO .....**

7.1. INTRODUCCIÓN.....	213
7.1.1. Estimación anticipada de los volúmenes de producción .....	214
7.1.1.1. Pronóstico mediante métodos subjetivos.....	215
7.1.1.2. Pronóstico mediante métodos objetivos .....	215
7.1.2. Calibres comerciales y su distribución .....	216
7.2. MATERIALES Y MÉTODOS.....	217

7.2.1. Estimación subjetiva de volúmenes de producción.....	217
7.2.2. Estimación objetiva de volúmenes de producción .....	218
7.2.2.1. Datos .....	218
7.2.2.2. Análisis de los datos .....	218
7.2.3. Distribución de calibres.....	220
7.3. RESULTADOS Y DISCUSIÓN .....	223
7.3.1. Estimación anticipada de volúmenes de producción.....	223
7.3.2. Distribución de calibres comerciales .....	226
7.4. RESUMEN DE RESULTADOS .....	228
7.5. BIBLIOGRAFIA .....	230

## **CAPÍTULO VIII**

<b>CONCLUIONES Y RECOMENDACIONES.....</b>	<b>232</b>
8.1. CONCLUSIONES .....	232
8.1.1. Determinación de componentes del modelo .....	233
8.1.1.1. Número de árboles .....	233
8.1.1.2. Peso medio de los frutos.....	233
8.1.1.3. Número de frutos por árbol .....	235
A nivel parcelario .....	235
A nivel de cuenca o región.....	235
8.1.2. Precisión del modelo de pronóstico .....	236
8.1.3. Precisión de las tablas de crecimiento .....	237
8.2. VERIFICACIÓN DE HIPÓTESIS .....	237
8.3. RECOMENDACIONES .....	238

## INDICE DE TABLAS

Tabla 2.1. Temporadas en las que se relevó información y características de los huertos de tangor Murcott ( <i>C. sinensis</i> x <i>C. reticulata</i> ) estudiados.....	36
Tabla 2.2. Temporadas en las que se relevó información y características de los huertos naranjo dulce ( <i>C. sinensis</i> ) estudiados .....	37
Tabla 3.1. Modelos de tipo sigmoideo utilizados para describir crecimiento de frutos	49
Tabla 3.2. Reparametrizaciones de los modelos Logístico y Gompertz .....	50
Tabla 3.3. Técnica de análisis (modelos no lineales de crecimiento sigmoideo, MNLCs; modelos no lineales mixtos, MNLM), huertos, temporadas, fechas de plena flor, de inicio de mediciones del diámetro ecuatorial de los frutos, expresado días después de plena flor .....	64
Tabla 3.5. Criterios de selección de modelos (raíz cuadrada del cuadrado medio del error, $CM\epsilon$ , y no linealidad intrínseca, NI) para los diferentes modelos estudiados.....	71
Tabla 3.6. Efecto parámetro para cada fruto (por combinación temporada-huerto) en las reparametrizaciones de los modelos Logístico (L) y Gompertz (G).....	73
Tabla 3.7. Estimación de los parámetros $\alpha$ , $\beta$ y $\gamma$ del modelo logístico en su 5 <sup>ta</sup> reparametrización por temporada y huerto .....	75
Tabla 3.8. Modelos de efecto fijo y efectos aleatorios propuestos para evaluar el patrón crecimiento de tangor Murcott.....	78
Tabla 3.9. Valores de AIC de los modelos con efectos fijos y aleatorios asociados a los parámetros $\alpha$ , $\beta$ y $\gamma$ temporada y huerto .....	79
Tabla 3.10. Parámetros estimados para el modelo no lineal mixto, por temporada y huerto. ....	80
Tabla 3.11. Estimación de componentes de varianza (Var%) y desvío estándar (SD) de los efectos aleatorios: Huerto y Temporada.....	82
Tabla 4.1. Modelo de tabla de crecimiento .....	98



Tabla 4.2. Identificación de huertos, temporadas, fechas de plena floración, de cosecha y días desde plena floración a cosecha .....	101
Tabla 4.3. Rangos de diámetro ecuatorial (mm) mínimo, máximo y medio, para cada tamaño comercial. ....	102
Tabla 4.4. Estimaciones de los parámetros del modelo logístico L5, para cada tamaño comercial según Unión Europea. $\alpha$ , $\beta$ y $\gamma$ parámetros del modelo, asíntota superior, ordenada al origen (intercepto), tasa de crecimiento y número de frutos por rango de calibre.....	105
Tabla 4.5. Tabla de crecimiento de frutos de tangor Murcott por calibre comercial de acuerdo con las normas de comercialización de la Unión Europea. Incremento del diámetro ecuatorial (mm) en función de los días después de plena floración (DDPF), TC: tamaño comercial según Unión Europea .....	108
Tabla 4.6. Estimación de los parámetros del modelo potencial, error estándar y p-valor. ....	111
Tabla 5.1. Huertos seleccionados para la estimación .....	133
Tabla 5.1. Continuación.....	134
Tabla 5.2. Fórmulas para el cálculo de volumen de copa correspondientes a cada método del marco contador (VolM1, VolM2, VolM3, VolM4, VolM5, VolM6) y sus respectivas referencias .....	138
Tabla 5.3. Coeficientes de correlación de Pearson ( $r$ ) entre los diferentes métodos de estimación de carga (medición en secciones predefinidas de copa: VolM1, VolM2, VolM3, VolM4, VolM5, VolM6; recuento total: RT) y la carga real por especie y momento de estimación (ME1: abril y ME2: junio) .....	140
Tabla 5.4. Porcentajes de error entre la carga estimada y la real para los diferentes métodos de estimación de carga (medición en secciones predefinidas de copa: VolM1, VolM2, VolM3, VolM4, VolM5, VolM6; recuento total: RT) por especie y momento de estimación (1: Abril y 2: Junio) .....	141

Tabla 5.5. Prueba T pareada de los recuentos de frutos por árbol entre momentos de estimación (ME1 y ME2), $X_{dif}$ : promedio de las diferencias, $S_{xdif}$ : desvío estándar de la diferencia, p-valor, obtenidos mediante el método recuento total .....	142
Tabla 5.6. Carga estimada, real y corregida por coeficiente de fruto oculto para tangor Murcott. ....	143
Tabla 5.7. Componentes de varianza para las estimaciones de carga en tangor Murcott .....	144
Tabla 5.8. Componentes de varianza para las estimaciones de carga en naranja dulce .....	145
Tabla 5.9. Componentes de varianza de la carga real a cosecha para tangor Murcott	145
Tabla 5.10. Componentes de varianza estimada de la carga real a cosecha para naranja dulce.....	146
Tabla 6.1. Huertos seleccionados para la instancia 1.....	176
Tabla 6.2. Huertos seleccionados para la instancia 2.....	177
Tabla 6.3. Descripción de las variables que integran cada conjunto de datos.....	179
Tabla 6.4. Descripción de las variables que integran cada conjunto de datos.....	185
Tabla 6.5. Frecuencias de las variables categóricas (Vd, variedad; Suelo, arenoso-arcilloso; Irr, presencia o ausencia de riego).....	185
Tabla 6.6. Parámetros introducidos en cada método de Aprendizaje Automático empleado (RT, árboles de regresión; ANN, redes neuronales artificiales; KNN, k-vecinos más cercanos; SVM, máquinas de soporte vectoriales; RF, bosques aleatorios).....	188
Tabla 6.7. Coeficientes de correlación de Pearson (r) entre diferentes variables y la carga a cosecha, y valores p asociados, en naranja dulce y tangor Murcott.....	190
Tabla. 6.8. Características de cada método empleado para predicción de carga mediante M5-PRIME en naranja dulce (conjuntos de datos utilizado, complejidad de los modelos, errores en las predicciones y correlación entre carga estimada y real).....	193
Tabla. 6.9. Características de cada método empleado para predicción de carga mediante M5-PRIME en tangor Murcott (conjuntos de datos utilizado, complejidad de los modelos, errores en las predicciones y correlación entre carga estimada y real) .....	193

Tabla 6.10. Valores p de la prueba t pareada para las diferencias medias de MAE y RMSE entre svm1 y otros algoritmos para cada conjunto de datos.....	197
Tabla 6.11. Variables seleccionadas del nuevo conjunto de datos d2 filtrado. ....	200
Tabla 7.1. Huertos utilizados para la estimación de volúmenes de producción y detalle de la información obtenida .....	218
Tabla 7.2. Rangos de diámetro ecuatorial (mm) mínimo, máximo y promedio, para cada tamaño comercial de frutos.....	220
Tabla 7.3 Matriz de confusión.....	222
Tabla 7.4 Valores Índice Kappa y su clasificación.....	223
Tabla 7.5. Estimación de volúmenes de producción (tn) en huertos de tangor Murcott .....	223
Tabla 7.6. Porcentajes de error de estimación de volúmenes de producción (tn) de tangor Murcott.....	224
Tabla 7.7. Predicción de los volúmenes de producción de tangor Murcott con información parcelaria, meteorológica y satelital.....	225
Tabla 7.8. Matriz de confusión de las frecuencias predichas y observados a campo para los diferentes calibres comerciales .....	227

## INDICE DE FIGURAS

Figura 1.1. Países productores de cítricos.....	2
Figura 1.2. Regiones Citrícolas de la Argentina .....	4
Figura 1.3. Modelo de pronóstico de producción .....	8
Figura 2.1. Localización geográfica de los huertos comerciales estudiados en las provincias de Corrientes y Entre Ríos, Argentina.....	27
Figura 2.2. Localización de los huertos estudiados en la provincia de Corrientes(a) y en la provincia de Entre Ríos (b) .....	38
Figura 3.1. Fases de crecimiento de frutos cítricos (Avanza, 2011) .....	48
Figura 3.2. Croquis de los huertos en estudio, selección de plantas.....	65
Figura 3.3. (a) Registro de floración con el objeto de determinar la fecha de plena flor (80% flores abiertas), (b) Identificación de árboles para registro del crecimiento de los frutos seleccionados (c) Identificación de frutos y registro del diámetro con calibre digital. ....	66
Figura 3.4. Curvas promedio de crecimiento de frutos tangor Murcott (diámetro ecuatorial vs días después de la plena floración DDPF) ajuste del modelo logístico en su 5ta reparametrización, en huertos y temporadas .....	71
Figura 3.5. Interacción de los factores temporada y huerto sobre los parámetros del modelo Logístico L5: a. Asíntota superior ( $1/\alpha$ ), b. Intercepción $1 / (\alpha + e^\beta)$ y c. Tasa de crecimiento ( $-\ln\gamma$ ).....	76
Figura 4.1. Curvas medias de crecimiento de frutos tangor Murcott por calibre comercial, según normas de comercialización de la Union Europea.....	106
Figura 4.2. Porcentajes de fruta mal clasificada en su tamaño comercial de tangor Murcott en función de los días después de floración (DDPF) .....	110
Figura 4.3. Valores observados y predichos del peso (g) en función del diámetro ecuatorial (mm) en frutos de tangor Murcott.....	111
Figura 4.4. Ajuste a la distribución normal mediante Q-Qplot .....	112

Figura 4.5. Gráficos de residuales en función a predichos del modelo potencial seleccionado para la relación peso (g)- diámetro (mm).....	112
Figura 4.6. Relación diámetro-peso, tamaños comerciales, número de frutos por cajas de 10 y 15 kg para tangor Murcott .....	114
Figura 5.1. Plantas seleccionadas para el recuento de frutos.....	135
(a) Huerto 1.5, (b) Huerto 1.6, (c) Huerto: 3.1 .....	135
Figura 5.2. Determinación de la carga mediante el recuento total .....	136
Figura 5.3. Uso del marco de muestreo .....	137
Figura 6.1. Ejemplo de un Árbol de decisión. Archivo <i>Iris</i> . (Balzarini et al., 2008).....	161
Figura 6.2. Representación gráfica de una red neuronal .....	167
Figura 6.3. Estructura de una red neuronal multicapa (Norasyikin et al., 2012) .....	167
Figura 6.4. Ejemplo de frontera de separación para un problema bidimensional (izquierda) o tridimensional (derecha).....	169
Figura 6.5. Diferentes separadores lineales para un conjunto de datos.....	169
Figura 6.6. Margen máximo de separación para un conjunto de datos.....	170
Figura 6.7. El truco de Kernel .....	171
Figura 6.8. Flujo de trabajo utilizado en el proceso .....	182
Figura 6.9. Extracción de los valores de NDVI en cada huerto.....	184
Figura 6.10. Errores obtenidos utilizando los diferentes conjuntos de variables para naranjas.....	191
Figura 6.11. Errores obtenidos utilizando los diferentes conjuntos de variables para tangor Murcott.....	192
Figura 6.12. Gráficos de caja que representan: <b>a</b> , MAE conjunto d1; <b>b</b> , RMSE conjunto d1; <b>c</b> , MAE conjunto d2; <b>d</b> , RMSE conjunto d2; <b>e</b> , MAE conjunto d3; <b>f</b> , RMSE conjunto d3 .....	196
Figura 6.13. Gráficos de caja que representan: <b>a</b> , valores MAE y <b>b</b> , RMSE obtenidos con svm1 para diferentes conjuntos de datos (d1, d2, d3). .....	198

Figura 6.14. Gráficos de caja que representan: <b>a</b> , MAE y <b>b</b> , RMSE, comparando smv1 con d2 y d2 filtrado .....	201
Figura 7.1. Modelo de pronóstico de producción .....	213
Figura 7.2. Distribución de calibres comerciales estimados y reales en tangor Murcott .....	226

## RESUMEN

Los objetivos de esta tesis fueron definir modelos de estimación para anticipar los volúmenes de producción y la distribución de tamaño de frutos a cosecha en mandarino tangor Murcott, en la provincia de Corrientes, Argentina. Complementariamente también se definieron los modelos para naranjo dulce Valencia late y Salustiana.

Se trabajó con el modelo propuesto: *Producción (tn) = número árboles x peso medio de frutos x carga*, para el cual se determinaron sus componentes a partir de la información obtenida en diferentes temporadas en huertos comerciales de tangor Murcott (*Citrus reticulata* × *C. sinensis*) y naranjo dulce (*C. sinensis*). Si bien la propuesta original del plan de tesis se centraba en el pronóstico de producción de tangor Murcott, se pudieron obtener datos de campo e imágenes satelitales para huertos de naranjo dulce de las variedades Valencia late y Salustiana. Por ello, se incluye como objetivo específico la definición de la metodología para estimación de carga en cítricos, que contempla además de tangor Murcott, a estas dos variedades de naranjo dulce.

El *peso medio de frutos* ( $y$ ) quedó definido por las curvas de crecimiento obtenidas en función del diámetro ecuatorial correspondiente a los días después de plena floración ( $x$ ). La variabilidad de huerto y temporada, fueron considerados mediante modelos no lineales mixtos y modelos sigmoidales. Las curvas de crecimiento de frutos se definieron por la quinta parametrización del modelo Logístico  $y = \frac{1}{0,0140 + e^{(-2,5866) * 0,9831^x}}$ . Con la curva de crecimiento para cada tamaño comercial se construyeron tablas de crecimiento. La relación entre el diámetro ecuatorial ( $y$ ) y el peso de los frutos ( $z$ ) se estudió mediante regresión no lineal, y se ajustó al modelo  $z = 0,00186 * y^{2,66435}$ , que permitió convertir los diámetros medidos a pesos estimados.

La estimación de carga de frutos se realizó sobre una muestra del 2 a 3% de árboles por huerto en tangor Murcott y naranjo dulce Valencia late y Salustiana. En los árboles seleccionados se midió altura, diámetro mayor, menor y radio de copa, número de frutos por árbol en abril y junio mediante recuento total, corregido por factor de fruto oculto, y mediante recuentos en secciones de copa y extrapolado al total considerando

volúmenes de cuerpos geométricos. Al momento de la cosecha se registró la carga por árbol, tomada como valor de comparación.

Para determinar la variabilidad de las cargas estimada y real a cosecha se utilizó el modelo  $y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \varepsilon_{ijk}$ , el mismo se estudió con efectos aleatorios de temporada, huerto y árbol.

La precisión de cada método se evaluó calculando el error porcentual (*% Error*), los momentos de estimación se compararon con la prueba de t-Student para muestras pareadas y la relación entre la carga estimada y la real se estableció por coeficientes de correlación de Pearson. Se logró estimar la carga mediante recuento total de frutos en junio, corregido por coeficiente de fruto oculto, con un error de 9,85%.

Buscando procedimientos accesibles y de bajo costo para estimar la carga se combinaron diferentes conjuntos de variables y técnicas de aprendizaje automático.

Se comparó la precisión obtenida solamente con datos de campo mediante árboles de regresión, usando el algoritmo M5-Prime y con el agregado de información climática y de sensores remotos utilizando árboles de regresión, máquinas de vectores soporte, redes neuronales artificiales, métodos vagos, bosques aleatorios y para la selección de los caracteres relevantes se utilizó el método Borda. Cada conjunto de datos se dividió en subconjuntos de entrenamiento y prueba (75% y 25%), se repitió 50 veces el proceso de entrenamiento con el modelo de validación cruzada con 10 cajas. La precisión se evaluó calculando el error cuadrático medio (RMSE), el error absoluto medio (MAE) y el error porcentual (*% Error*).

Se logró estimar la carga por árbol en tangor Murcott, y naranjas dulces tardías Salustiana y Valencia con un 3,99% de error, utilizando una máquina de vectores soporte basado en la especie, la edad de los árboles, las condiciones de riego, la precipitación en el periodo de maduración de los frutos, la humedad en el periodo de crecimiento de los frutos, la reflectancia roja e infrarroja cercana en febrero, y el índice normalizado de vegetación, la reflectancia infrarroja cercana y roja en diciembre.

Finalmente se estimaron los *volúmenes de producción* mediante informantes calificados con errores porcentuales entre 6,22 y 9,57% y con los métodos objetivos aplicando el modelo propuesto con las componentes estimadas previamente, se



obtuvieron errores medios de 5,96% cuando se efectuaron estimaciones por huerto basadas en información parcelaria, y de 7,58% cuando se realizaron pronósticos globales para todos los huertos empleando información parcelaria, meteorológica y satelital.

Se estimó la distribución de calibres comerciales esperada a la cosecha en base a las tablas de crecimiento definidas. La correspondencia entre la distribución estimada y real se evaluó mediante una matriz de confusión multiclase, y se obtuvo una precisión global de 0,784 y un  $\kappa = 0,721$ , lo que indica una buena estimación del modelo.

## ABSTRACT

The objectives of this thesis were to define an estimation model to anticipate the production volumes and the fruit size distribution at harvest for tangor Murcott, in Corrientes province, Argentina. Furthermore, the models for Valencia late and Salustiana sweet oranges were also defined.

The proposed model was:  $Production(tn) = \text{number of trees} \times \text{average weight of fruits} \times \text{load}$ , its components were determined by the information obtained from different seasons in commercial orchards of tangor Murcott (*Citrus reticulata* × *C. sinensis*) and sweet orange (*C. sinensis*). Although the original proposal in the thesis plan was focused on prediction of production for tangor Murcott, field data and satellite images were obtained from sweet orange orchards of Valencia late and Salustiana varieties. Because of it, the definition of a methodology for estimating load in citrus is included as a specific objective, embracing, in addition to tangor Murcott, to these two varieties of sweet orange.

The average weight of fruits ( $y$ ) was defined by growth curves obtained as a function from the equatorial diameter corresponding to days after the full flowering ( $x$ ). The variability of orchard and season variability were considered through mixed nonlinear models and sigmoidal models. The fruit growth curves were defined by the fifth parameterization of Logistic model  $y = \frac{1}{0,0140 + e^{(-2,5866) * 0,9831^x}}$ . Growth tables were developed from the growth curve of each commercial size. The relation between equatorial diameter ( $y$ ) and weight of fruits ( $z$ ) was studied by a nonlinear regression, and it fit within the potential model  $z = 0,00186 * y^{2,66435}$ , this, allowed to convert the measured diameters into estimated weights.

The fruit load estimation was done from a sample of 2 to 3% trees for each orchard in tangor Murcott and sweet orange Valencia late and Salustiana. In the selected trees, height, larger and smaller diameter of crown and crown radius, number of fruits per tree, in April and June were measured by total count, it was corrected with the hidden fruit factor, and by counts in cup sections, and it was extrapolated for the total

considering volumes of geometric bodies. During harvesting, the load per tree was recorded, which was taken as a comparison value.

To determine the variability in estimated and real loads in the harvest, the final model was used  $y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \varepsilon_{ijk}$  it was studied with random effects for season, orchard and tree.

The precision of each method was assessed by calculating the percentage error (% Error), the estimation moments were compared with the t-Student test for paired samples and the relationship between the estimated and the real load was established by the correlation coefficient of Pearson. The load has been estimated by total fruit count in June, and it was corrected by the coefficient of hidden fruit, with an error of 9.85%.

Looking for accessible and low-cost procedures to estimate the load, different sets of variables and machine learning techniques were combined.

The accuracy obtained was compared only with field data using regression trees, using the M5-Prime algorithm, and with the addition of climate and remote sensing information using regression trees, support vector machines, artificial neural networks, vague methods, random forests and for the selection of the relevant characters the Borda method was used. Each dataset was divided into training and testing subsets (75% and 25%), and the training process was repeated 50 times with the 10-box cross-validation model. Accuracy was evaluated by calculating Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and percentage error (% Error).

It was possible to estimate the load per tree in tangor Murcott, and Salustiana and Valencia late sweet oranges with an error of 3.99%, using a support vector machine based on the species, the age of the trees, the irrigation conditions, the precipitation in the ripening period of the fruit, humidity in the fruit growth period, red and near-infrared reflectance in February, and normalized vegetation index, near-infrared and red reflectance in December.

Finally, the production volumes were estimated by qualified informants with percentage errors between 6.22 and 9.57% and with the objective methods applying the proposed model with its components previously estimated, they were obtained average

errors of 5.96% when estimates were made per orchard based on parcel information, and 7.58% when global forecasts were made for all orchards using orchards meteorological and satellite information.

The distribution of commercial sizes expected at harvest was estimated in based on the defined growth tables. Correspondence between both estimated and real distribution was evaluated using a multiclass confusion matrix, it was obtained an overall precision of 0.784 and a  $\kappa = 0.721$ , which indicates a good estimation of the model.

# CAPITULO I

## PLANTEAMIENTO DEL PROBLEMA, ANTECEDENTES

### 1.1. ORIGEN DE LOS CÍTRICOS

Se denomina cítricos a diversas especies del género *Citrus*, originarias del sudeste asiático, en una vasta área ubicada entre 0° y 30° de latitud norte, que incluye sur y sudeste de China, India, Myanmar, Tailandia, Filipinas, Borneo y Sumatra, entre otros. Es allí donde se ha encontrado la mayor cantidad de especies cítricas. En esas latitudes, con altas temperaturas en las zonas bajas, pero templadas a templadas-frías en las colinas adyacentes a los Montes Himalaya, límite entre China e India, surgieron especies e híbridos naturales, muchos de ellos actualmente conocidos (Palacios, 2013).

La hipótesis más aceptada por la comunidad científica con respecto al origen filogenético de los cítricos es que la mayoría de las especies del género *Citrus* son probablemente híbridos directos o híbridos sucesivos de cuatro especies ancestrales: *C. medica* L. (cidro), *C. reticulata* Blanco (mandarina), *C. maxima* Burm. Merr. (zamboa) y *C. micrantha* Wester (papeda) (Ancillo & Medina, 2014).

## 1.2. PRODUCCIÓN DE CÍTRICOS

### 1.2.1. En el mundo

En la actualidad, el cultivo de los cítricos se extiende a todos los continentes, en dos franjas bien definidas que abarcan, en el hemisferio Norte, desde los 41° a 16° N, y en el hemisferio Sur, desde los 11° a 35° S (ver Figura 1.1).

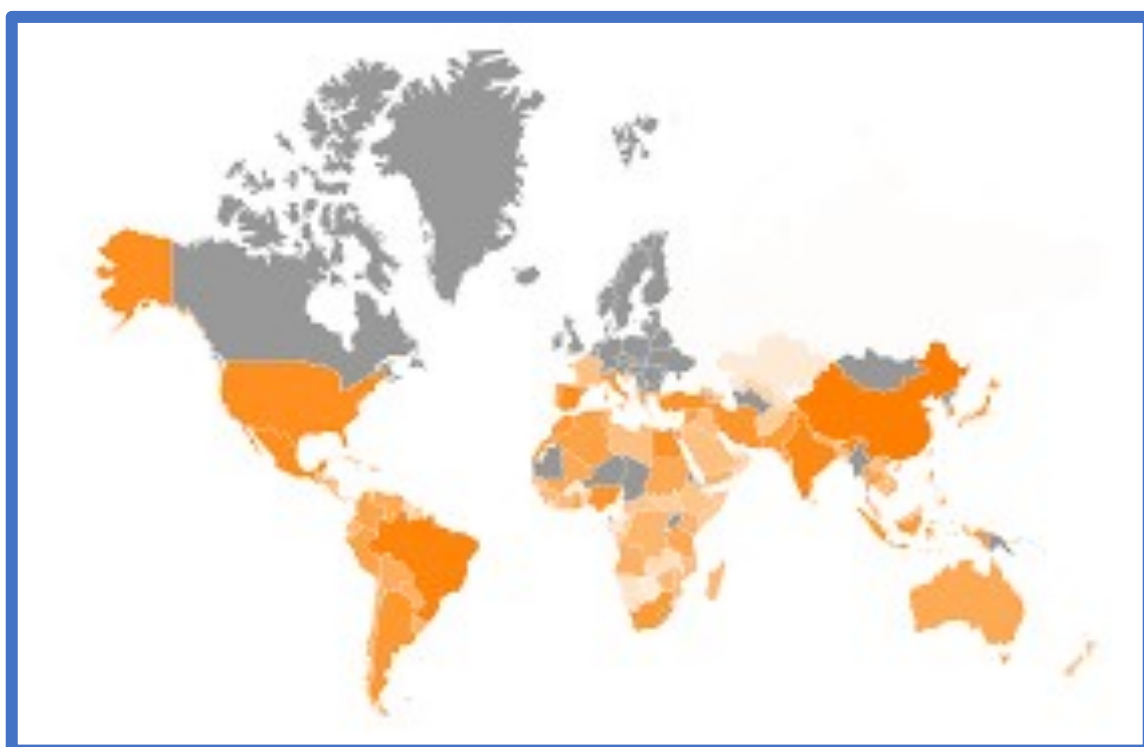


Figura 1.1. Países productores de cítricos.

Fuente: <https://www.atlasbig.com/es-mx/paises-por-produccion-total-de-citricos>

Los cítricos más cultivados son el naranjo dulce (*Citrus sinensis*, L) y diversas especies agrupadas dentro del término de mandarinos que agrupa diversas especies (*C. reticulata*, Blanco, *C. unshiu* Marcovitch, *C. nobilis* Loureiro, *C. deliciosa* Tenore, *C. Clementina* Hort. ex Tanaka, *C. reshni* Hort. ex Tanaka, *C. sunki* Hort. ex Tanaka), y sus híbridos como el tangor, un híbrido interespecífico (*C. reticulata* x *C. sinensis*).

En la campaña 2018/19 la producción total de los frutos cítricos a nivel mundial fue de 84,89 millones de t, de los cuales 31,00 millones de toneladas (t) corresponden a

mandarinas. El mayor productor fue China, con un volumen de 21,20 millones de t, seguida de la Unión Europea con 2,94 millones de t.

La producción mundial de naranjas en 2018/19 fue estimada en 54,3 millones de t, la mayor de los últimos ocho años, con un aumento interanual de 6,3 millones de t, debido principalmente a las condiciones meteorológicas muy favorables en EUA y Brasil. El primer productor mundial es Brasil con el 37%, seguido de China con el 13,3% y la UE con el 12%, que representan en conjunto el 62% de la producción mundial de naranjas.

### 1.2.2. En Argentina

En Argentina los cítricos ocupan el primer lugar en volumen de producción de frutas, seguidos por los frutales de pepita y de carozo. En Argentina la producción de mandarinas fue de 431 miles de t (Federcitrus, 2019).

En la temporada 2017/2018, el país se situó octavo en la producción mundial de cítricos, con 1.007 miles de t de naranjas, 431 miles de t de mandarinas, 1.989 miles de t de limones y 114 miles de t de pomelos (Federcitrus, 2019).

Diferentes regiones productoras del país tienen condiciones ecológicas ideales para el crecimiento de las diferentes especies cítricas. El desarrollo de los cultivos en el país se extiende principalmente a dos regiones: Noroeste (NOA), donde se producen naranjas, pomelos y limones, especialmente estos últimos en la provincia de Tucumán, y Noreste (NEA), donde predominan los cultivos de naranjas y mandarinas, que, a través de diferentes variedades orientadas a las preferencias del mercado, se cosechan y exportan prácticamente durante todo el año<sup>1</sup>.

En el NEA la principal área productora de cítricos es la denominada Mesopotamia Argentina, que comprende las provincias de Entre Ríos, Corrientes y Misiones y limita con los países de Paraguay, Brasil y Uruguay.

---

<sup>1</sup> <http://biblioteca.cfi.org.ar/wp-content/uploads/sites/2/2015/11/informe-sector-citricola.pdf#:~:text=En%20la%20Argentina%20pueden%20distinguirse%20dos%20regiones%20productoras,Corrientes%20y%20Misiones%2C%20as%20C3%AD%20como%20en%20Buenos%20Aires.>

*Corrientes*: concentra la mayor producción de naranjas (48% del total nacional), también produce mandarinas (18%), pomelos (16%) y limones (7%). Los cultivos se ubican en los departamentos de Bella Vista, Saladas, San Miguel, Murucuyá, General Alvear y Monte Caseros

*Entre Ríos*: produce la mayor cantidad de mandarinas (44% del total nacional), también pomelos (27%, destinados principalmente para elaboración de jugos), naranjas (12%) y limones (5%). Las quintas se ubican sobre la costa del Río Uruguay, en las localidades de Concordia, Chajarí y Federación.

*Misiones*: representa el 5,8% del área citrícola del país. La producción de naranjas y mandarinas es la principal actividad citrícola, gran parte de la producción se destina principalmente a industria.

Esta región citrícola también incluye, zonas de menor desarrollo en las provincias de Santa Fe y el sector limítrofe norte de la provincia de Buenos Aires y áreas bien definidas en las provincias de Formosa y Chaco.



Figura 1.2. Regiones Citrícolas de la Argentina



## 1.3. PLANTEAMIENTO Y DESCRIPCIÓN DEL PROBLEMA

### 1.3.1. Importancia de los pronósticos anticipados de producción

La fruticultura, al igual que toda actividad agrícola, es de alto riesgo, que unido a la gran interdependencia de los mercados genera la necesidad de contar con información anticipada y confiable sobre la producción y los precios, y la planificación se convierte en una herramienta de trabajo. Sin embargo, la agricultura en general y la fruticultura en particular, como actividades económicas, están regidas por un sinnúmero de factores internos y externos que hacen muy dificultoso cualquier intento de planificación (Winter, 1969).

Una de las herramientas necesarias para poder competir en los mercados nacionales e internacionales, es la predicción de los volúmenes a producir. Estimar la producción en cada temporada es importante para influir en las decisiones, que afectan tanto la comercialización como las recomendaciones en las prácticas de cultivo, y en la formulación con la suficiente antelación, de las estrategias necesarias a seguir (Bergh, 1982).

En países como Argentina, cuya economía depende en gran medida de la producción de cultivos, las estimaciones tempranas de cosecha son un requisito elemental para generar el presupuesto nacional, planificar estrategias para atender a desastres naturales y anticipar la demanda para el transporte y el almacenamiento (Lopresti *et al.*, 2015).

La incertidumbre en el pronóstico anual del volumen de cosecha en los cultivos perennes depende principalmente de los muy cambiantes rendimientos por hectárea. Las condiciones climáticas tienen una influencia directa en la producción, originando grandes variaciones interanuales en los volúmenes de fruta producida, con los consiguientes problemas en términos del uso eficiente de los equipos de procesamiento y de la disponibilidad de productos para desarrollar y mantener mercados (Winter, 1969).

Las especies frutales, y en particular los cítricos, presentan características especiales que dificultan las tareas de estimación de cosecha. Una de ellas es que son cultivos perennes de bastante longevidad, donde las variaciones en las técnicas culturales y cultivares definen una mayor variabilidad en las producciones. Cuando las plantaciones declinan, los productores en vez de arrancar y renovar las mismas por su dudosa rentabilidad, realizan injertos, replantes, etc. Generalmente la producción no resulta consecuencia de una planificación racional, sino de expectativas comerciales, produciendo cambios de algunas variedades por otras nuevas en función de los precios. Todos estos aspectos se traducen en una gran heterogeneidad en las plantaciones en cuanto a variedad, tamaño, edad, sanidad y, por lo tanto, rendimientos (Palacios, 2013).

Es por ello importante lograr la mayor anticipación y exactitud posible en los pronósticos de cosecha de frutas, con el fin de incrementar la base de información para la toma de decisiones de varios sectores: administradores de empresas frutícolas, empaques e industrias de jugos de fruta, que sufren el impacto de los cambios en los niveles de aprovisionamiento de productos. También permite tomar resoluciones más acertadas, a nivel empresarial, tanto en términos de calendarios de cosecha y equipos de procesamiento, como la comercialización o las necesidades de almacenamiento. A nivel gubernamental regional o nacional, este conocimiento permite formular políticas y planificar estrategias tendientes a la apertura de mercados sobre una base cierta de la producción que se puede ofrecer (Folwell *et al.*, 1994).

El Instituto Nacional de Investigación Agropecuaria (INIA) de Salto, Uruguay, a través del Programa Nacional de Citricultura, ha desarrollado técnicas de previsión de productividad y calidad de cítricos con el fin de adecuar volúmenes y calidad a las exigencias del mercado. Una de ellas es el muestreo de frutos a campo siguiendo protocolos que permiten tomar decisiones anticipadas de manejo. Han desarrollado también modelos de estimación de rendimiento, en función a las características de la floración y el desarrollo de frutos (Otero *et al.*, 2005).

En nuestro país, se han desarrollado y se vienen ejecutando diversos programas de pronóstico de producción para frutales.

En la provincia de Mendoza, en 1996, se inició un programa de pronóstico de cosecha frutícola, con el objetivo de estimar con anterioridad, la cosecha de manzano, peral, membrillero, duraznero, ciruelo, damasco, cerezo, almendro, olivo y nogal<sup>2</sup>.

La Facultad de Ciencias Agrarias de la Universidad Nacional del Comahue, a través de un convenio con los gobiernos provinciales de Río Negro y Neuquén, viene realizando trabajos de relevamientos de información, procesamiento y elaboración del pronóstico de las principales variedades de frutos de pepita existentes en la región, desde el año 1992<sup>3</sup>.

Avanza (2011), ha desarrollado una propuesta metodológica para la predicción de cosecha de naranjo dulce variedad Valencia late para una zona de la provincia de Corrientes, Argentina. Propuso un modelo basado en mediciones objetivas a campo, con determinación de la superficie o densidad y estimación de la carga y el tamaño medio de frutos a cosecha, en diferentes huertos.

Para estimar anticipadamente el volumen de producción es necesario conocer los factores que intervienen en ella, y así construir un modelo que represente la realidad en forma simplificada. De esta manera es posible considerar que la producción de frutales en un determinado momento se puede representar en función de esos factores.

### 1.3.2. Métodos para la estimación de producción

El modelo de predicción de producción que se propone en esta tesis está basado en el conocimiento de tres elementos fundamentales:

- (i) La cantidad real de plantas en producción, o alternativamente la superficie efectiva implantada con la variedad en estudio.
- (ii) La carga de frutos, expresada en número de frutos por planta o por hectárea.
- (iii) El tamaño (o peso) del fruto medio esperado al momento de la cosecha.

---

<sup>2</sup> Disponible en: <https://www.idr.org.ar/>

<sup>3</sup> Disponible en: <https://docplayer.es/42642541-Pronosticos-de-produccion.html>.

Para determinar el número real de plantas en cada estrato o la superficie efectiva, se requiere de un censo frutícola del área donde se va a realizar la predicción, mientras que los otros dos componentes surgen de mediciones a campo realizadas sobre muestras convenientemente seleccionadas o son obtenidos mediante sensores remotos.

Con la información precedente, la estimación de producción de la variedad en cuestión para un huerto o área determinada estará dada por el modelo:

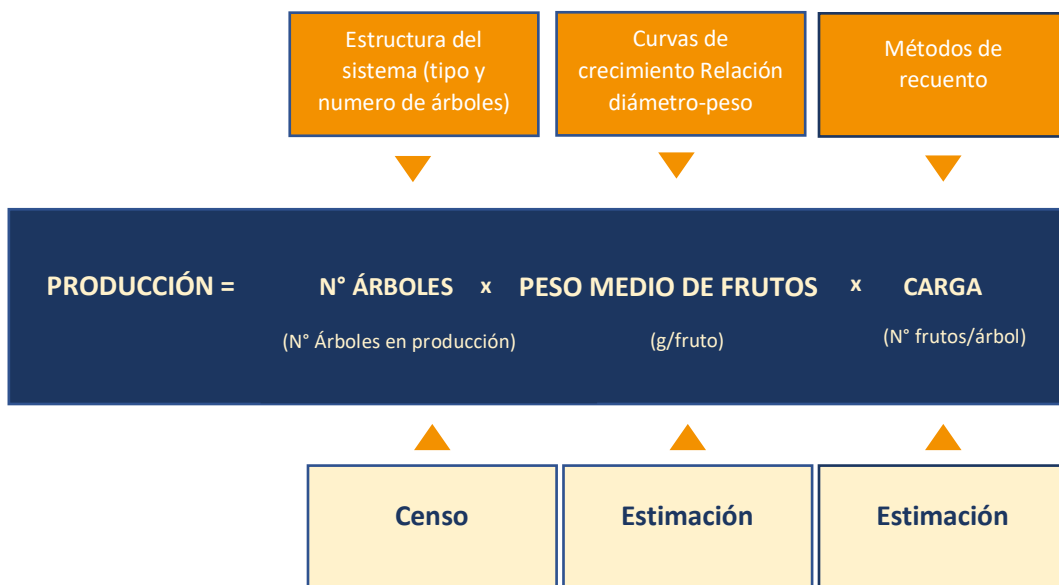


Figura 1.3. Modelo de pronóstico de producción

## 1.4. OBJETIVOS

Las naranjas dulces son los cítricos más consumidos mundialmente, pero en los últimos años las mandarinas y sus híbridos presentan una demanda creciente. La disponibilidad de estas frutas está asegurada durante todo el año gracias a los tratamientos de post cosecha, al desarrollo de variedades tardías y a la producción en ambos hemisferios. Entre los híbridos de mandarinas, los tangores se encuentran entre los más requeridos por los mercados internacionales, por lo que su cultivo se está incrementando rápidamente. En el nordeste argentino las naranjas dulces y las mandarinas constituyen exportaciones agrícolas de relevancia (Feder Citrus, 2019), por lo que en esta tesis se

decidió trabajar con tangor Murcott y posteriormente se incorporaron naranjo dulce Valencia late y Salustiana.

#### 1.4.1. General

Definir un modelo para la estimación anticipada de los volúmenes de producción y distribución de tamaños de fruto a cosecha en mandarino tangor Murcott, para la provincia de Corrientes, Argentina. Complementariamente se definieron modelos para naranjo dulce Valencia late y Salustiana.

#### 1.4.2. Particulares

- ✓ Hallar el mejor modelo para describir las curvas de crecimiento de frutos de tangor Murcott.
- ✓ Construir tablas de crecimiento para diferentes diámetros de fruta con destino al mercado externo, que permitan estimar el tamaño de éstas en cualquier momento previo a la cosecha, con la finalidad de ser utilizadas en las predicciones de volúmenes de producción y en el raleo de frutos.
- ✓ Estimar el peso de frutos a cosecha en función del diámetro, mediante el empleo de un modelo matemático que relacione ambos caracteres.
- ✓ Definir metodología para la estimación de la carga ( $N^{\circ}$  frutos. árbol<sup>-1</sup>) en cítricos.
- ✓ Evaluar la precisión del modelo de pronóstico y las tablas de crecimiento elaboradas para tangor Murcott.

Si bien la propuesta original del plan de tesis se centraba en el pronóstico de producción de mandarino tangor Murcott, se logró incorporar huertos de naranjo dulce de las variedades Salustiana y Valencia late, ubicados en las provincias de Corrientes y Entre Ríos, y recabar información de campo e imágenes satelitales, por lo que se incluyó como objetivo específico la definición de metodología para la estimación de carga en

cítricos, que contempla, además de mandarino tangor Murcott estas dos variedades de naranjo dulce.

## 1.5. ANTECEDENTES

### 1.5.1. Número de árboles

El inventario de árboles identificados por cultivar, patrón, edad y ubicación es una información necesaria para realizar pronósticos de producción, se actualiza anualmente a través de conteos físicos o con empleo de fotografías aéreas o imágenes satelitales. Los árboles se agrupan por estratos productivos de edades, en correspondencia con la entrada de producción y con el incremento de la edad de la plantación (Bentarcourt *et al.*, 2013).

En el año 1996, el Instituto de Desarrollo Rural (IDR) comenzó con el desarrollo de las primeras herramientas para la realización de pronósticos de producción de las especies frutales de mayor importancia en la provincia de Mendoza. Para ello necesitan contar con información actualizada de superficie cultivada, edad de las plantas, variedades implantadas, estado de los montes, entre otros. Desde la temporada 1997/98 y hasta la fecha, se realiza la estimación de producción de duraznos y ciruelas para industria, duraznos para fresco, ciruelas para fresco, peras y manzanas. Posteriormente se fue agregando, adaptando y validando metodología para nuevas especies, como cerezo, damasco y olivo. El método utilizado se basa en mediciones realizadas a campo y ajustes a partir de información de temporadas anteriores.

El tamaño muestral, es decir la cantidad de parcelas y plantas a medir, se calcula previendo un error menor al 10 por ciento. Para definir la componente de superficie de las especies monitoreadas (edad de las plantaciones, estado, variedad, densidad de plantación, etc.), se utiliza el censo general más actual y censos específicos posteriores con ajustes por edad de plantación y estado de los montes. Se obtienen volúmenes de producción esperados para la provincia de Mendoza por especie y oasis productivo y la

distribución de calibres esperada a cosecha. Estos datos se difunden todos los años a mediados de noviembre para damasco y cereza y la primera semana de diciembre para el resto de los frutales <sup>4</sup>.

Para la determinación del número de frutos por árbol o por hectárea, en el pronóstico de cosecha de frutales de pepita y carozo desarrollado en Mendoza, Argentina, Gil (2003) sugiere realizar un muestreo en dos etapas. Previamente agrupa las parcelas de las diferentes fincas en estratos homogéneos respecto a la variedad y edad de los árboles. Luego, en la primera etapa de muestreo selecciona al azar el 12% del total de parcelas, y en la segunda etapa selecciona aleatoriamente árboles. Para esta última etapa utiliza el criterio de seleccionar menos del 1% del número promedio de árboles presentes por parcela.

Avanza *et al.* (2010) determinaron que, para estimar la carga al momento de cosecha, con una precisión entre el 89 y el 90 %, se requiere un tamaño de muestra entre 18 y 21 árboles por huerto.

### 1.5.2. Peso de fruto medio

Los métodos basados en mediciones a campo para estimar el peso del fruto medio a cosecha se pueden subdividir en los que predicen el tamaño de frutos a partir de modelos estocásticos y los que utilizan curvas de crecimiento. Los primeros se basan en el supuesto de que existe una relación entre el tamaño de los frutos (peso) y la carga de los árboles. Las previsiones de producción de frutales de pepita realizadas por el Centro de Servizi Ortofrutticoli para la Unión Europea, se basan en modelos estocásticos para la estimación del peso medio de frutos. En cambio, en Argentina para las mismas especies frutales, dicha relación (peso vs. carga), no se ha logrado establecer dada la heterogeneidad de las plantaciones, por lo que se utilizan curvas de crecimiento para la modelización del crecimiento de los frutos en función del tiempo y poder así realizar la

---

<sup>4</sup> Pronósticos por temporada y cultivo, disponible en: <https://www.idr.org.ar/pronostico-de-cosecha/>.

estimación del peso de los frutos al momento de cosecha (Álvarez *et al.*, 2001, Bramardi *et al.*, 2005).

Varias funciones matemáticas han sido propuestas para modelar curvas decrecimiento.

Del Valle (1989) utilizó el modelo de Von Bertalanffy como ecuación diferencial para describir las curvas de incremento anual, con el objetivo de determinar la edad y el crecimiento de dos especies de árboles tropicales; encontró que la versatilidad de la ecuación del modelo se evidencia con curvas de crecimiento y, dependiendo de los datos, se pueden obtener curvas sigmoidales de tipo Logístico, en roble.

Demirsoy & Demirsoy (2007), utilizaron modelos de regresión lineal múltiple para predecir el peso y volumen de frutos de diferentes cultivares de durazno, a partir de la altura, largo y ancho de los mismos en época temprana. Para ambos modelos, las estimaciones de los parámetros resultaron significativas y obtuvieron valores de  $R^2 > 0,99$ .

Con el fin de estimar el peso de frutos de diferentes cultivares de frambuesa a través de características de la caña del sistema de conducción (número de frutos por caña, altura y diámetro de caña, rendimiento por caña), Eyduran *et al.* (2008) seleccionaron ecuaciones de regresión lineal múltiple con valores de  $R^2$  entre 0,85 y 0,99 de acuerdo con el cultivar.

La generación de curvas de crecimiento de los frutos en base al peso requiere la determinación del peso individual de los frutos en mediciones anticipadas a la cosecha, lo que implica la remoción de los frutos del árbol desde etapas tempranas. La estimación del peso en función del tamaño, expresado a través del diámetro ecuatorial o polar, es una metodología de amplio uso por la estrecha relación que existe entre las variables (Esposti, *et al.*, 2008; Avanza, 2011) y las técnicas de muestreo sencillas y no destructivas que se utilizan.

Alvarez & Boche (1999) indicaron que, en los frutos de carozo y en algunas bayas con semillas, el crecimiento presenta más de un punto de inflexión mostrando dos etapas que suelen representarse separadamente por modelos sigmoideos simples. Analizaron y compararon tres modelos dobles sigmoides obtenidos como una generalización de los



modelos sigmoideos simples, y ajustaron datos de crecimiento en diámetro de una variedad de Nectarin tardío en Neuquén (Argentina), encontrando que la función logística generalizada resulta ser el modelo más adecuado para la descripción del crecimiento de los frutos.

En la región del Alto Valle, Argentina, en parcelas productivas de peras William's, Zon *et al.* (2011) midieron el peso de frutos y describieron un patrón exponencial para el período que va desde el cuajado del fruto hasta la cosecha comercial, al considerar el diámetro de los frutos definieron un patrón sigmoideo. En la misma región se ha utilizado un modelo logístico para describir el crecimiento de peras William's y Packham's Triumph (Bramardi *et al.*, 1998), desde los 20 días desde plena floración hasta pasada la cosecha comercial. También en manzanas Red Delicious y Granny Smith de ciclo medio y largo se ajustaron patrones sigmoideos de crecimiento realizando mediciones sucesivas de diámetro de frutos (Stangaferro *et al.*, 2001).

Godoy *et al.* (2008), indicaron que el crecimiento de los frutos del arándano suele presentar un patrón de doble sigmoide. Y estudiaron dicho crecimiento con el objetivo de caracterizar y comparar cultivares que difieren en el tiempo de maduración, utilizando modelos mixtos no lineales exponenciales. Los modelos mixtos exponenciales ajustaron con precisión los patrones de crecimiento de los arándanos y pusieron de manifiesto las diferencias entre los cultivares de maduración temprana y tardía.

En cítricos, específicamente naranjos de variedades tardías como Valencia late, se describe un crecimiento sigmoidal del diámetro de los frutos con 3 fases, similar al observado en frutales de pepita, destacándose que naranjo tiene mayor longitud de ciclo (Avanza *et al.*, 2008).

En caso de que las curvas de crecimiento se confeccionen en base al diámetro de frutos, para realizar pronósticos de cosecha expresados en kilos se debe transformar ese diámetro a peso. Para ello Esposti *et al.* (2008) estimaron el peso a cosecha en frutos de naranjo Valencia y Hamlin, en función al diámetro, aplicando una ecuación de regresión lineal (con un coeficiente  $R^2 = 0,83$ ). Bramardi (1995), ajustó modelos potenciales a la relación peso- diámetro de frutos de pera William's y Packham's Triumph, con coeficientes de determinación ( $R^2$ ) del 99%. Además, Avanza (2011) estableció que el

peso de los frutos de naranjo Valencia late, puede ser estimado a través del diámetro ecuatorial mediante un modelo potencial, que mostró un buen ajuste ( $R^2 = 0,94$ ).

Las curvas de crecimiento de frutos son también muy útiles para ajustar prácticas de raleo y predecir calibres a cosecha. Pérez de Villareal *et al.* (2021), registraron el crecimiento de los frutos en ciruelo, evaluando semanalmente: peso fresco, diámetro longitudinal y ecuatorial de frutos y carozos y relación diámetro longitudinal/diámetro ecuatorial. El peso fresco ajustó a un modelo logístico doble ( $r = 0,96$ ), los diámetros longitudinal y ecuatorial a modelos cuadráticos ( $r = 0,97$ ), la relación entre ambos a una curva polinomial de tercer grado ( $r = 0,96$ ) con una disminución de la relación a través del tiempo.

### 1.5.3. Carga

La determinación del número de frutos a producir por árbol o por hectárea puede ser realizada por diversos caminos.

#### 1.5.3.1. Informantes calificados

En Estados Unidos, el Servicio de Estadísticas (NASS) del Departamento de Agricultura (USDA)<sup>5</sup>, provee proyecciones de producción y/o estimaciones de producción y uso de 19 especies frutales no cítricas, 7 especies cítricas y 6 frutas secas. La metodología de encuestas utilizada varía por producto y por estado y puede ser categorizada como: No probabilística, probabilística o de medición objetiva. Estos métodos están ordenados en función creciente al grado de confiabilidad y del costo. El NASS y las contrapartes del Estado Federal utilizan los tres métodos, dependiendo del producto y de la disponibilidad de recursos.

---

<sup>5</sup> Información y Servicios del United States Department of Agriculture (USDA, EUA).

Dumanski y Onofrei (1989) señalan que muchas veces el rendimiento del cultivo se estima por las opiniones de expertos, mediante la evaluación visual de las condiciones del cultivo, como color, vigor de las plantas, densidad de plantación. Esta evaluación (ocular) se puede combinar con mediciones de campo y fórmulas empíricas, conocidas como el método de evaluación de expertos, que se puede aplicar a una escala relativamente grande, sin embargo, la estimación visual del rendimiento de los cultivos requiere no solo una familiaridad práctica sino también técnica con el potencial de rendimiento de diferentes variedades de cultivos en diferentes entornos. Por lo tanto, la precisión de la evaluación del rendimiento, en este método, dependerá en gran medida del nivel de experiencia del personal involucrado en la evaluación (Casley y Kumar, 1988).

En España, la Consellería de Agricultura, Pesca y Alimentación de la Generalitat Valenciana, realiza previsiones anuales de la producción de la cosecha citrícola de la campaña entrante de las provincias de Alicante, Castellón y Valencia, por especies y variedades. Emplean una metodología estadística denominada “marco de áreas” con la variante denominada “segmentos” para adaptarse a la metodología común aplicada en las estadísticas agrícolas que se basan en un muestreo aleatorio del territorio<sup>6</sup>.

En Sudáfrica, en la temporada 2003/04, se comenzó a implementar una propuesta metodológica para la estimación de producción de frutos de carozo y uvas basada en: 1) primera estimación teórica; 2) debate con representantes técnicos que puedan arrojar más idea sobre la reducción o aumento de la producción; 3) segunda estimación de producción (normalmente un mes después de la primer estimación), en la cuál es posible añadir más detalles sobre tamaño de frutos, y habrá mayor seguridad en el recuento de frutos por regiones y sectores. El número total de estimaciones dentro de cada sector de frutas dependerá de cada cultivar, pero en la mayoría de los casos es necesario llevar a cabo una tercera o incluso cuarta estimación en función de la temporada de producción (Louw & Fourie, 2003).

En Zimbabue, entre 2011 y 2012, se realizaron encuestas para determinar las pérdidas en naranja y las percepciones de los agricultores sobre la cadena de valor del

---

<sup>6</sup> Consellería de Agricultura, Pesca y Alimentación, Generalitat Valenciana (Valencia, España).

suministro de naranjas dulces. Los datos se recogieron mediante cuestionarios tipo Likert administrados por un entrevistador y entrevistas informales sobre: prácticas de gestión de los huertos, infestación por plagas, actividades de manipulación de la fruta y prácticas de comercialización. La muestra del estudio provenía de dos pueblos seleccionados al azar en cada uno de los distritos administrativos con una importante producción de naranjas dulces (Musasa *et al.*, 2013).

En Sao Paulo, Brasil, a través de encuestas por muestreo, anualmente se realiza la estimación de producción por planta de los principales cultivares de naranjo dulce para industria. Estas encuestas se realizan a mediados de abril, para conocer la predicción de cosecha y establecer estrategias de comercialización del sector (Paulino *et al.*, 2007).

#### *1.5.3.2. Información a nivel parcelario*

Desde el año 2000, en Australia se realiza la estimación nacional de producción de naranjas Navel y Valencia, mandarinas, limones y pomelos, para las tres mayores regiones productoras en conjunto con el Australian Citrus Growers Federation (ACGF). La metodología consiste en determinar densidad y tamaño de frutos, y otras características de las plantaciones mediante tres operarios. La base de datos incluye encuestas a productores, estadísticas de plantación, densidad de frutos por sector de copa, y medidas estadísticas (Australia Citrus Growers ACGF, 2019).

Conocer anticipadamente el volumen de producción en los cítricos permite a los productores tomar decisiones de gestión adecuadas para la próxima temporada. Se recomienda realizar estimaciones de la carga dos veces durante la temporada de cultivo, una temprana y otra tardía. La estimación temprana ayudará a determinar si son necesarias prácticas como el aclareo manual, mientras que la tardía proporciona un rendimiento estimado para la temporada. El tamaño de fruta es el factor más importante a la hora de determinar el rendimiento en el mercado y la estimación de la carga permite decidir las prácticas de gestión que darán el mejor tamaño de fruta.

Los métodos de estimación se basan en el *ojo experimentado*, donde los cultivadores experimentados se hacen una idea de la carga del cultivo a partir del tiempo que han

pasado en los huertos y, considerando los rendimientos de temporadas pasadas pueden estimar la carga de la nueva temporada (ACGF, 2019).

Los procedimientos se pueden agrupar en:

- (i) Estimación de un árbol: los agricultores hacen un recuento de frutos en un árbol representativo del huerto y luego multiplican ese valor por el número total de árboles.
- (ii) Marcos de recuento: se utiliza un marco de recuento para determinar la cantidad de fruta en sectores del árbol. Estos datos, junto con información de densidad y tamaño de los árboles, permitirá estimar el rendimiento de fruta.

Cualquiera de los métodos elegidos se basa en que la estimación de la carga al momento de la cosecha se hace en función al recuento de la densidad de fruta por el número de árboles.

Según Avilán *et al.* (2001) una manera de estimar la producción en frutales de tipo arbóreo es la utilización del "índice de fructificación", que permite establecer una relación entre la superficie de producción, área foliar o follaje y la producción de frutos. Para ello, se tiene en cuenta que la copa se asemeja a un cuerpo geométrico y se determina su volumen o la superficie lateral o interna de la misma. Se realiza el conteo de los frutos en el árbol y se establece la relación con el volumen o la superficie antes determinada ( $N^{\circ}$  frutos. $m^3$  o  $N^{\circ}$ fruto.V $m^2$ ). Si se quiere expresar los resultados en unidades de peso, se toma una muestra de los frutos y se establece el peso de estos.

En Sudáfrica, Button (2009) estableció que, para estimar la producción, el número de frutos por árbol es un parámetro importante para predecir el rendimiento o la carga de la cosecha. El número de frutos por árbol en el momento de la cosecha viene determinado tanto por la intensidad de la floración como por el cuajado alcanzado. La estimación del número de frutos sólo debe realizarse después del periodo de caída fisiológica de los frutos. No obstante, en muchos casos una gran proporción de la fruta continúa cayendo después del período de caída fisiológica de la fruta y hasta el momento de la cosecha.

Existen diferentes formas de determinar el número de frutos por árbol (Verreyne, 2009).

- (i) Contar el número de frutos por árbol de uno o más árboles representativos por huerto o unidad.
- (ii) Extraer todos los frutos de uno o varios árboles representativos después del periodo fisiológico de caída de los frutos. Estos frutos se cuentan, se determina el diámetro de estos y se calcula su distribución de tamaños. A partir del porcentaje de frutos de cada categoría de tamaño y del peso medio de los frutos de cada categoría en el momento de la cosecha, se puede predecir la carga ( $\text{kg. árbol}^{-1}$  y  $\text{t.ha}^{-1}$ ).

El rendimiento viene determinado tanto por el número de frutos por árbol como por el tamaño de estos. Las cosechas abundantes con un mayor número de frutos generalmente dan lugar a frutos más pequeños y, si se identifican en una fase temprana de la temporada, se pueden aplicar estrategias de reducción del número de frutos, como el aclareo o raleo manual. La reducción de la cosecha en esta fase no sólo puede dar lugar a frutos más grandes, sino que puede reducir la incidencia de la alternancia anual en la fructificación, característica de los cítricos (Verreynne, 2009).

El número de frutos por árbol es útil bajo condiciones uniformes, en donde las diferencias de tamaño de los árboles sean insignificantes, pero cuando se trata de plantaciones con tamaños variables de árboles es esencial normalizar la carga frutal por tamaño del árbol. Es así como se ha utilizado la cantidad de fruta en relación con el área de sección transversal de tronco, aunque su relación con el tamaño del árbol se va perdiendo con el tiempo, debido a las intervenciones de poda que se realizan cada año. (Verreynne, 2009).

Una estimación que ha ido tomando importancia a través del tiempo es la eficiencia de convertir la luz en energía dentro de la fruta ( $\text{kg de fruta por unidad de luz interceptada}$ ), existiendo una directa relación entre la luz interceptada por la copa y la productividad por unidad de superficie. Este índice permite realizar comparaciones minimizando las diferencias entre densidad, forma y tamaño de los árboles (Reginato *et al.*, 2007).

En Argentina, Reeb & Tassile (2007), en cultivos de manzanas y peras establecieron que la carga de los frutos, expresada en unidades por planta o por hectárea, es un

parámetro que debe estimarse cada año dado que su valor depende de las condiciones que caracterizan a la temporada. Debe determinarse para cada especie y variedad por separado, estratificando por sistema de conducción y edad del monte debido a la heterogeneidad. Además, a los efectos de considerar el error incurrido por los relevadores al momento de contabilizar de manera visual la carga de los árboles, introducen un factor de corrección de fruto oculto, cuyo valor es obtenido de manera experimental y es particular para las distintas variedades.

Marini (2019) comprobó que el tamaño de los frutos al momento de la cosecha se correlacionaba bien con el tamaño de estos en todas las fechas de muestreo, pero, como se esperaba, la correlación mejoraba en las últimas fechas de muestreo. Una manera de estimar el tamaño medio de los frutos en la cosecha es dividir el rendimiento total por árbol entre el número de frutos del árbol, pero cosechar, contar y pesar todos los frutos de un árbol puede llevar tiempo y es costoso, por lo que los investigadores desearían recoger una muestra de frutos a partir de la cual se pueda estimar el tamaño medio de los mismos <sup>7</sup>.

### *1.5.3.3. Información de sensores remotos*

En las últimas décadas, debido a los avances en informática, robótica, inteligencia artificial y sensores remotos, se han desarrollado diferentes opciones para estimar el rendimiento de los cultivos dentro de la llamada "agricultura de precisión". Estos métodos semiautomatizados utilizan diferentes sistemas de captura de imágenes y aplican técnicas de segmentación Teixidó *et al.* (2018).

Los sensores remotos permiten estimar la producción con mayor antelación y menor coste que las técnicas tradicionales basadas en la información de campo. Se han probado con éxito diversos métodos de captura de imágenes, por ejemplo, Wang *et al.* (2021),

---

<sup>7</sup> Disponible en: <https://extension.psu.edu/fruit-harvest-estimating-apple-yield-and-fruit-size>.

desarrolla una tecnología de para imágenes ópticas de teledetección de alta resolución basada en el aprendizaje profundo.

Zeeshan *et al.* (2016) presentaron una técnica para estimar rendimiento de naranjas a partir de las imágenes de los árboles. Utilizaron la segmentación k-means para el reconocimiento de la fruta, que segmenta la imagen con exactitud, permitiendo así una estimación más precisa del rendimiento. Crearon un conjunto de datos que contiene 83 imágenes de árboles con cítricos de tres campos diferentes y detectaron los frutos en el árbol con una precisión del 91,3%. Además, encontraron una fuerte correlación entre el recuento de frutos manual y el automatizado, obteniendo coeficientes de determinación  $R^2$  de hasta 0,99.

Gracias a la gran variedad de sensores a bordo de los satélites civiles, los valores espectrales y las propiedades temporales de la superficie terrestre pueden captarse y controlarse a diferentes escalas espaciales y temporales, Begué *et al.* (2018). De hecho, las imágenes satelitales de baja resolución se han utilizado ampliamente para el seguimiento de los cultivos y la previsión del rendimiento durante más de 30 años y desempeñan un papel importante en un número creciente de sistemas operativos.

La combinación de una alta frecuencia temporal con su amplia cobertura geográfica, generalmente asociada a un bajo coste por unidad de superficie, hacen de estas imágenes una opción conveniente. Además, los sistemas tradicionales siguen siendo cruciales para garantizar la disponibilidad de series temporales largas, tal y como necesitan la mayoría de los métodos de predicción de rendimientos utilizados hoy en día (Teixidó *et al.*, 2018).

Hace varios años se reportó diferentes modelos de cámaras disponibles actualmente en el mercado con diversas características (por ejemplo, blanco y negro, RGB, térmica, etc.). Después de los sensores de color, las cámaras RGB se utilizaron mayoritariamente en todos los sistemas de detección (para captar el color), lo que facilitó el proceso de detección. Las imágenes térmicas también se han utilizado en algunos trabajos de detección de frutas (Gongal *et al.*, 2015). Aquí, la característica de cada objeto (es decir, rama, tallo, etc.) se detecta en función de la temperatura, ya que las frutas tienen una temperatura más alta que los objetos de fondo. La imagen espectral es un modelo de cámara de nueva generación (es decir, sensores multispectrales e hiperspectrales)



que se utiliza actualmente. Proporciona información adicional relacionada con los detalles espectrales (en cada espacio de color) junto con las características del color. Incluso si el color de la fruta y el fondo (como las hojas y el tallo) son iguales, la detección de la fruta puede realizarse utilizando la información espectral (Feng *et al.*, 2019).

Zhao *et al.* (2016) detectaron los píxeles de fruta de cítricos verdes inmaduros utilizando el método de la suma de la diferencia transformada absoluta, emplearon el clasificador SVM para eliminar los falsos positivos (basado en características texturales) y obtuvieron los valores de precisión y recuerdo de 0,88 y 0,80, respectivamente.

Los sensores remotos permiten hacer estimaciones de producción de forma anticipada y con un costo menor que las técnicas tradicionales. Además de su objetividad, la observación por teledetección proporciona datos homogéneos que pueden ser registrados geográfica y temporalmente. Sin embargo, el uso de la teledetección no es tan eficaz para estimar el rendimiento de los cultivos como para estimar la superficie ocupada, donde la observación es más directa y donde encontrar un método aplicable a diferentes cultivos, zonas y las temporadas de crecimiento es más viable.

## **1.6. HIPÓTESIS**

Se establece como hipótesis de trabajo que es posible definir un modelo que permita estimar en forma anticipada los volúmenes de producción y la distribución de tamaños de frutos a cosecha en tangor Murcott con un grado de precisión aceptable.

## 1.7. BIBLIOGRAFÍA

1. Ancillo, G. & Medina, A. 2014. Monografías Botánicas. Los Cítricos. Universidad de Valencia, España. p: 156.  
[http://www.jardibotanic.org/fotos/pdf/publicacion\\_2\\_84\\_LOS\\_CITRICOS-ESP.pdf](http://www.jardibotanic.org/fotos/pdf/publicacion_2_84_LOS_CITRICOS-ESP.pdf)
2. Avilán, L.; Marín, C.; Rodríguez, M.; Ruiz, J. 2001. Crecimiento, floración y producción del mango sometido a diferentes tratamientos en plantaciones de alta densidad. *Agronomía Tropical*. 51(1):29-47.
3. Álvarez, A; Boche, S. 1999. Modelos matemáticos para describir crecimientos doble sigmoideos en frutos de nectarin tardío (c.v. Sun Grand). *AgroSur*. 27 (1): 21-28.
4. Alvarez, A.; Bramardi, S.; Stangaferro, S. 2001. Modelos de predicción anticipada del peso medio de los frutos en producción de frutos de pepita. X Congreso Latino Americano de Biometría. Campinas. Brasil.
5. ACGF, Australian Citrus Growers Federation. 2019. Fruit Size Management Guide. Part.2 (en línea) [http://www.dpi.nsw.gov.au/\\_data/assets/pdf\\_file/0006/138831/Fruit-Size-Guide-PART-2.pdf](http://www.dpi.nsw.gov.au/_data/assets/pdf_file/0006/138831/Fruit-Size-Guide-PART-2.pdf) [Consulta: Junio, 2021].
6. Avanza, M.; Bramardi, S.; Mazza, S. 2008. Statistical Model to Describe Growth Pattern in Sweet Orange “Valencia Late”. *Spanish Journal of Agricultural Research*, 6(4):577–585.
7. Avanza M.M.; Bramardi, S.; Mazza, S.M. 2010. Optimal sample size for evaluate the pattern growth of Valencia late orange fruit. *Revista Brasileira de Fruticultura*. 34 (4): 1154-1163.
8. Avanza, M.M. 2011. Desarrollo de una propuesta metodológica para la predicción de cosecha de naranjo dulce (*Citrus sinensis* L. Osbeck) var. Valencia late en la provincia de Corrientes, Argentina. Tesis Doctoral de la Universidad Nacional del Nordeste, Área Recursos Naturales.
9. Begué, A.; Arvor, D.; Bellon, B.; Betbeder, J.; De Abelleira, D.; Ferraz, R.; Lebourgeois, V.; Lelong, C.; Simões, M.; Verón, S. 2018. Remote Sensing and Cropping Practices: A Review. *Remote Sens*. 10 (1): 99.

10. Bentarcout, M.; García-Álvarez, M.; Reyes-Rivas N. 2013. Pronóstico de Producción en Frutales. *CitriFrut*. 3 (11).
11. Bergh, O.1982. Prediction of apple fruit size: a promising model. *Agroplanta*, 14: 43-46.
12. Álvarez, A., Boche, S. 1999. Modelos matemáticos para describir crecimientos doble sigmoideos en frutos de nectarin tardío (c.v. Sun Grand). *AgroSur*, 27 (1): 21-28.
13. Bramardi, SJ. 1995. Modelos de predicción de tamaños de fruta a la cosecha en base a mediciones sucesivas durante el período de crecimiento. Peras cvs. William'S y Packman's Triumph en la región del Alto Valle de Río Negro y Neuquén. Tesis para optar el grado de Magister Scientiae- Área Biometría. Escuela para graduados Convenio Facultad de Agronomía- UBA e Instituto Nacional de Tecnología Agropecuaria. 169 pp.
14. Bramardi, SJ; Castro, H; Zanelli, ML. 1998. Fruit growth pattern of pear cv. Bartlett and Packham's Triumph to improve hand thinning. *Acta Horticulturae* (ISHS). 475: 283-293. CAB Abstracts. AGRICOLA Database. ISBN 978-90-66059-20-7. ISSN 0567-7572. DOI: 10.17660/Acta Horticulturae. 475.34.
15. Bramardi, S.; Tassile, V.; Reeb, P.; De Bernardin, F. 2005. Comparación de métodos para la predicción anticipada del peso de fruta a cosecha: Curvas de crecimientos vs. Modelos estocásticos. X Reunión del Grupo Argentina de Biometría, Corrientes, Argentina, p. 6.
16. Button, J. 2009. Fruit development in *Citrus paradisi* (Macf. W. Marsh) seedless in five localities with special reference to the rind. Thesis. Dept. of Horticulture, University of Natal, Pietermaritzburg, South Africa.
17. Casley, D.J., Kumar K. 1988. The collection, analysis and use of monitoring and evaluation data. Johns Hopkins University Press for the World Bank, Baltimore
18. Del Valle, J.I. 1989. La ecuación del crecimiento de Von Bertalanffy en la determinación de la edad y el crecimiento de árboles tropicales. *Revista Facultad Nacional de Agronomía*. Medellín, Colombia. 39 (1).

19. Demirsoy, L.; Demirsoy, H. 2007. Characteristics of some local and standard sweet cherry cultivars grown in Turkey. *Journal of the American Pomological Society*. 57. 128-136.
20. Dumanski, J.; Onofrei C. 1989. Techniques of crop yield assessment for agricultural land evaluation. *Soil Use Manage* (5):9-15
21. Esposti, M.D.D.; de Siqueira, D.L.; Cecon. P.R. 2008. Crescimento de frutos da tangerineira Ponca (*Citrus reticulata* Blanco). *Revista Brasileira de Fruticultura*. 30 (3): 657-661.
22. Eyduran S.P.; Eyduran, E.; Agaoglu, Y.S. 2008. Estimation of fruit weight by cane traits for eight American blackberries (*Rubus fruticosus* L.) cultivars. *African Journal of Biotechnology*. 7 (17): 3031-3038.
23. Federcitrus, 2019. La Actividad Citrícola Argentina (en línea). <http://www.federcitrus.org.ar/actividad-citricola-2018.pdf>. [Consulta: Junio, 2021].
24. Feng, J., Zeng, L., and He, L. 2019. Apple fruit recognition algorithm based on multi-spectral dynamic image analysis. *Sensors* 19:949.
25. Folwell, R.J.; Santos, D.E; Spayd, S.E.; Porter, L.H.; Wells, D.S. 1994. Statistical Technique for Forecasting Concord Grape Production. *American Journal Enology and Viticulture*. 45: 63-70.
26. Gil, F. 2003 Curso de Capacitacion NEA. Técnicas de estimación para cítricos. Instituto de Desarrollo Rural. Programa de Pronóstico Frutícola. 53 pp.
27. Godoy, C; Monterubbianesi, G; Tognetti, J. 2008. Analysis of highbush blueberry (*Vaccinium corymbosum* L.) fruit growth with exponential mixed models. *Scientia Horticulturae*. 115: 368–376.
28. Gongal, A., Amatya, S., Karkee, M., Zhang, Q., and Lewis, K. (2015). Sensors and systems for fruit detection and localization: a review. *Computer Electronic Agriculture* .116: 8-19.
29. Lopresti, M.F.; Di Bella, C.M.; Degioanni, A.J. 2015. Relationship between MODIS-NDVI data and wheat yield: A case study in Northern Buenos Aires province, Argentina, *Information Processing in Agriculture*. 2 (2): 73-84.

30. Louw, D.; Fourie, M. 2003. Crop estimate methodology for fruit, (en línea) <http://www.dfpt.co.za/fruitlog/Phase%203%20reports/Crop%20estimate%20methodology.pdf>. [Consulta: Junio, 2021].
31. Marini, R.P; Schupp, R.J.; Baugher, T.A; Crassweller, R. 2019. Sampling Apple Trees to Accurately Estimate Mean Fruit Weight and Fruit Size Distribution. HortScience. 54 (6): 1017-1022.
32. Musasa, S.; Mvumi, B.; Manditsera, F.; Chinhanga, J.; Musiyandaka, S.; Chigwedere, C. 2013. Postharvest orange losses and small-scale farmers' perceptions on the loss causes in the fruit value chain: a case study of Rusitu Valley, Zimbabwe. Food Science and Quality Management 18: 1-8.
33. Otero, A.; Goñi, C.; Carrau, F.; Zefferino, E.; Brennan, M. 2005. Previsión de productividad y calidad en citrus: herramientas para una gestión sustentable. Revista INIA – Hortofruticultura. 4: 21-24.
34. Palacios, J. 2013. Citricultura. Editor Talleres Gráficos ALFA BETA. S. A. Argentina. ISBN: 9789-8743-266. 518 pp.
35. Paulino, S.E.P.; Mourão Filho, F.D.E.; Maia, A.N, Avilés, T.E.C.; Dourado Neto, D. 2007. Agrometeorological models for 'Valencia' and 'Hamlin sweet 'oranges to estimate the number fruits per plant. Scientia Agricola, 64 (1):1-17.
36. Pérez de Villareal, A.; Confalone, A.; Yommi, A.; Díaz, K. 2021. Modelos de regresión para predecir el crecimiento del fruto japones (*Prunus salicina*L. cv. Soledad). 41° Congreso de Horticultura. Horticultura Argentina 41 (104): Ene. - Abr. 2022. ISSN de la edición on line 1851-9342. <http://id.caicyt.gov.ar/ark:/s18519342/krpyakbek>
37. Reeb, P. & Tassile, V. 2007. Estudios Estadísticos. Pronósticos de Producción. (<https://docplayer.es/42642541-Pronosticos-de-produccion.html> )
38. Reginato, G.; García de Cortázar, V.; Robinson, T. 2007. Predicted crop value for nectarines and cling peaches of different harvest season as a function of crop load. American Society for Horticultural Science 42(2):239-245.

39. Stangaferro, S., Álvarez, O., and Bramardi, S. (2001). Modelos de predicción anticipada del peso medio de los frutos en producción de manzanas. In - VELAEM, A., editor, X Congreso Latinoamericano de Biomatemática. ALAB - VELAEM.
40. Verreyne, S. 2009. Fruit size and crop load prediction for citrus. Citrus Research International, Department of Horticultural Science, University of Stellenbosch ([sv@cri.co.za](mailto:sv@cri.co.za)).
41. Teixidó, P.; Gómez-Galán, J.A.; Gómez-Bravo, F.; Sánchez-Rodríguez, T.; Alcina, J.; Aponte, J. 2018. Low-Power Low-Cost Wireless Flood Sensor for Smart Home Systems. *Sensors*. 18 (11):3817.
42. Wang, Y.; Yang, H.; Luo, L.; Hu, X.; Lan, H. 2021. Innovation of agricultural digital business model based on remote sensing image target detection and embedded network. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-021-03258-x>.
43. Winter, F. 1969. Méthode des densités de charge, modèle d'analyse et de prévision de la production de fruit à pépins. Luxembourg, Office Statique des Communautés Européennes. Série Etudes de Statique Agricole (5). 137 pp.
44. Zeeshan, M.; Sheikh, Z.; Ahmad, R.S.; Asad, S. 2016. Detection and counting of On-Tree Citrus Fruit for Crop Yield Estimation, *International Journal of Advanced Computer Science and Applications*. 7 (5): 519-523.
45. Zhao, C., Lee, W. S., and He, D. (2016). Immature green citrus detection based on colour feature and sum of absolute transformed difference (SATD) using colour images in the citrus grove. *Computer Electronic Agriculture*. 124:243-253.
46. Zon, K; Sepúlveda, G; Giménez, G; Dussi, MC. 2011. Fruit Growth and Cell Evolution in Williams pears. *Acta Horticulturae*, 909.

## CAPÍTULO II

### DESCRIPCIÓN DEL AREA Y MATERIAL DE ESTUDIO

#### 2.1. ÁREA DE ESTUDIO

Las principales zonas productivas de cítricos en el Nordeste Argentino se concentran fundamentalmente en las provincias de Corrientes y Entre Ríos. En esta tesis se trabajó con información relevada de huertos comerciales ubicados en los departamentos de Saladas, Bella Vista y Concepción, ubicados hacia el centro y oeste de la provincia de Corrientes, y el departamento de Concordia en el nordeste de la provincia de Entre Ríos (Figura 2.1).

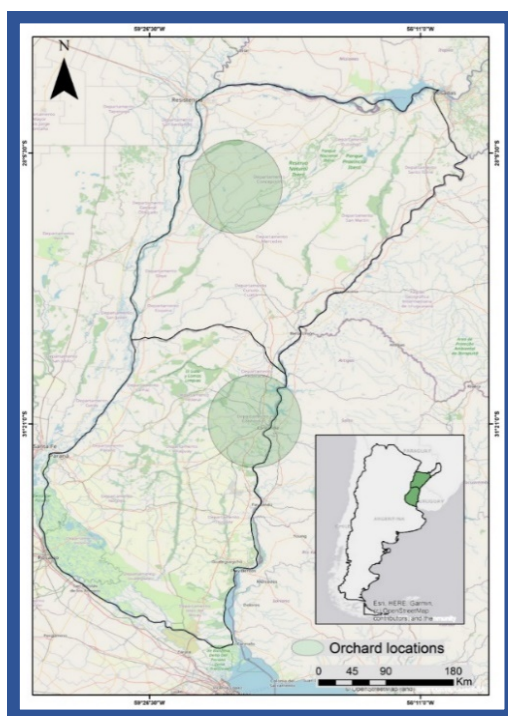


Figura 2.1. Localización geográfica de los huertos comerciales estudiados en las provincias de Corrientes y Entre Ríos, Argentina.

### 2.1.1. Provincia de Corrientes

La provincia de Corrientes cuenta con una superficie de 88.886 km<sup>2</sup> y forma parte de la Mesopotamia Argentina y, en forma más amplia, de la región denominada Litoral. El relieve de la región corresponde en su mayoría al de una dilatada llanura, rodeada por cursos de agua que le confieren un aspecto insular, con abundancia de canales interiores y grandes zonas lacustres. Limita al norte y al oeste con el río Paraná, que la separa del Paraguay y de las provincias de Chaco y Santa Fe; su linde este está marcado por el río Uruguay, que la separa de Brasil y Uruguay; los ríos Guayquiraró y Mocoretá y los arroyos Basualdo y Tunas marcan su límite sur con la provincia de Entre Ríos; al nordeste los arroyos Itaembé y Chimiray junto con un tramo de "límite seco" de 30 km, la separan de la provincia de Misiones (Castro, 1991).

Topográficamente la provincia puede considerarse dividida por el sistema Iberá y su prolongación en el río Corriente, en dos zonas: la occidental y la oriental. La zona occidental comprende un terreno llano con ondulaciones de arena rojiza (médanos consolidados) rodeadas de bañados, esteros y pantanos a los que concurren terrenos arcillosos y perfiles con escasa pendiente que retienen las aguas pluviales. Esta zona a su vez se subdivide en: a) Triángulo de la Capital, línea imaginaria que une Itá Ibaté con Bella Vista y Capital, sobre las terrazas del río Paraná; b) Cuenca del Iberá, una extensa superficie lacustre que se localiza en el centro y norte de la provincia, cuyo caudal responde a fenómenos periódicos de inundaciones y sequías; y c) Bajos del Río Corriente, donde se advierten cordones arenosos paralelos a la dirección de los principales cursos de agua, como los ríos Santa Lucía, Batel, Batelito y Corriente. Lomas y depresiones confieren al paisaje un relieve ondulado, con suelos planos donde se desarrollan extensos "malezales" (Castro, 1991).



### 2.1.2. Provincia de Entre Ríos

La provincia de Entre Ríos se ubica en el sur de la región mesopotámica, en el noreste de la Argentina. Su territorio abarca una superficie total de 78.347 km<sup>2</sup>, de los cuales el 20% corresponde a áreas de Delta, Pre-delta e Islas, 8,4% a cauces y valles inundables de los principales ríos y arroyos y la superficie restante, a tierra firme (Sasal *et al.*, 2015).

El paisaje general está delimitado por los ríos principales: Paraná y Uruguay. El río Gualaguay divide el territorio provincial en dos amplias regiones: la occidental donde la Cuchilla de Montiel constituye la principal divisoria de aguas y en la oriental la Cuchilla Grande (Sasal *et al.*, 2010).

## 2.2. SUELO Y CLIMA

### 2.2.1. Provincia de Corrientes

La provincia de Corrientes presenta gran variedad fitogeográfica, aunque muy modificada por la actividad humana, producto del contacto de distintas formaciones vegetales: sabanas con herbazales, especies hidrófilas en las áreas pantanosas, bosques en galería en las riberas de los ríos y superficies aisladas de bosques en medio de pastizales (Carnevali, 1994).

El clima de la región es subtropical húmedo, muy cálido en verano, pero con heladas en invierno. Presenta frecuentes excesos hídricos en otoño y primavera, y moderados y eventuales déficits principalmente en verano. La temperatura media anual varía entre 21º C al Norte y 19º C al Sur. Las temperaturas estivales e invernales relativamente elevadas y su variación anual definen al clima correntino. Las lluvias son abundantes y frecuentes, superando los 1.500 mm anuales en el NE descendiendo gradualmente hasta menos de 1.000 mm en el ángulo SO. La principal característica de este régimen es su irregularidad. La distribución anual de precipitaciones tiene dos máximos, en

primavera y en otoño, y un mínimo en invierno. El número de días con precipitaciones varía muy poco entre los meses del año; entre 6 y 8 al mes, que totalizan alrededor de 80 a 100 días de ocurrencia de lluvias por año (Carnevali, 1994; Escobar *et al.*, 1996).

La humedad relativa media anual oscila, en todas las localidades de la provincia, entre 70 y 75%, siendo mínimos los valores en verano, y máximos en invierno, según una variación inversamente proporcional a la temperatura. Esta elevada humedad promedio es el resultado de la enorme cantidad de cuerpos de agua que caracterizan el territorio provincial, no sólo por los ríos y arroyos, pero también por lagunas, esteros y cañadas de variada extensión (Escobar *et al.*, 1996).

En la región occidental, se observan suelos poco desarrollados, sobre sedimentos modernos y otros que han adquirido un fuerte desarrollo con la formación de horizontes A-B-C. En las condiciones climáticas actuales, los suelos están sometidos a lavado continuo; este proceso de lixiviación induce a la formación de suelos ácidos, migración de arcillas hacia horizontes subyacentes, con presencia de horizontes argílicos bien desarrollados, tanto en Molisoles como Alfisoles y para los suelos poco desarrollados se detectan lamelas texturales (B incipiente), en Entisoles con texturas arenosas (Escobar *et al.*, 1996).

Los huertos comerciales del departamento de Saladas se localizan en un paisaje de planicie suavemente ondulada conformando albardones, lomadas parcialmente disectadas y planos hidromórficos a la vera del río Paraná y afluentes. Se ubican en relieve normal (albardón), en loma a media loma alta, con pendientes de 1 a 3%. La vegetación está alterada por el uso agrícola, con preponderancia de *Setaria geniculata*, *Schizachyrium sp.*, *Paspalum notatum* y *Cynodon dactylon*. Son suelos medianamente fértiles, con mediano contenido de bases de cambio, con buen nivel de calcio, magnesio y potasio y relativamente moderados en contenido de materia orgánica. Presentan buena aptitud agrícola con algunas limitaciones por susceptibilidad a la erosión hídrica. Son suelos muy utilizados en citricultura, cultivos anuales, forestación, horticultura y como campo natural de pastoreo (Escobar *et al.*, 1996).

Los huertos comerciales del departamento de Bella Vista corresponden a suelos francos gruesos que presentan como limitante la susceptibilidad a la erosión hídrica, eólica y el drenaje deficiente. El paisaje es de una extensa planicie suavemente

ondulada, compuesta de lomadas con orientación noreste a suroeste, disectadas por caños de drenaje, con depresiones (paleocauces) e innumerables lagunas. Se localizan en relieve normal, posición de loma a media loma alta, con pendientes entre 1 y 1,5%. La vegetación más común, está compuesta por *Digitaria* sp., *Cynodon* sp. y *Sporobolus* sp., y está alterada por el uso agrícola (Escobar *et al.*, 1996).

Los huertos comerciales ubicados en el departamento de Concepción se ubican en relieve normal, en posición de loma con pendientes de 1 a 2%. El escurrimiento es medio, de permeabilidad moderadamente rápida y es bien drenado. En general son suelos de baja fertilidad natural y retención de humedad, además de ser susceptibles a la erosión hídrica y eólica, pero con excelentes condiciones físicas, sin ningún impedimento para la penetración del agua y de las raíces. Tienen bajos tenores de materia orgánica y bases de cambio, aunque con mejoras en profundidad (Escobar *et al.*, 1996).

### 2.2.2. Provincia de Entre Ríos

La provincia de Entre Ríos está comprendida dentro del clima templado húmedo de llanura. El área se caracteriza por su condición de planicie abierta, sin restricciones a la influencia de los vientos y con inundaciones periódicas como consecuencia del desborde de los ríos. El promedio anual de lluvias es de 1.000 mm y los valores de humedad relativa se mantienen elevados todo el año. La temperatura media anual es de 17° C y la amplitud térmica de 12,9° C, valores que señalan la influencia de la proximidad oceánica y la de grandes masas de aguas de la red pluvial (Pausich *et al.*, 2013).

Los suelos predominantes en la localidad de Concordia corresponden al orden Molisol, seguidos en cuanto a superficie ocupada por Vertisoles y Entisoles. Los suelos arenosos profundos de la costa del río Uruguay deben su origen a depósitos de este río. Los suelos mestizos se desarrollaron a partir de arenas eólicas del río Uruguay. Los Entisoles son profundos, de textura arenosa a arenosa franca, presentan bajo porcentaje de materia orgánica, de baja fertilidad, permeables y que sufren sequías. Se encuentran entre los suelos más aptos de la provincia para la implantación de citrus. Los molisoles

aparecen al alejarse del río Uruguay y pueden presentarse combinados con arena. Dentro de este grupo se encuentran los "gley subhúmicos" que son Molisoles muy húmedos típicos de los bañados de altura del norte entrerriano. Los Vertisoles aparecen cuando pierden importancia la influencia de las arenas. Muestran microrrelieve y buen contenido de minerales como fuente de nutrientes y alto porcentaje de materia orgánica (Pausich *et al.*, 2013).

## 2.3. MATERIAL DE ESTUDIO

### 2.3.1. Mandarinas

Se conoce como mandarinas a los frutos de algunas especies del género *Citrus* (*C. reticulata*, Blanco, *C. unshiu* Marcovitch, *C. nobilis* Loureiro, *C. deliciosa* Tenore, *C. Clementina* Hort. ex Tanaka, *C. reshni* Hort. ex Tanaka, *C. sunki* Hort. ex Tanaka), como así también sus híbridos.

Los híbridos de mandarino tienen un alto grado de adaptación a un rango de condiciones climáticas, desde desérticas hasta subtropicales. Son muy específicos en cuanto a requerimientos climáticos para producir frutas de calidad. Se caracterizan además por ser muy apreciados por los consumidores de cítricos, especialmente por su sabor dulce y facilidad de pelado. Los tangores son híbridos entre mandarino (*Citrus reticulata*) y naranjo dulce (*C. sinensis*), su nombre se forma por las primeras letras de sus parentales en su nombre inglés *tang* (tangerine) y *or* (orange) (Palacios, 2013).

En esta tesis se trabajó con tangor Murcott, cuya autoría se le atribuye a Charles Murcott Smith en Florida, presenta árboles vigorosos y abiertos, con amplia ramificación y pequeñas espinas en sus ramas. Se trata de una mandarina tardía, sensible al rajado de los frutos al final del verano y a la enfermedad fúngica denominada Alternaria, requiere un manejo cuidadoso tras la recolección para evitar, entre otras cuestiones que se reseque la corteza (Garavello *et al.*, 2019 a).

Los frutos son de tamaño medio y presentan una piel de color naranja rojizo, suave que puede ser pelada de forma relativamente fácil. Su pulpa es también de un color naranja intenso y ofrece un jugo de gran calidad y muy buen sabor, en el que el nivel de acidez y azúcares está equilibrado. La maduración se produce en los meses de junio-julio y se comercializa después que otras variedades (Satsuma, Ortanique, entre otras), dado que puede mantenerse en planta hasta diciembre. Uno de los principales inconvenientes que presenta es la sensibilidad al frío, que afecta a la fruta que se produce en el extremo de las ramas (Palacios, 2013).

### 2.3.2. Naranjas

Se denomina naranjas a las frutas cítricas obtenidas del naranjo dulce (*C. sinensis* L), del naranjo amargo (*C. aurantium*) y de algunos híbridos de estas especies. Son hesperidios carnosos de cáscara más o menos gruesa y endurecida, y su pulpa está formada típicamente por once gajos u hollejos llenos de jugo, con altos contenidos de vitamina C, flavonoides y aceites esenciales.

Se cultivan desde antiguo como árbol ornamental y para obtener fragancias de sus frutos, que son más pequeños y dulces que los pomelos o toronjas y más grandes, aunque menos perfumadas, que las mandarinas. Existen numerosas variedades de naranjas, la mayoría híbridos producidos con las especies *C. maxima* (pamplemusa), *C. reticulata* (mandarina) y *C. medica* (cidro) (Palacios, 2013).

Entre las naranjas dulces, en esta tesis se seleccionaron dos variedades, Valencia late por ser la naranja dulce de mayor producción en la región y Salustiana por el avance que ha mostrado en los últimos años, presentándose como promisorias en un futuro cercano.

#### 2.3.2.1. Variedad Valencia late

La naranja Valencia late (Valencia tardía), si bien tiene su origen en la California del siglo XIX, es una de las variedades dulces más empleadas en España para la extracción

de jugo. Debido a que su cosecha es posterior a otras naranjas dulces, su consumo es popular luego de pasada la época de las denominadas de ombligo. Pese a su nombre, esta variedad no guarda ninguna relación de origen con la Comunidad Valenciana, si bien es una de las variedades que más se cultivan en ella, junto a la Navel y la Navelina. Tiene un tamaño pequeño característico, y una gran cantidad de zumo que en ocasiones puede ser ligeramente ácido. Forma parte de las naranjas de la variedad blanca, llamada así porque su color no es tan intenso (Garavello *et al.*, 2019 b).

#### *2.3.2.2. Variedad Salustiana*

Salustiana es una variedad de las naranjas denominadas Navel y se produjo por una mutación espontánea en la provincia de Valencia debido al clima y condiciones de cultivo. Es una naranja típicamente española.

Al igual que las naranjas Valencia, tiene un color naranja claro, no el naranja rojizo que caracteriza a las Navelinas. Su forma es bastante achatada, su piel más fina que otras variedades y su tamaño suele ser mediano. Su punto fuerte es la gran cantidad de jugo que presenta lo que, unido a su intenso sabor dulce, las hace perfectas para consumir en forma de zumo. Los frutos de Salustiana son muy sensibles al frío (Garavello *et al.*, 2019 b).

### **2.3.3. Los huertos cítricos**

#### *2.3.3.1. Caracterización general de los huertos*

La producción cítrica puede identificarse como intensiva, dado que se aplican técnicas para maximizar la producción y, generalmente, los productores disponen de

superficies pequeñas. El Censo Citrícola de la Provincia de Corrientes<sup>8</sup> permitió determinar, que el Departamento de Monte Caseros cuenta con 741 establecimientos citrícolas de un tamaño promedio de 32 has, mientras que en el resto de la provincia se relevaron 618 establecimientos, con un tamaño promedio de 18 has. La edad media de las plantas se estableció entre 14 y 16 años. En la provincia de Corrientes el 48% de las quintas tienen un tamaño menor a 10 has y un 43% se encuentra entre 10 y 50 has.

En la provincia de Entre Ríos se realizó un Censo Citrícola en el año 2016<sup>9</sup>, del cual se pudo determinar la distribución de tamaños de los establecimientos. Para el Departamento de Federación, el 59% de los establecimientos cuenta con menos de 15 has, el 20% entre 15 y 25 has, el 15% entre 25 y 50 has y solamente el 6% más de 50 has. En el Departamento de Concordia el 47% de los establecimientos cuenta con menos de 15 has, el 15% entre 15 y 25 has, el 22% entre 25 y 50 has y solamente el 16% más de 50 has.

Con la finalidad de contribuir a un mejor control de enfermedades y plagas y a los efectos de la certificación de huertos comerciales para la exportación de fruta a mercados con restricciones cuarentenarias por parte del Servicio Nacional de Sanidad y Calidad Agroalimentaria (SENASA), los establecimientos citrícolas dividen su superficie en huertos (o lotes) de no más de 4 has, rodeados generalmente por cortinas rompevientos formadas por especies arbóreas de rápido crecimiento (principalmente de varias especies de *Pinus* y *Eucalyptus*).

#### 2.4.2. Huertos bajo estudio

Los datos experimentales se obtuvieron en diferentes temporadas de huertos comerciales de tangor Murcott (*C. reticulata* × *C. sinensis*) y naranjo dulce (*C. sinensis*), ubicados en las provincias de Corrientes y Entre Ríos, Argentina. Los árboles recibieron las prácticas de poda, fertilización y tratamientos sanitarios habituales de la zona. En las

---

<sup>8</sup> Censo Citrícola de la Provincia de Corrientes, año 2012.

<sup>9</sup> Censo Citrícola de la Provincia de Entre Ríos, año 2016.

Tablas 2.1. y 2.2., se describen las temporadas estudiadas y las características de los huertos donde se obtuvo la información empírica.

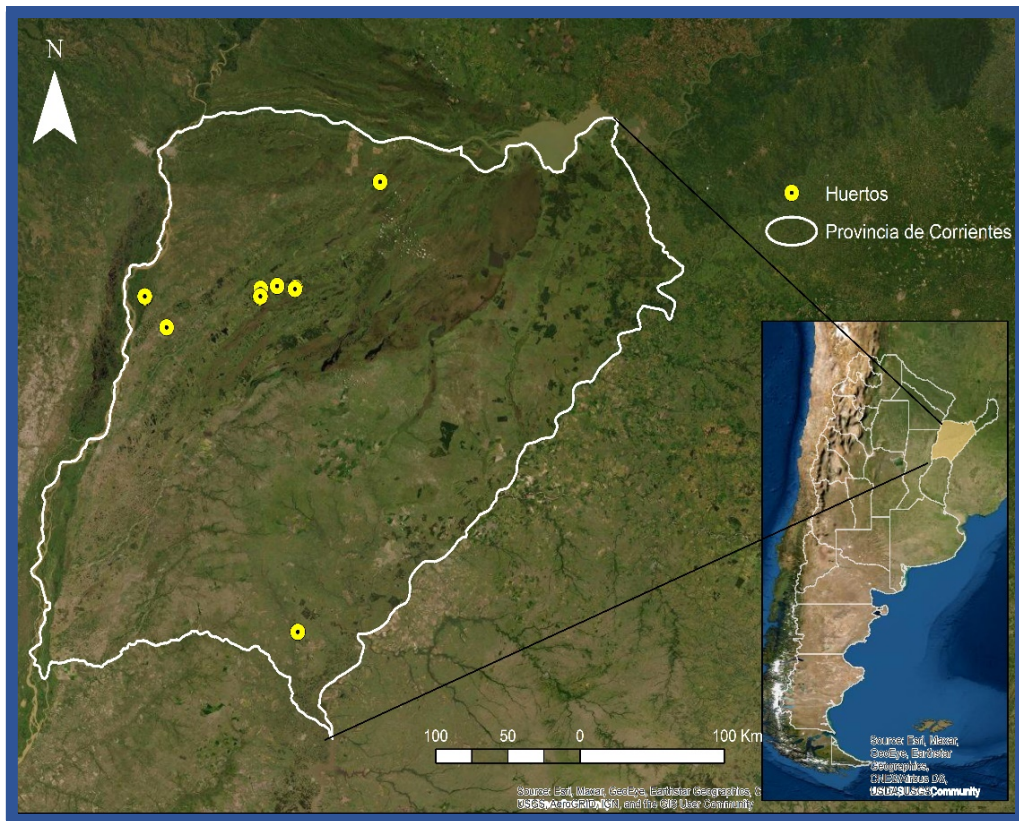
Tabla 2.1. Temporadas en las que se relevó información y características de los huertos de tangor Murcott (*C. sinensis* x *C. reticulata*) estudiados

Temporada	Huerto	Provincia y departamento	Portainjerto	Marco plantación	N° árboles	Edad	Riego	Suelo
2010/11	1.1	Corrientes Concepción	<i>Poncirus trifoliata</i>	7 x 3,5	582	8	Si	Arenoso
2014/15	1.2	Corrientes Concepción	<i>Poncirus trifoliata</i>	7 x 3,5	361	15	Si	Arenoso
2014/15	1.3	Corrientes Concepción	<i>Poncirus trifoliata</i>	7 x 3,5	361	15	Si	Arenoso
2015/16	1.4	Corrientes Concepción	<i>Citrus maxima</i>	6 x 5	1089	15	Si	Arenoso
2015/16	1.5	Corrientes Concepción	<i>Poncirus trifoliata</i>	6 x 2	1304	11	Si	Arenoso
2010/11 2011/12 2013/14	2.1	Corrientes Saladas	<i>Poncirus trifoliata</i>	7 x 3,5	733	7	Si	Arcilloso
2010/11 2013/14 2014/15	3.1	Corrientes Bella Vista	<i>Poncirus trifoliata</i>	6 x 3	815	18	Si	Arenoso
2015/16	3.2	Corrientes Bella Vista	<i>Poncirus trifoliata</i>	7 x 3,5	949	7	Si	Arenoso
2011/12	4.1	Corrientes Concepción	<i>Poncirus trifoliata</i>	6 x 2	733	10	Si	Arenoso
2007/08 2008/09 2009/10	5.1	Entre Ríos Concordia	<i>Poncirus trifoliata</i>	7 x 3,5	583	11	Sí	Arenoso
2007/08 2008/09 2009/10	6.1	Entre Ríos Concordia	<i>Poncirus trifoliata</i>	6 x 3	357	12	No	Arenoso
2007/08 2008/09 2009/10	7.1	Entre Ríos Concordia	<i>Poncirus trifoliata</i>	6 x 4	1344	18	No	Arenoso
2007/08 2008/09 2009/10	8.1	Entre Ríos Concordia	<i>Poncirus trifoliata</i>	7 x 3	1070	30	Si	Arenoso

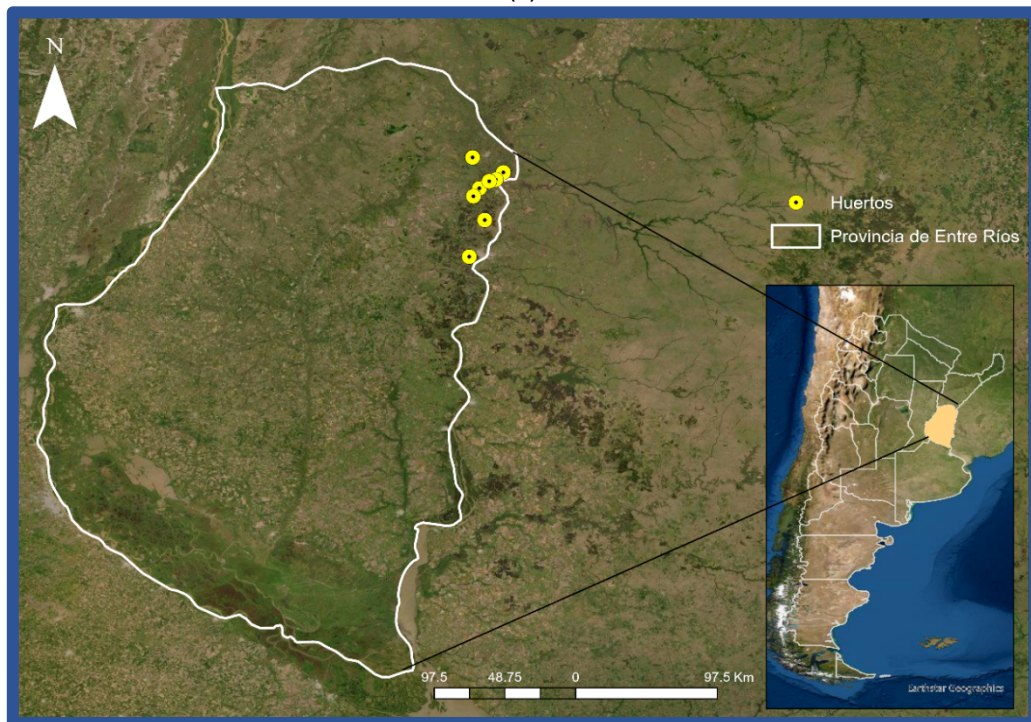


Tabla 2.2. Temporadas en las que se relevó información y características de los huertos naranjo dulce (*C. sinensis*) estudiados

Temporada	Huerto	Provincia y departamento	Variedad y portainjerto	Marco plantación	N° árboles	Edad	Riego	Suelo
2005/06	1.1	Corrientes Concepción	Valencia late <i>Citrus limonia</i>	4 x 7	357	12	No	Arenoso
2005/06	2.1	Corrientes General Paz	Valencia late <i>Citrus limonia</i>	4 x 8	312	15	No	Arenoso
2006/07 2007/08 2008/09 2009/10	3.1	Entre Ríos Concordia	Valencia late <i>Poncirus trifoliata</i>	6 x 4	644	16	Sí	Arenoso
2007/08 2008/09 2009/10	4.1	Entre Ríos Concordia	Valencia Late <i>Poncirus trifoliata</i>	6 x 4	461	23	Sí	Arenoso
2006/07 2007/08 2008/09 2009/10	5.1	Entre Ríos Concordia	Valencia Late <i>Poncirus trifoliata</i>	7 x 5	312	30	Sí	Arenoso
2006/07 2007/08 2008/09 2009/10	6.1	Entre Ríos Concordia	Valencia Late <i>Poncirus trifoliata</i>	8 x 3,5	918	27	No	Arcilloso
2006/07 2007/08 2008/09 2009/10	6.2	Entre Ríos Concordia	Valencia Late <i>Poncirus trifoliata</i>	7 x 3,5	940	17	No	Arcilloso
2006/07 2007/08 2008/09 2009/10	7.1	Entre Ríos Concordia	Valencia Late <i>Poncirus trifoliata</i>	6 x 3	1062	10	Si	Arenoso
2013/14	8.1	Corrientes Concepción	Salustiana <i>Citrus limonia</i>	7 x 5,5	1354	5	No	Arenoso
2013/14	9.1	Corrientes Concepción	Salustiana <i>Citrus limonia</i>	7 x 5,5	1743	5	No	Arenoso



(a)



(b)

Figura 2.2. Localización de los huertos estudiados en la provincia de Corrientes(a) y en la provincia de Entre Ríos (b)

## 2.5. BIBLIOGRAFÍA

1. Carnevali, R. 1994. Fitogeografía de la provincia de Corrientes. Editores: INTA-Gobierno de la provincia de Corrientes. 342 pp.
2. Castro, G. 1991. Caracterización Agroclimática. De la Provincia de Corrientes. Consejo Federal de Inversiones. Tomo 1. 1° Etapa. 90 pp.
3. Escobar, E.H.; Ligier, H.D.; Melgar, R.; Matteio, H.; Vallejos, O. 1996. Mapa de suelos de la provincia de Corrientes 1:500.000. Área de producción Vegetal y Recursos Naturales. E.E.A. INTA-Corrientes, 315 pp.
4. Garavello, M.; Beltran, V.; Kornowki, M. 2019 b. Catálogo de Variedades de Mandarinas. Colección Investigación, Desarrollo e Innovación. INTA Ediciones. [E.E.A. Concordia](#), [C.R. Entre Ríos](#), [E.E.A. Bella Vista](#), [E.E.A. Montecarlo](#). Ministerio de Agroindustria. Presidencia de la Nación. 7 pp.
5. Garavello, M.; Beltran, V.; Kornowki, M. 2019 a. Catálogo de Variedades de Naranjas. Colección Investigación, Desarrollo e Innovación. INTA Ediciones. [E.E.A. Concordia](#), [C.R. Entre Ríos](#), [E.E.A. Bella Vista](#), [E.E.A. Montecarlo](#). Ministerio de Agroindustria. Presidencia de la Nación. 17 pp.
6. Palacios, J. 2013. Citricultura. Editor Talleres Gráficos ALFA BETA. S. A. Argentina. ISBN: 9789-8743-266. 518 pp.
7. Pausich, G.; Gvozdzenovich, J.; Pioto, C. 2013. Caracterización de zonas y subzonas Entre Ríos. Red de Información Agropecuaria Nacional. Proyecto Agroecología. EEA Paraná-Centro Regional Entre Ríos. 29 pp.
8. Sasal M.C.; Castiglioni, MG.; Wilson, M.G. 2010. Effect of crop sequences on soil properties and runoff on natural rainfall erosion plots under no tillage. Soil Till Res. 108: 24-29.

## CAPITULO III

### MODELIZACIÓN DEL CRECIMIENTO DE LOS FRUTOS

#### 3.1. INTRODUCCIÓN

La modelización del crecimiento de los frutos se realiza mediante las denominadas curvas de crecimiento, que definen una función matemática para describir la evolución del tamaño (o peso) de los frutos a través del tiempo, y permiten su estimación en cualquier momento dentro del lapso estudiado. Varias funciones matemáticas han sido propuestas para modelar dichas curvas de crecimiento, que generalmente responden a modelos de tipo sigmoideo o doble sigmoideo, según la especie de que se trate.

#### 3.1.2. Crecimiento y desarrollo de los frutos

Se define crecimiento como el aumento irreversible del volumen de una célula, tejido, órgano o individuo, acompañado de un aumento de masa. Este fenómeno se produce por una secuencia de tres fases sucesivas: división, expansión y diferenciación celular (Taiz & Zeiger, 2004). El desarrollo vegetal se entiende como la sucesión progresiva de etapas que conducen a establecer la morfología propia del organismo adulto, a través de una serie de cambios cualitativos que se dan en diferentes fases fenológicas, y por los cuales pasa un organismo durante su ciclo ontogénico (Andrade *et al.*, 1996).

El crecimiento y desarrollo de los frutos cítricos ocurre en un período de tiempo relativamente largo que, en regiones subtropicales y templadas, comienza en la primavera y se completa en otoño (para el hemisferio sur de marzo a mayo) en las variedades tempranas, en invierno (junio a agosto) en las variedades intermedias, y

hasta finales de primavera (septiembre a noviembre) en las más tardías (Agustí *et al.*, 2020).

El crecimiento de los frutos puede verse afectado por diversas causales externas, pero también por factores intrínsecos de la propia planta.

La disponibilidad de energía y de agua en una región determina su potencial de productividad agrícola. La energía radiante, la temperatura y la humedad afectan el crecimiento y desarrollo de los cultivos. La duración de las fases del ciclo de desarrollo de los vegetales es condicionada por la temperatura y por el tiempo que ella permanece dentro de límites específicos (Pereira *et al.*, 2002).

El desarrollo de los frutos depende de factores ambientales y de las prácticas culturales. La acumulación de metabolitos en el fruto y su tasa de crecimiento están directamente asociadas a la temperatura. El crecimiento y desarrollo de las plantas acontecen dentro de un marco térmico definido (temperaturas cardinales): un nivel mínimo, debajo del cual no hay actividad; un rango óptimo en el cual el crecimiento y desarrollo alcanzan su mayor ritmo; y un nivel máximo sobre el cual cesa la actividad funcional. Estas temperaturas cardinales varían entre las especies, cultivares e incluso entre las fases ontogénicas (Andrade *et al.*, 1996).

Los frutos son susceptibles a las altas temperaturas, diversos estudios demostraron que su crecimiento disminuye si éstas superan los 30° C; atribuyéndose tal efecto a la reducción de la actividad fotosintética, lo cual mengua el aporte de carbohidratos al fruto en desarrollo. Por otro lado, durante el período de maduración, la ausencia de crecimiento del fruto se halla asociada también a bajas temperaturas (Agustí *et al.*, 2020). El rango de temperaturas óptimas para el desarrollo del cultivo de cítricos se encuentra entre 23° C y 34° C, con un máximo de 39° C y un mínimo de 13° C (Pimentel *et al.*, 2007).

Los cítricos presentan cierta tolerancia a la falta de luminosidad, sin embargo, el crecimiento de los frutos situados en el interior del árbol es más lento que el de los frutos ubicados en el exterior. El desarrollo del fruto también depende de las características del suelo, del aporte de agua por riego, de la provisión de nutrientes y del tipo de pie de injerto (Agustí *et al.*, 2003).

Existen factores internos de la planta que determinan el tamaño final de los frutos, como ser el componente genético, ya que el tamaño del fruto es una característica propia de cada variedad, aunque puede variar según la edad de la planta, condiciones climáticas, de manejo y fundamentalmente la disponibilidad de nutrientes en la planta (Agustí *et al.*, 2003). En relación con dicha disponibilidad, algunas técnicas para aumentar el tamaño de los frutos son el rayado de ramas y la aplicación de auxinas de síntesis al final de la caída fisiológica de frutos (Rebolledo-Roa, 2017).

El tamaño del fruto está inversamente relacionado al número de frutos por árbol, y más aún con el número de flores producidas, debido a la competencia entre órganos en desarrollo por productos de la fotosíntesis y elementos minerales (Gravina, 1999). Los mandarinos tienden a presentar grandes floraciones y altas cargas, lo que lleva a obtener frutos de pequeño calibre, razón por la cual el raleo o aclareo de frutos es una práctica común en este cultivo.

### 3.1.3. Maduración de los frutos

La maduración de los frutos se define como el conjunto de cambios externos e internos ocurridos cuando el fruto alcanza su máximo tamaño y completa su desarrollo, tales como coloración, firmeza, sabor, textura, aumento de azúcares solubles, descenso de almidón y de ácidos libres, entre otros. A partir de entonces comienza el proceso de senescencia y abscisión del fruto. En los cítricos la maduración interna y externa son procesos fisiológicos diferentes (Iglesias *et al.*, 2007).

El proceso de maduración externa en cítricos se refiere a la conversión de cloroplastos en cromoplastos que consiste en la progresiva pérdida de clorofila y el aumento en el contenido de carotenoides, lo que se manifiesta en el cambio de color de la cáscara al pasar del color verde al naranja. Este proceso es reversible y está influenciado por las condiciones ambientales, la disponibilidad de nutrientes y hormonas (Palacios, 2013).

Las condiciones ambientales tales como temperatura, luz y humedad son factores determinantes en la coloración. Temperaturas por debajo de 13° C provocan el cambio

de coloración del fruto (por degradación de clorofilas y síntesis de carotenoides). La luz afecta la coloración del fruto, ya que es necesaria para la síntesis de carotenoides y antocianinas. Los frutos situados en el exterior de la copa tienen una coloración más intensa.

La humedad relativa afecta el grosor de la piel, especialmente en las últimas fases del desarrollo del fruto. Así pues, elevados valores de ésta, en combinación con temperaturas altas en otoño favorecerán a la incidencia del bufado (hinchamiento de la corteza que la separa de la pulpa). La radiación también afecta la calidad de la fruta al influir sobre su coloración y firmeza, provocando daños en la piel de los frutos cuando su intensidad supera ciertos límites, como el denominado golpe de sol en mandarina Satsuma, que consiste en un aplanamiento asimétrico de la zona del fruto expuesta al sol, en la que la corteza adquiere una textura reseca y dura y una coloración amarillenta, desprovista de clorofila, las glándulas de aceites esenciales se rompen y las células se necrosan (Gariglio *et al.*, 2012). El viento produce lesiones sobre la corteza como consecuencia de la rozadura de ramas, hojas o tallos. También puede provocar desórdenes fisiológicos cuando es de elevada intensidad (Agustí *et al.*, 2003).

Los cítricos son frutos no climatéricos, por lo cual no muestran aumentos respiratorios o de síntesis de etileno durante la maduración. En esta etapa los cambios de composición química como la hidrólisis de polisacáridos y componentes de la pared celular, de ácidos orgánicos, aroma y pigmentación, se dan más gradualmente que en los frutos climatéricos (Sozzi, 2007). Durante la maduración hay una disminución de la acidez debido al catabolismo del ácido cítrico y el aumento en los azúcares. Entre el 75 y el 85 % de los sólidos solubles totales del jugo de los cítricos son azúcares (sacarosa, glucosa y fructosa en proporción 2:1:1) y el resto son ácidos orgánicos, proteínas, lípidos y minerales. La acidez se debe principalmente al ácido cítrico, otros ácidos orgánicos presentes en el jugo son los ácidos málico, succínico e isocítrico.

La relación entre sólidos solubles totales y acidez constituye el índice de madurez y es la base para determinar la madurez comercial. La acidez es determinante del pH del jugo, el cual varía entre 2 para limones y 5 para mandarinas y naranjas. La alternancia de temperaturas afecta la calidad interna, ante un alto régimen térmico día/noche los frutos poseen baja acidez, atribuida a la respiración de los ácidos orgánicos. Ambos

parámetros son reducidos cuando en el periodo previo a la cosecha se presentan lluvias (Agustí *et al.*, 2003).

#### 3.1.4. Estimación del tamaño de los frutos a la cosecha

Estimar la producción de manera temprana permite a los fruticultores definir estrategias de manejo de cultivos y planificar aspectos logísticos para su comercialización, facilitando la mejora de la competitividad en los mercados nacionales e internacionales (Bergh, 1982). Para predecir la producción frutícola se requiere la definición de la cantidad de fruta a obtener, pero también de su tamaño. Ajustar un modelo al crecimiento de la fruta permite encontrar la función adecuada que sintetiza la información obtenida de las mediciones tomadas de un solo individuo, favoreciendo la predicción del tamaño de fruta en cualquier momento antes de la cosecha, y estudiar la tasa de crecimiento a lo largo del proceso (Avanza *et al.*, 2008). Las curvas de crecimiento describen la evolución en el tiempo de órganos o individuos vegetales o animales y conocerlas permite estimar el tamaño de los frutos en la cosecha.

Las curvas que representan el crecimiento de los frutos varían entre las especies, el desarrollo de los cítricos en general sigue una curva de crecimiento sigmoidea simple caracterizada por tres períodos bien diferenciados.

Los procesos de crecimiento que generan curvas sigmoideas son bastante frecuentes en biología, agricultura, ingeniería y economía. Tales curvas comienzan en un punto fijo y aumentan su tasa de crecimiento monótonamente hasta que alcanzan un punto de inflexión, después de esto, la tasa de crecimiento disminuye y la función se acerca asintóticamente a un valor final (Ratkowsky, 1983). Los modelos apropiados para su descripción son modelos no lineales, que también tienen la ventaja de que sus parámetros tienen una interpretación práctica directa, en consecuencia, son ampliamente utilizados en estudios de crecimiento de frutos de diferentes especies (Carneiro *et al.*, 2014; Fernandez, 2014).



### 3.1.4.1. Regresión no lineal

La técnica estadística de regresión permite investigar y modelar la relación entre variables. Los modelos de regresión lineal proporcionan un marco rico y flexible que se adapta a los requisitos de muchos analistas; sin embargo, no son apropiados en todas las situaciones. Varios problemas surgen cuando la respuesta y las variables predictoras están relacionadas a través de una función no lineal. Cuando se aplica el método de los mínimos cuadrados a esos modelos, las ecuaciones normales resultantes son no lineales, por lo tanto, estimar en forma exacta los parámetros se hace difícil o imposible (Montgomery, 2004).

Se dice que un modelo de regresión es lineal, cuando lo es respecto a sus parámetros. Estos modelos pueden representarse matemáticamente de la siguiente manera:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_{p-1} x_{p-1,i} + \varepsilon_i$$

Los modelos de regresión lineal no necesariamente deben incluir a las p-1 variables independientes en grado uno, sino que éstas pueden estar elevadas a distintas potencias, e incluirse sus interacciones. Estos modelos continúan siendo lineales en los parámetros:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{1i}^2 + \beta_3 x_{2i} + \beta_4 x_{1i} x_{2i} + \varepsilon_i$$

A la gran cantidad de modelos que son lineales en sus parámetros se suman todos aquellos no lineales en los parámetros, pero que mediante una transformación pueden aparecer como lineales, por ejemplo, el modelo de regresión exponencial, el cuál es linealizable mediante la aplicación de logaritmos a ambos miembros:

$$y_i = \gamma_0 \cdot e^{\gamma_1 x_i} \cdot \varepsilon_i$$

Hay muchas situaciones en las que puede no ser adecuado un modelo lineal de regresión, por ejemplo, cuando se tiene el conocimiento directo de la forma de la relación entre la variable respuesta y la regresora, quizás con base en la teoría que gobierna los fenómenos (Montgomery, 2004). Es el caso del proceso de crecimiento de frutos de naranjo que responde a una función sigmoidea en forma de S, no lineal en sus parámetros (Agustí, 2000).

Los modelos de regresión no lineal son aquellos que son no lineales en sus parámetros, ni tampoco pueden serlo por transformación. Por ejemplo, el modelo de regresión exponencial con término de error aditivo:

$$y_i = \theta_0 \cdot e^{\theta_1 x_i} + \varepsilon_i \quad y_i = \theta_0 + e^{\theta_1 x_i} + \varepsilon_i$$

Se utiliza el símbolo  $\theta$  para representar un parámetro en un modelo no lineal, para mostrar la diferencia con el caso lineal.

En general, el modelo de regresión no lineal se escribe de la siguiente forma:

$$y_i = f(x, \theta) + \varepsilon$$

donde  $\theta$  es el vector de  $p \times 1$ , de parámetros desconocidos, y  $\varepsilon$  es un error aleatorio no correlacionado, cuyo  $E(\varepsilon) = 0$  y cuya  $Var(\varepsilon) = \sigma^2$ . También se supone que los errores tienen distribución normal, como en regresión lineal.

El procedimiento más objetivo para ajustar una función a un conjunto de datos se conoce como el método de los mínimos cuadrados, el cual consiste en una técnica de análisis numérico encuadrada dentro de la optimización matemática, en la que, se intenta encontrar la función que mejor se aproxime a los datos, de acuerdo con el criterio de mínimo error cuadrático. En su forma más simple, intenta minimizar la suma de cuadrados de las diferencias entre los puntos generados por la función y los correspondientes en los datos.

Este procedimiento presenta ciertas propiedades bajo modelos lineales (Draper & Smith, 1981):

- a)  $b$  es un estimador de  $\beta$  que minimiza la suma de cuadrados del error independiente de cualquier propiedad distribucional de los errores;
- b) los elementos de  $b$  son funciones lineales de las observaciones  $y_1, y_2, \dots, y_i$  y proveen estimadores insesgados de los elementos de  $\beta$  que tienen varianza mínima;
- c) si los errores se distribuyen  $\varepsilon_i \sim N(0, \sigma^2)$  entonces  $b$  es el estimador máximo verosímil de  $\beta$ .

Un punto importante es que los estimadores por mínimos cuadrados de un modelo de regresión no lineal no tienen las mismas propiedades que los estimadores correspondientes a un modelo lineal. Sólo asintóticamente, es decir, cuando el tamaño de la muestra aumenta a infinito, las propiedades de los estimadores de modelos no lineales se aproximan a las de los estimadores de modelos de regresión lineal (Draper & Smith, 1981).

Cuando la muestra es finita y aunque las  $y_i$  se distribuyan normalmente, con varianza finita y desconocida (es decir, que un estimador mínimo-cuadrático es también máximo-verosímil), como  $\hat{\theta}$  no es combinación lineal de las  $Y_i$ , en general, no tendrá distribución normal. Entonces  $\hat{\theta}$  no será un estimador insesgado de  $\theta$  y podrá no presentar varianza mínima. Por todo ello, un estimador por máxima verosimilitud, en un modelo no lineal, tiene propiedades desconocidas para tamaños de muestras finitos. Sin embargo, cuando el tamaño de muestra se incrementa,  $\hat{\theta}$  tenderá a tomar un comportamiento que provea errores independientes e idénticamente distribuidos en forma normal (este es el caso en que los estimadores mínimos cuadrados coinciden con los de máxima verosimilitud). En general, no es posible presentar una idea de cuán grande debe ser el tamaño de la muestra para aproximarse a las propiedades asintóticas; en algunos modelos se alcanzarán dichas propiedades con tamaños de muestra pequeños y en otros modelos las propiedades asintóticas serán pobremente alcanzadas con grandes tamaños de muestra (Ratkowsky, 1983).

#### *3.1.4.1.1. Modelos de crecimiento sigmoidal*

Al tener las curvas de crecimiento de los frutos una forma sigmoidea, los modelos adecuados para su descripción son los modelos no lineales, que además tienen la ventaja de que sus parámetros tienen interpretación práctica directa. Así, los modelos no lineales son ampliamente utilizados en los estudios de crecimiento de frutos de diferentes especies (Carneiro *et al.*, 2014; Fernández *et al.*, 2014).

En el caso del crecimiento de los frutos, las curvas de crecimiento permiten estimar el tamaño que tendrán en el momento de la recolección. Estas curvas varían con la especie que se trate, en el caso de los cítricos, en general corresponde a una forma sigmoidea simple (Cañizares *et al.*, 2012; Rebolledo-Roa, 2017) que va desde la antesis hasta la maduración, caracterizada por tres períodos bien diferenciados (ver Figura 3.1):

- ✓ Fase I o de crecimiento exponencial: desde antesis hasta el final de la caída fisiológica de frutos (de aproximadamente dos meses de duración), caracterizada por un rápido crecimiento, básicamente de la corteza (exocarpio, mesocarpio y endocarpio), provocado por la división celular, donde se define el número de células del fruto y, por lo tanto, su tamaño potencial.
- ✓ Fase II o de crecimiento lineal: desde la caída fisiológica de frutos hasta antes del cambio de color, con duración más corta en variedades precoces (2-3 meses) o más larga en las tardías (5-6 meses). Se caracteriza por la expansión celular, desarrollo de lóculos y aumento del contenido de jugo en las vesículas.
- ✓ Fase III o de maduración: presenta una reducida tasa de crecimiento, la corteza adquiere pigmentación debido a la degradación de la clorofila y la síntesis de carotenoides, se produce la maduración interna con el aumento de sólidos solubles y de compuestos nitrogenados, a la vez que disminuyen los ácidos libres debido a su dilución y metabolización.

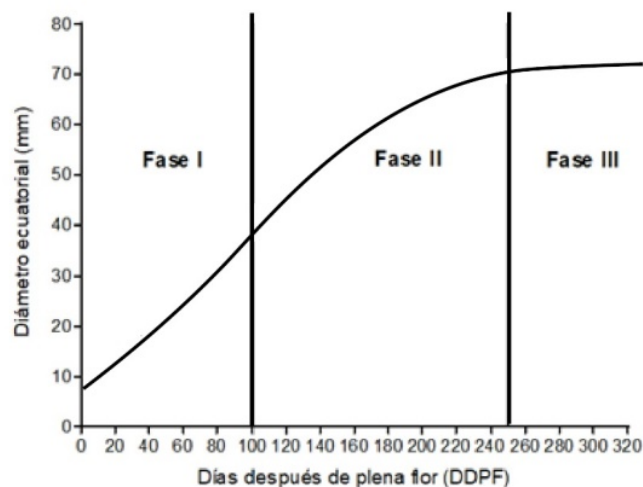


Figura 3.1. Fases de crecimiento de frutos cítricos (Avanza, 2011)

Se han propuesto varias funciones matemáticas para modelar curvas de crecimiento. Los más utilizados son Logístico, Gompertz, Richards (1959), Morgan-Mercer-Flodin (1975) (MMF) y los modelos derivados de la distribución de Weibull (Ratkowsky, 1983), los que se describen a continuación.

Tabla 3.1. Modelos de tipo sigmoideo utilizados para describir crecimiento de frutos

Modelo	Expresión Matemática
Logístico (L1)	$y = \frac{\alpha}{1 + \exp(\beta - \gamma x)}$
Gompertz (G1)	$y = \alpha \exp\left[-\exp(\beta - \gamma x)\right]$
Morgan-Mercer-Flodin (MMF)	$y = \frac{\beta\gamma + \alpha x^\delta}{\gamma + x^\delta}$
Richards	$y = \frac{\alpha}{\left[1 + \exp(\beta - \gamma x)\right]^{\frac{1}{\delta}}}$
Weibull	$y = \alpha - \beta \exp^{-\gamma x^\delta}$
Von Bertalanffy	$y = \alpha - \beta \gamma^x$
Brody	$y = \alpha + (1 - \beta * \exp^{(-\gamma * x)})$

$y$  = diámetro ecuatorial de los frutos;  $x$  = DDPF,  $\alpha$ ,  $\beta$ ,  $\gamma$ , parámetros

Se estudian también distintas reparametrizaciones de los modelos originales, algunas de las cuales se presentan a continuación.

Tabla 3.2. Reparametrizaciones de los modelos Logístico y Gompertz

Modelo	Expresión Matemática
Logístico	L2 $y = \frac{1}{\alpha + \beta \exp(-\gamma x)}$
	L3 $y = \frac{1}{\alpha + \beta \gamma^x}$
	L4 $y = \frac{\alpha}{1 + \exp(\beta) \gamma^x}$
	L5 $y = \frac{1}{\alpha + \exp(\beta) \gamma^x}$
	L6 $y = \frac{\alpha}{1 + \beta \exp(-\gamma * x)}$
	L7 $y = \frac{\alpha}{1 + \left(\frac{1}{\beta} - 1\right) * e^{-\gamma * x}}$
	Gomperzt
G3 $y = \alpha * e^{-e[\gamma * (\beta - x)]}$	
G4 $y = \alpha * e^{-\beta * e^{(-\gamma * x)}}$	

$y$  = diámetro ecuatorial de los frutos;  $x$  = DDPF;  $\alpha$ ,  $\beta$ ,  $\gamma$ , parámetros.  
 Nomenclatura de reparametrizaciones según Ratkowsky (1989)

En los distintos modelos, como así también en sus distintas parametrizaciones, se utilizaron las letras griegas  $\alpha$ ,  $\beta$ ,  $\gamma$  y  $\delta$  para representar a los parámetros de cada modelo. Sus interpretaciones corresponden a los siguientes aspectos de las curvas de crecimiento:

$\alpha$  : es el parámetro relacionado con la asíntota superior. En los modelos logístico original, Richards, MMF y Weibull el valor de  $\alpha$  es la asíntota superior, mientras que en la primer reparametrización del modelo logístico la asíntota superior será la inversa de  $\alpha$  , y en el de Gompertz la asíntota será  $e^\alpha$ .

$\beta$  : este parámetro está relacionado con la ordenada al origen de la función. En el modelo de MMF corresponde directamente al valor de la ordenada al origen, mientras

que en el de Weibull, indica la diferencia entre la asíntota superior y la ordenada al origen. En todos los casos, este parámetro determina la posición de la intercepción de la función con el eje de ordenadas.

$\gamma$  : este parámetro está relacionado con la velocidad de crecimiento de la curva, desde un valor inicial determinado por la magnitud de  $\beta$  hasta un valor final determinado por la magnitud de  $\alpha$  .

$\delta$  : este parámetro, sólo presente en los modelos de cuatro parámetros, proporciona una mayor flexibilidad al modelo para ajustar datos, en comparación con los modelos de tres parámetros.

Varios autores han utilizado diferentes parametrizaciones para describir el crecimiento de órganos y desarrollar curvas de crecimiento para los frutos (Amancio *et al.*, 2014; Andrade, 2014; Pereira, 2014; Prado *et al.*, 2013; Sousa *et al.*, 2014; Terra *et al.*, 2010).

Para describir el crecimiento de la fruta de gulupa (*Passiflora edulis*), Carvajal *et al.* (2012) encontraron que el modelo que mejor se ajustaba a los datos de crecimiento era el Logístico; en este modelo, la tasa de cambio en el diámetro del fruto está determinada por la magnitud del crecimiento y disminuye hasta que se vuelve casi nula, cuando el fruto alcanza su tamaño máximo.

Utilizando medidas de no linealidad, Bramardi (1995) seleccionó el modelo Logístico en la tercera parametrización definida por Ratkowsky (1989), para describir el crecimiento de peras William's y Packham's Triumph, con el objetivo de predecir el tamaño final de la fruta.

Cañizares *et al.* (2012) encontraron que el desarrollo de lima Tahití siguió un modelo sigmoideal desde la antesis hasta la maduración, caracterizada por tres fases: Fase I, desde la caída del estilo hasta el aumento de la masa fresca y seca y el diámetro del fruto, caracterizado por una ligera pendiente curva. Esta fase abarca de 7 a 42 días después de la floración completa (DDPF), el crecimiento es lento debido a la división celular que ocurre en casi todas las células que constituirán el fruto maduro. La fase II se caracteriza por un rápido crecimiento, las células se diferencian en los diversos tipos de tejido, se produce el alargamiento celular, y las vesículas se llenan de líquido, esta

fase abarca entre 49 y 196 DDPF. En fase III el tamaño de la fruta se ha reducido y el color verde en la corteza es menos intenso. Esta fase se caracteriza por una curva de pendiente leve continua y abarca de 203 a 245 DDPF.

Avanza *et al.* (2008) relatan que la combinación modelo-datos-parametrización más adecuada para describir el patrón de crecimiento del naranja tardío de Valencia es el modelo logístico de la quinta parametrización (Ratkowsky ,1983). El modelo se definió para tres tamaños de fruta, pequeña, mediana y grande, y determinó que la variación de tamaño dependía del huerto.

#### *3.1.4.1.2. Estimación de parámetros en modelos de regresión no-lineal*

Los parámetros de los modelos no lineales son, en general, difíciles de estimar. El método usualmente utilizado consiste en minimizar directamente la suma de los residuos al cuadrado a través de un procedimiento iterativo. En un modelo de regresión no lineal, al menos una de las derivadas de la función del valor esperado con respecto a los parámetros depende al menos de uno de sus parámetros. El uso de parametrizaciones (forma en la que aparece un parámetro en el modelo) es común en los modelos no lineales. Según Cordeiro (2009), muchas veces el modelo no se expresa en una forma paramétrica que facilite la rápida convergencia del proceso iterativo utilizado para estimar los parámetros, por lo que es necesario encontrar la reparametrización más adecuada.

Un método muy usado en los algoritmos de cómputo para regresión no lineal es la linealización de la función no lineal, seguida por el método iterativo del algoritmo de Levenberg-Marquardt (LMA), también conocido como el método de mínimos cuadrados amortiguados, se utiliza para resolver problemas de mínimos cuadrados no lineales. Estos problemas de minimización surgen especialmente en el ajuste de curvas de mínimos cuadrados. El LMA se usa en muchas aplicaciones de software para resolver problemas genéricos de ajuste de curvas. Sin embargo, como ocurre con muchos algoritmos de ajuste, el LMA solo encuentra un mínimo local, que no es necesariamente el mínimo global. El LMA interpola entre el algoritmo de Gauss-Newton (GNA) y el



método de descenso de gradiente. El LMA es más robusto que el GNA, lo que significa que en muchos casos encuentra una solución incluso si comienza muy lejos del mínimo final. Para funciones de buen comportamiento y parámetros de inicio razonables, el LMA tiende a ser un poco más lento que el GNA. El LMA también se puede ver como Gauss-Newton utilizando un enfoque de región de confianza.

El algoritmo de Levenberg-Marquardt es un procedimiento iterativo. Para iniciar una minimización, el usuario debe proporcionar una estimación inicial del vector de parámetros  $\beta$ . En los casos con solo un mínimo, una conjetura estándar no informada como  $\beta^T = (1, 1, \dots, 1)$  funcionará bien; en los casos con múltiples mínimos, el algoritmo converge al mínimo global solo si la estimación inicial ya es algo cercana a la solución final.

En cada paso de iteración, el vector parámetro  $\beta$  se reemplaza por una nueva estimación  $(\beta + \omega)$ . Para determinar  $\beta$ , la función  $f(x_i, (\beta + \omega))$  se aproxima por su linealización:

$$f(x_i, (\beta + \omega)) \approx f(x_i, \beta) + J_i \omega_i$$

Donde:

$$J_{i=} = \frac{\partial f(x_i, \beta)}{\partial \beta}$$

Es el gradiente (fila-vector en este caso) de  $f$  con respecto a  $\beta$

#### *3.1.4.1.3. Obtención de valores iniciales de los parámetros*

Para ajustar un modelo no lineal de regresión se requieren valores de arranque  $\theta_0$  de los parámetros del modelo. Esto es, valores de  $\theta_0$  cercanos a los valores verdaderos de los parámetros, minimizarán las dificultades de convergencia. Una mala elección podría causar la convergencia hacia un mínimo local de la función, y pasaría completamente desapercibido el que se haya obtenido una solución subóptima.

En los modelos no lineales de regresión, con frecuencia los parámetros tienen algún significado físico, lo cual ayuda mucho para obtener los valores de arranque. También se puede graficar la función de valor esperado para varios valores de los parámetros, para familiarizarse con el comportamiento del modelo, y la forma en que los cambios en los parámetros afectan ese comportamiento (Montgomery, 2004).

Para el caso de los modelos donde el valor de  $\alpha$  corresponde directamente a la asíntota superior, un estimador inicial razonable para este parámetro, indicado con  $\alpha^{(0)}$ , puede obtenerse visualmente cuando se grafican los datos (Ratkowsky, 1983, 1989).

Para obtener las estimaciones iniciales de  $\beta^{(0)}$  y  $\gamma^{(0)}$  se debe reordenar la ecuación del modelo elegido y aplicar sobre ella las transformaciones necesarias para que resulte en un modelo simple, lineal en los parámetros a estimar, de esta manera serán fácilmente hallados por el método clásico de aplicación de ecuaciones normales.

Existe una amplia variedad de métodos de obtención de buenos valores iniciales que son presentados por Ratkowsky (1983).

A continuación, se detalla una de las metodologías para obtener los valores iniciales de los parámetros de los modelos no lineales de tipo sigmoidales según Ratkowsky (1983):

**Modelo Gompertz:** Para obtener  $\alpha_{(0)}$  se grafican los datos en un sistema de coordenadas X-Y y se determina visualmente la asíntota superior. Para estimar los otros valores iniciales se aplica la siguiente transformación:

$$Z_0 = \log \left[ -\log \left( \frac{y}{\alpha_0} \right) \right] = \beta - \gamma x$$

A partir de dicha transformación se realiza el ajuste de regresión lineal simple y se obtienen  $\beta_0$  y  $\gamma_0$ .

**Modelo Logístico:** Considerando el modelo original:  $y = \frac{\alpha}{1 + \exp(\beta - \gamma X)}$ , para que el estimador mínimo cuadrado se aproxime a un comportamiento lineal, es suficiente representar los datos en un gráfico de Y versus X y obtener una estimación visual de asíntota superior; y luego, reordenando la función:

$$Z_0 = \log\left(\frac{\alpha_0}{y} - 1\right) = \beta - \gamma x$$

se obtienen las estimaciones de  $\beta_0$  y  $\gamma_0$  a partir de la regresión lineal de  $Z_0$  sobre  $x$ .

**Modelo Richards:** Considerando que la función del modelo:

$y = \frac{\alpha}{[1 + \exp(\beta - \gamma x)]^{1/\delta}}$  es considerado como una extensión del modelo logístico, se

utilizan los mismos procedimientos, pero suministrando un valor de  $\delta$  fijo. Luego, reordenando la función se obtiene:

$$Z = \log\left[\left(\frac{\alpha_0}{y}\right)^{\delta_0} - 1\right] = \beta - \gamma x$$

donde  $Z$  es lineal en  $\beta$  y  $\gamma$ , y es análoga a la expresión obtenida para el modelo logístico excepto por la inclusión de  $\delta$ . Las estimaciones de  $\alpha_0$  y  $\delta_0$  se pueden obtener visualmente a partir del gráfico de  $Y$  versus  $X$ , donde  $\alpha_0$  corresponde a la asíntota superior y  $\delta_0$  es el punto de inflexión  $(x_f, y_f)$ , donde la pendiente deja de incrementarse para comenzar a decrecer; matemáticamente es donde la derivada segunda es cero. Entonces la derivada segunda con respecto a  $x$ , resultante de la expresión igual a cero es:

$$x_f = \frac{\beta - \log \delta}{\gamma}$$

y luego, sustituyendo en la función del modelo se obtiene:  $y_f = \alpha(1 + \delta)^{-1/\delta}$ , de donde se despeja  $\delta_0$ .

**Modelo Morgan Mercier Flodin:** Como se ha sugerido anteriormente, se pueden estimar los valores iniciales visualmente a partir de un gráfico de  $Y$  versus  $X$ . De tal gráfico, se puede estimar  $\alpha_0$  (asíntota superior) y  $\beta_0$  que es la ordenada al origen (valor de  $y$  cuando  $x = 0$ ). Luego de determinar  $\alpha_0$  y  $\beta_0$ , de la derivada segunda del modelo respecto de  $x$  evaluada en el punto de inflexión se obtiene:

$$x_f = \left[ \frac{\gamma(\delta-1)}{\delta+1} \right]^{1/\delta} \quad y_f = \frac{\beta_0(\delta+1) + \alpha(\delta-1)}{2\delta}$$

de donde se pueden estimar  $\gamma_0$  y  $\delta_0$ .

**Modelo Weibull:** Como en los modelos anteriores, se estima a partir del gráfico  $Y$  versus  $X$ , el valor inicial  $\alpha_0$  y el de  $\beta_0$  como la diferencia entre  $\alpha_0$  y la ordenada al origen ( $\beta_0 = \alpha_0 - y_{\text{int}}$ ).

Luego a partir de la siguiente transformación:

$$Z = \log \left[ -\log \left( \frac{\alpha_0 - y}{\beta_0} \right) \right] = \log \gamma + \delta \log x$$

se aplica regresión lineal simple y se estiman los valores de  $\gamma_0$  y  $\delta_0$ .

Todos los procedimientos descritos están ajustados para las parametrizaciones iniciales de los modelos. Para encontrar estimadores iniciales para una reparametrización en particular de un modelo dado, se debe analizar qué relación existe entre los parámetros de la parametrización inicial con la reparametrización en estudio.

#### 3.1.4.1.4. Criterios de selección de modelos

Para medir la adecuación de un modelo a un conjunto de datos, en estas familias de curvas, un elemento para tener en cuenta son las medidas de no-linealidad. Los modelos de crecimiento planteados en este Capítulo corresponden a modelos no-lineales, es decir que no pueden ser linealizados, razón por la cual se proceden a procesos iterativos para la estimación de los parámetros funcionales. Sus estimadores no tendrán precisamente las propiedades de máxima verosimilitud, esto es ser insesgados, de mínima varianza y con distribución normal, alcanzándolas sólo de manera asintótica.

Las medidas de no linealidad fueron desarrolladas para medir el alejamiento de la linealidad en una combinación modelo-diseño experimental-parametrización.

Si se asume que la relación entre la respuesta ( $y$ ) y la variable independiente ( $x$ ) puede ser representada por la siguiente ecuación:

$$y_i = f(x_i, \theta) + \varepsilon_i$$

donde  $\theta$  es el conjunto de parámetros desconocidos y  $\varepsilon_i$  es el término aleatorio del error, los estimadores mínimos cuadráticos de  $\theta$  se encuentran minimizando la suma de cuadrados de residuales:

$$\phi(\theta) = \sum_{i=1}^n (y_i - f(x_i, \theta))^2 = \|y - f(X, \theta)\|^2$$

Esta suma de errores al cuadrado ( $\phi(\theta)$ ), geoméricamente representa la distancia entre el vector  $y$  de respuesta y  $f(X, \theta)$  en un espacio muestral  $n$ -dimensional, donde se utiliza un eje de coordenadas para cada observación. A medida que  $\theta$  va cambiando, el vector  $f(X, \theta)$  va determinando una superficie  $p$ -dimensional, llamada curva solución en el espacio muestral. En general, el vector  $y$  observado no será un punto sobre esta curva solución, y el estimador mínimo cuadrático  $\hat{\theta}$  es el valor del vector de parámetros tal que  $f(X, \hat{\theta})$  es el punto sobre la curva solución que más se acerca al vector  $y$  observado. En un modelo lineal la curva solución es lineal, siendo una recta en el caso de un parámetro, un plano para dos parámetros y un hiperplano para tres o más parámetros; y los puntos o líneas para dimensión superior a dos, llamadas líneas paramétricas, sobre la curva  $\hat{y}$  correspondientes a incrementos constantes de los parámetros están igualmente espaciados.

El grado de curvatura de la curva solución en la zona del parámetro estimado ha sido llamada por Beale (1960) y Bates & Watts (1980) *no-linealidad intrínseca (NI)* y queda definida para una combinación particular de conjuntos de datos y modelo especificado, el nombre intrínseca es adecuado puesto que no se altera por la parametrización.

Otro aspecto en que un modelo no-lineal difiere de los modelos lineales se refiere a los valores que toma  $\hat{y}$  sobre la curva de solución cuando el valor de los parámetros sufre incrementos iguales.

Cuando se ajusta el modelo  $E(y) = x^\theta$  al incrementar  $\theta$  en valores constantes, los incrementos de la curva solución no resultan igualmente espaciados como debería ser en un modelo lineal. El grado en que este espaciamiento es desigual en un modelo no-lineal define la segunda componente de un modelo no-lineal, denominada *efecto*

*parámetro (EP)* de la no-linealidad. Bates & Watts (1980) mostraron que la no linealidad de los modelos puede ser separada en dos componentes o medidas: *no linealidad intrínseca (NI)* asociada con la curvatura de la curva solución y, la *no linealidad efecto parámetro (EP)* asociada con el hecho de la proyección de las líneas paramétricas sobre el plano tangente a la solución.

Estas medidas explican, en forma cuantitativa, la adecuación de la aproximación lineal; una aproximación lineal inadecuada indicaría que son dudosos los resultados de la inferencia (Montgomery, 2004).

Para el cálculo de la no-linealidad, Bramardi (1995) desarrolló un programa en el procedimiento IML del SAS, que consiste en las siguientes subrutinas:

- I. Estimación de parámetros de los modelos de regresión no-lineal.
- II. Evaluación de las derivadas primeras y segundas del modelo respecto de los parámetros para cada medición de los frutos en el tiempo.
- III. Construcción de las matrices de derivadas primeras  $n \times p$  y de los arreglos correspondientes a las derivadas segundas  $n \times p \times p$ . En el cual en esta parte el SAS no permite trabajar con arreglos en tres dimensiones, se crearon las  $n$  fases  $p \times p$  que conforman el mismo, y se trabajó sobre ellas.
- IV. Descomposición QR de la matriz de derivadas primeras.
- V. Cálculo de los arreglos de aceleración
- VI. Procesos iterativos para encontrar la máxima curvatura correspondiente a efecto parámetro (EP) y no-linealidad intrínseca (NI).

Para determinar la magnitud e impacto de la no linealidad sobre una particular región de confianza, las medidas (NI y EP), deben ser comparadas con el valor crítico  $1/\sqrt{F}$ , donde  $F = F(p, v, \alpha)$  con  $v = n - p$ ,  $n$ = número de observaciones en el tiempo y  $p$ = número de parámetros (Ratkowsky, 1989).

#### *3.1.4.2. Modelos no lineales mixtos*

El análisis de efectos fijos es una de las técnicas más comunes de inferencia estadística, utilizada en los casos donde el investigador diseña los niveles de interés a

estudiar para cada factor. En contraste, los efectos aleatorios, o los efectos de la unidad experimental, son los factores cuya variación se debe a la adhesión de algunos de los grandes miembros de la población potencial. Cuando las observaciones múltiples se registran en la misma unidad, cualquier factor o coeficiente que puede variar de una unidad a otra se considera como aleatorio. Así que una media de una población es fija, porque todas las unidades de esa población tienen esto en común. Sin embargo, las unidades individuales pueden presentar desviaciones aleatorias de la media poblacional.

Un modelo mixto es el que incorpora efectos fijos y aleatorios. Las medidas repetidas a menudo se incluyen como efectos aleatorios, porque las mediciones se toman en la misma unidad experimental, y los efectos fijos se basan en que la mayoría de los investigadores están interesados en hacer inferencias sobre un grupo de tratamientos sobre la base de los sujetos de la muestra.

Las aplicaciones de modelos con supuestos simplistas para la componente del error no tienen en cuenta que en el caso de medidas longitudinales (repetidas en el tiempo sobre una misma unidad experimental) la estructura de la matriz de varianzas y covarianzas de los errores no se corresponde con tales asunciones. El hecho de realizar mediciones repetidas sobre el mismo sujeto implica que no es posible aleatorizar el factor tiempo, por lo que las medidas tomadas sobre un mismo individuo están autocorrelacionadas y, por lo tanto, no se cumple el supuesto de independencia de los errores (Schabenberger & Pierce, 2002).

Por lo general, los errores en los datos de series de tiempo tienen correlación seriada, esto es,  $E(\varepsilon_i, \varepsilon_{i+j}) \neq 0$ . Se dice que esos términos de error están autocorrelacionados. La presencia de autocorrelación en los errores tiene diversos efectos sobre el procedimiento ordinario de regresión por mínimos cuadrados (Montgomery, 2004). Estos efectos se resumen a continuación:

- i Los coeficientes de regresión ordinaria por mínimos cuadrados siguen siendo insesgados, pero ya no son estimados con varianza mínima. Se dice que esos estimadores son ineficientes.

- ii Cuando los errores están auto correlacionados positivamente, el cuadrado medio de los residuales, puede subestimar mucho a  $\sigma^2$ , en consecuencia, los errores estándar de los coeficientes de regresión pueden ser muy pequeños, los intervalos de confianza son más cortos de lo que deberían ser, y las pruebas de hipótesis acerca de los coeficientes individuales de regresión pueden indicar que uno o más de los regresores contribuyen en forma significativa al modelo, cuando en realidad no lo hacen.
- iii Los intervalos de confianza y las pruebas de hipótesis basados en las distribuciones t y F ya no son adecuados, en el sentido estricto.

Los modelos de efectos mixtos se suelen utilizar cuando los datos constan de grupos de observaciones. Los parámetros de efectos fijos describen los patrones generales de los datos y los parámetros de efectos aleatorios describen grupos específicos. Si el modelo no es lineal en los parámetros, se denomina modelo de efectos mixtos no lineales (Davidian & Giltinan, 2003). Los datos de crecimiento en los que la trayectoria de cambio no es lineal son un tipo común de datos modelados con modelos de efectos mixtos no lineales, con el sujeto tratado como un conglomerado y las mediciones repetidas como las observaciones individuales anidadas dentro del conglomerado. En tal situación, los efectos fijos describen la trayectoria de cambio de la población, mientras que los efectos aleatorios reflejan la variabilidad entre individuos (Stegman *et al.*, 2017).

El modelo general para un modelo de efectos mixtos es:

$$y_{it} = f(z_{it}, b_{ki}) + e_{it}$$

Con:

$$b_{ki} = \beta_k + r_{ki}$$

Donde  $y_{it}$  es el resultado de interés para el individuo  $i=1, \dots, n$  medido repetidamente en tiempos  $t = 1, \dots, T_i$ , siguiendo la forma funcional  $f$ , que depende de los parámetros  $b_{ki}$  para  $k = 1, \dots, k$  y  $z_{it}$  es el valor de la métrica de tiempo. Los efectos aleatorios,  $r_{ki}$  y  $e_{it}$ , tienen medias de 0, y normalmente se supone que siguen una distribución normal con sus respectivas matrices de covarianza (Pinheiro & Bates, 2000).



Los criterios de bondad de ajuste utilizados son los siguientes (Burnham & Anderson, 2002).

- ✓ Criterio de información de Akaike

$$AIC = -2\log (MV) + 2(q + p)$$

- ✓ Criterio de información bayesiano de Schwartz

$$BIC = -2\log (MV) + 2(q + p) \log (n)$$

Donde  $MV$  es máxima verosimilitud;  $q$  el número de parámetros de la estructura de covarianzas;  $p$  el rango de la matriz de diseño  $X$ ; y  $n$  el número de observaciones.

Los modelos mixtos han sido empleados para el ajuste de curvas de crecimiento en frutales. Tassile *et al.* (2002), estudiaron la aplicación de modelos no lineales mixtos en la descripción del crecimiento de frutos de pepita, durante varios ciclos productivos, en distintas localidades del Alto Valle de Río Negro y Neuquén. Analizaron la variabilidad de los frutos de una localización en torno al patrón de crecimiento general y desagregaron la componente de varianza incorporada por el factor año al considerar distintas temporadas. Lograron una mejor adecuación del modelo ajustado (Logístico), incorporando la estructura de covarianza correspondiente al modelo autorregresivo de primer orden AR (1), y considerando como factores aleatorios al año y fruto.

Bramardi *et al.* (1997) propusieron un modelo logístico de tercer parámetro para describir el patrón de crecimiento de peras William's y Packahm's Triumph. Este modelo y su gran capacidad de predicción se encuentra con las mejores propiedades en las medidas de no linealidad, por lo cual Giménez & Tassile (2015) establecieron que, para modelar el crecimiento pueden haber factores de variabilidad estocástica que no se ven directamente, como el lugar o la parcela, la genética y el clima, entre otros. Además, el seguimiento de los frutos en periodos de tiempo constantes provoca correlaciones entre las mediciones generando medidas longitudinales durante el registro de datos. Por eso propusieron el ajuste de modelos no lineales mixtos.

Estos modelos son adecuados para analizar datos correlacionados, comunes en varias disciplinas como la farmacocinética, la agricultura y la medicina. Por otra parte,

permiten modelar la variabilidad estocástica del desarrollo de la fruta en sus diversas fuentes a través de parámetros fijos y aleatorios. Los efectos aleatorios en cualquier parámetro son útiles no sólo para considerar la variabilidad en una población, sino también para los estimadores empíricos bayesianos. En este contexto, las covariables se explican a través de los parámetros fijos y los modelos de regresión no lineal, mientras que las fuentes tradicionales de heterogeneidad y correlación pueden considerarse mediante la inclusión de un modelo de efectos aleatorios mixtos.

Hoseien-Zadeh (2017), comparó cinco ecuaciones matemáticas no lineales mixtas y estimó los parámetros de los modelos Brody, Exponencial negativo, Logístico, Gompertz y von Bertalanffy, para describir las curvas de crecimiento en ovejas Moghani iraníes. Después de seleccionar el modelo mejor ajustado basado en criterios puramente estadísticos, estimaron los componentes de la varianza y los parámetros de la curva de crecimiento. El animal fue considerado como sujeto en los modelos. Se examinó la bondad de ajuste de los modelos mixtos no lineales utilizando el criterio de información de Akaike (AIC) y la varianza residual. El modelo Logístico proporcionó el mejor ajuste de la curva de crecimiento, debido a los valores más bajos de AIC y la varianza residual en comparación con otros modelos.

Ibiapina *et al.* (2020), examinaron la inclusión de efectos aleatorios en modelos no lineales, para luego identificar los modelos más adecuados y describir el crecimiento de pollos naturalizados. Se estimaron los parámetros de los modelos no lineales Gompertz, Logístico y Von Bertalanffy y se ajustaron mediante el método de Gauss-Newton. La varianza residual disminuyó en más de un 50% cuando se añadieron efectos aleatorios al modelo. Los mejores ajustes en la estimación de la curva de crecimiento en hembras asociando los efectos aleatorios a los tres parámetros de los modelos Gompertz y Logístico. La asociación de los efectos aleatorios con dos parámetros (peso asintótico y tasa de maduración) y con los tres parámetros del modelo Logístico proporcionó los mejores ajustes para los machos. La inclusión de efectos aleatorios en los modelos proporciona una mayor precisión en la estimación de la curva de crecimiento.

El crecimiento de los frutos en mandarina, y particularmente de tangor Murcott, ha sido poco estudiado, por lo que el objetivo de este Capítulo fue conocer la curva que

describe el crecimiento de los frutos de tangor Murcott y encontrar un modelo matemático para describir el patrón de dicho crecimiento, aplicando técnicas de regresión no lineal a datos de campo recopilados en la provincia de Corrientes, Argentina. Una vez seleccionado el modelo que mejor ajuste al crecimiento de estos frutos y estimados sus parámetros, utilizar modelos no lineales mixtos con el objeto de determinar si es posible mejorar el ajuste del modelo seleccionado.

### **3.3. MATERIALES Y MÉTODOS**

#### **3.3.1. Datos Experimentales**

En este Capítulo se trabajó con algunos de los huertos descritos en el Capítulo II. Se usaron dos selecciones diferentes de huertos para la modelización del crecimiento de los frutos de tangor Murcott mediante modelos no lineales de crecimiento sigmoideal y luego para la aplicación de modelos no lineales mixtos al modelo seleccionado. En la Tabla 3.1 se presentan los huertos empleados para cada técnica, se indican las temporadas en las que se relevó la información de campo en los mismos y las fechas de plena floración y de realización de mediciones del diámetro ecuatorial en frutos marcados.

Tabla 3.3. Técnica de análisis (modelos no lineales de crecimiento sigmoideal, MNLCS; modelos no lineales mixtos, MNLM), huertos, temporadas, fechas de plena flor, de inicio de mediciones del diámetro ecuatorial de los frutos, expresado días después de plena flor

Técnica	Temporada	Huerto	Fecha plena flor	Fecha inicio mediciones	Días después de plena flor
MNLCS	2010/11	1.1	26/09/2010	07/01/2011	103
		3.1	23/09/2010	06/01/2011	100
	2015/16	1.4	16/08/2015	17/12/2015	102
		3.2	26/09/2015	17/12/2015	82
MNLM	2010/11	1.1	26/01/2010	07/01/2011	103
	2014/15	1.2	30/09/2014	26/11/2014	59
		1.3	01/10/2014	26/11/2014	58
		1.4	31/09/2015	17/12/2015	73
	2015/16	1.4	31/09/2015	17/12/2015	73
		1.6	31/09/2015	17/12/2015	73
	2010/11	2.1	17/09/2010	06/01/2011	79
	2011/12		23/09/2012	01/12/2012	62
	2013/14		30/08/2013	03/12/2013	94
	2010/11	3.1	23/09/2010	06/01/2011	99
	2013/14		18/08/2013	03/12/2013	105
	2015/16	3.2	26/09/2015	17/12/2015	81
	2011/12	4.1	25/09/2011	01/01/2012	96

En cada huerto se realizó un censo, en el que se registró la cantidad de árboles presentes según tres tipos, sanos, enfermos y replantes, consignándose también los ausentes. A partir de esta información se determinó el número de árboles por huerto estableciendo una de las componentes del modelo de pronóstico (número de árboles por hectárea o por huerto).

En el perímetro de cada lote se descartó un borde compuesto por las tres primeras filas y en el resto se seleccionaron 10 árboles sobre las dos diagonales principales, mediante un muestreo sistemático con arranque aleatorio. En la Figura 3.1 se presenta, a modo de ejemplo, el croquis de los huertos 1.4 y 3.2 y el esquema de muestreo de árboles.



(a)

(b)

Figura 3.2. Croquis de los huertos en estudio, selección de plantas.

(a). Huerto 1.4

(b). Huerto 3.2

En cada árbol se identificaron 30 frutos, distribuidos en toda la copa, a los que se realizó el seguimiento del diámetro ecuatorial (mm) empleando calibre digital, con una frecuencia de 10 a 15 días (Figura 3.2). Las mediciones se iniciaron una vez terminados los manejos culturales que definen la carga y fueron referidas a días desde la fecha de plena floración (DDPF), que corresponde al 80% de flores abiertas.

### 3.3.1.1. Variables estudiadas

La modelización del crecimiento de los frutos se realizó considerando las siguientes variables:

**Variable independiente:** *Momentos de observación*, que se establecieron en días después de plena floración (DDPF) y se realizaron con una frecuencia quincenal. En la definición de los modelos, esta fue la variable independiente.

**Variable dependiente:** *Diámetro ecuatorial*, en los frutos seleccionados se midió el diámetro ecuatorial en mm (DE) con calibre digital, desde los 80 DDPF, hasta el momento de la cosecha, con una frecuencia entre 7 y 10 días.

### 3.3.1.2. Factores

**Huerto:** el huerto en el que se realizaron las mediciones fue considerado como un factor en la definición de los modelos.

**Temporada:** la temporada a la que corresponden las mediciones fue considerada como un factor en la definición de los modelos.



Figura 3.3. (a) Registro de floración con el objeto de determinar la fecha de plena flor (80% flores abiertas), (b) Identificación de árboles para registro del crecimiento de los frutos seleccionados (c) Identificación de frutos y registro del diámetro con calibre digital.

### 3.3.2. Análisis de datos

Se aplicaron técnicas de análisis exploratorio de datos mediante herramientas gráficas y se modelizó el crecimiento de los frutos mediante el ajuste de modelos sigmoidales.

Dada la complejidad de las técnicas de regresión no-lineal y sobre todo del cálculo de las medidas de no-linealidad, que dificultan el análisis sobre la totalidad de la muestra de frutos recolectada, se seleccionó al azar el registro de crecimiento de 8 frutos de cada huerto y temporada para evaluar la adecuación de la mejor combinación modelo-parametrización (Avanza *et al.*, 2008; Bramardi *et al.*, 1997).

Para la aplicación de los modelos no lineales mixtos, la cantidad de mediciones en las diferentes temporadas varió entre 11 y 22 a lo largo del año. El modelo utilizado fue el Logístico en su quinta reparametrización (L5), en el que los valores de los parámetros iniciales utilizados fueron obtenidos previamente tal como fue tal como fue descrito en punto 3.2.1.3.

#### 3.3.2.1. Modelización del crecimiento de frutos de tangor Murcott

Los modelos se ajustaron mediante el método computacional de Marquardt, a través del procedimiento PROC NLIN del paquete estadístico SAS (SAS University Edition, 2019). Los valores iniciales de los estimadores se calcularon utilizando la metodología propuesta por (Ratkowsky, 1983).

Con la información mencionada anteriormente se ajustaron los modelos estadísticos que se describieron en las (Tablas 3.1 y 3.2) (reparametrizaciones de las formas iniciales más frecuentemente utilizadas en curvas de crecimiento) (Ratkowsky, 1983).

Los modelos fueron evaluados para seleccionar los que mejor representan el equilibrio entre la capacidad de ajuste de los datos y la complejidad del análisis de éstos. Modelos con un número elevado de parámetros tienden a ajustar mejor a una base de

datos que aquellos con número bajo de parámetros; sin embargo, tienden a ser más inestables y a modelar la variabilidad de los datos más que su tendencia.

Los modelos no lineales se evaluaron mediante los criterios clásicos de bondad de ajuste: distribución de los residuos y error medio al cuadrado  $\sqrt{CM\epsilon}$ . Además, se evaluaron las magnitudes de NI y la significación y no linealidad de EP (Ratkowsky, 1989); estos últimos miden el grado de no-linealidad en un conjunto de datos-modelo-parametrización y si resultan no significativos, podría considerarse que el modelo se comporta como lineal con todas las propiedades que esto implica. Para el cálculo de las medidas de no linealidad se utilizó el programa desarrollado por Bramardi (1995), con el uso del PROC IML de SAS (SAS University Edition, 2019).

Para examinar las posibles diferencias debidas a huertos y temporadas, entre los parámetros obtenidos con la parametrización adecuada se realizó la prueba de razón de verosimilitud propuesta por (Regazzi & Silva, 2004) mediante el procedimiento NLIN-SAS (SAS University Edition, 2019).

Esta prueba consiste en estimar las sumas de cuadrados residuales del modelo completo (SQRcomp), que considera todos los parámetros  $\alpha$ ,  $\beta$  y  $\gamma$  para todas las combinaciones de lote y temporada, y el modelo restringido (SQRrest), donde se fija o se considera un parámetro para las diferentes combinaciones de lote-temporada. El estadístico resultante es un Chi-cuadrado, que puede estimarse según:

$$\chi^2 = -N \ln \left( \frac{SCR_{com}}{SCR_{rest}} \right), \quad \text{con } v = p_{comp} - p_{rest}$$

Donde:  $SCR_{com}$  es la suma de cuadrados residual del modelo completo;  $SCR_{rest}$  la suma de cuadrados residual del modelo restringido; y  $v$  los grados de libertad.

Esta prueba se realizó con los 32 frutos seleccionados de diferentes temporadas y huertos.

Se plantearon las siguientes hipótesis estadísticas:

$H_0$ : los parámetros del modelo de crecimiento de frutos no difieren entre temporadas y huerto.



H<sub>1</sub>: los parámetros del modelo de crecimiento de frutos difieren entre temporadas y huerto.

### 3.3.2.2. Análisis de modelos no lineales mixtos

Los modelos no lineales con efectos mixtos se ajustaron por el método de máxima verosimilitud

$$y = \frac{1}{(\alpha+a)+\exp(\beta+b)\gamma+g^x}$$

El proceso de ajuste fue para cada combinación de temporada y huerto. Se estimaron modelos considerando efectos fijos y aleatorios. El ajuste de los parámetros se realizó con el método de máxima verosimilitud restringida (REML), y los efectos aleatorios de la covarianza: no estructurados, para la selección del modelo se utilizaron los criterios AIC. Los análisis de datos se realizaron con el software (R Development Core Team<sup>10</sup>).

Se estimaron modelos, considerando efectos fijos y aleatorios, donde el sujeto es el fruto.

## 3.4. RESULTADOS Y DISCUSIÓN

A continuación, se presentan y discuten los resultados obtenidos al ajustar modelos de curvas sigmoideas al crecimiento de frutos de tangor Murcott, mediante las técnicas tradicionales de regresión no lineal, y luego aplicando modelos no lineales mixtos para optimizar los parámetros del modelo seleccionado previamente.

---

<sup>10</sup> Disponible en : <https://www.r-project.org>

### 3.4.1. Selección de modelos sigmoidales

En la Figura 3.4 se presenta el registro de los diámetros ecuatoriales (mm) como una función de los días después de la floración plena (DDPF).

Se puede decir que el crecimiento de los frutos de tangor Murcott presenta un comportamiento sigmoidal, donde se pueden diferenciar las tres fases, caracterizadas de la siguiente manera:

La fase I se encuentra incompleta, atribuible al comienzo de medición que corresponde entre los 83 y 104 días después de plena flor, momentos similares fueron referenciados por Avanza (2011). La fase II que corresponde al período que va desde la caída fisiológica de frutos hasta un poco antes de su cambio de color, se caracteriza por una expansión marcada de los tejidos, acompañada de un agrandamiento celular y la formación de un mesocarpo esponjoso, para esta variedad ocurre alrededor de los 230 DDPF. Por último, la fase III, se caracteriza por una reducida tasa de crecimiento, mientras el fruto se mantiene en el árbol y comprende cambios asociados a la maduración (Agustí *et al.*, 2003).

En el ajuste de los modelos Logístico, Gompertz, Von Bertalanfy y Brody no hubo problemas de convergencia con el método de Marquardt. En los modelos Richards y Weibull convergieron 14 y 15 frutos respectivamente de los 32 y el modelo MMF no convergió para ningún fruto. Por lo tanto, estos modelos con cuatro parámetros (MMF, Weibull, Richards) no se consideraron para su posterior análisis. La falta de convergencia del método de Marquardt en los modelos de cuatro parámetros puede ser debido a que en el rango de medición de observaciones del crecimiento de los frutos evaluados no presentó un comportamiento fuertemente sigmoideo, lo que implicaría una sobre parametrización de estos modelos para describir el patrón de crecimiento de los frutos de tangor Murcott. Similares resultados fueron encontrados por Bramardi *et al.* (1997), para frutos de pera de los cultivares William's y Packham's Triumph.

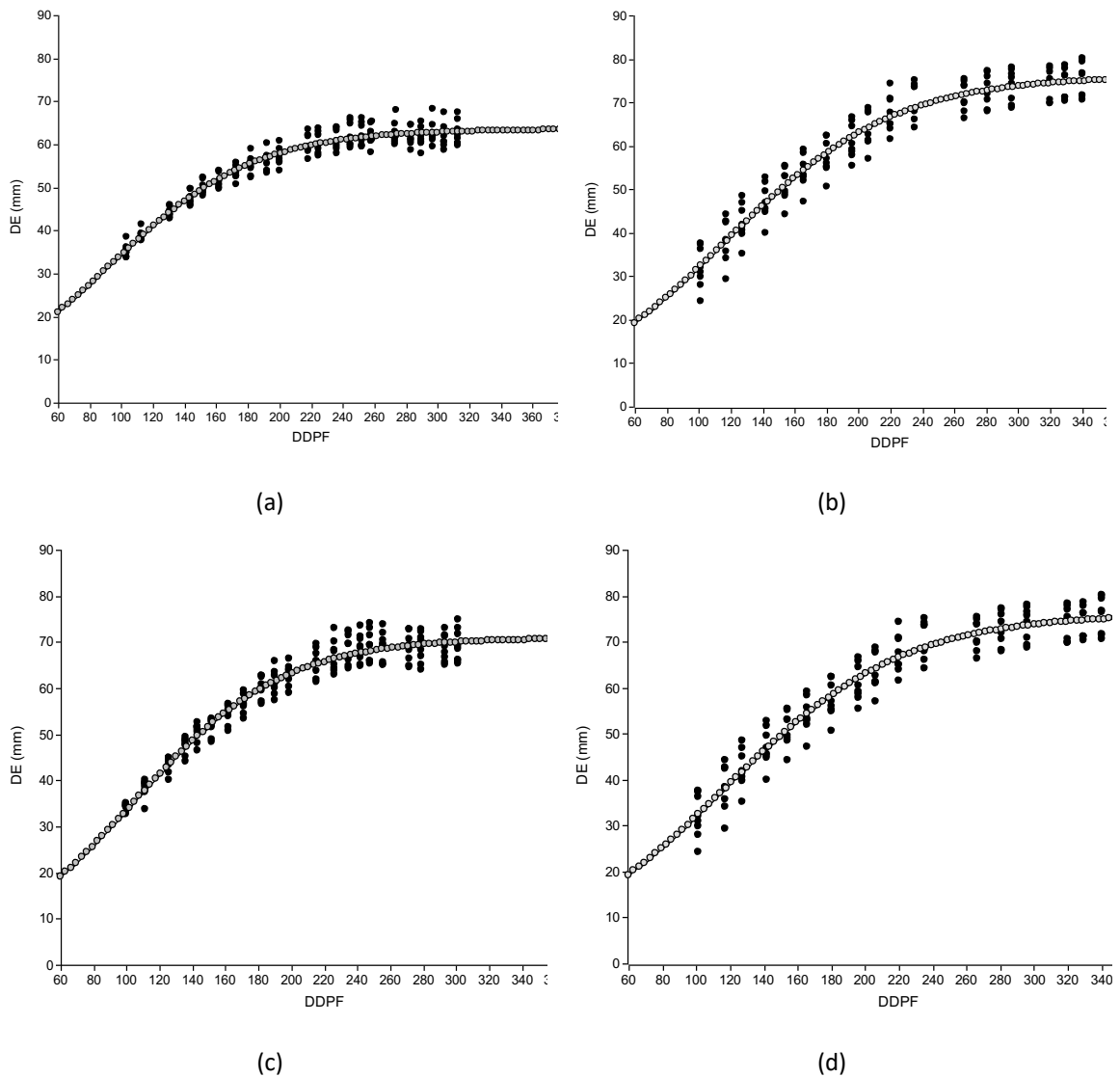


Figura 3.4. Curvas promedio de crecimiento de frutos tangor Murcott (diámetro ecuatorial vs días después de la plena floración DDPF) ajuste del modelo logístico en su 5ta reparametrización, en huertos y temporadas

a: Huerto 1.1-2010/11, b: Huerto 1.4-2015/16, c: Huerto 3.1-2010/11, d: Huerto 3.2 2015/16

En la siguiente tabla se presenta el comportamiento de los modelos estudiados según los criterios de selección utilizados.

Tabla 3.5. Criterios de selección de modelos (raíz cuadrada del cuadrado medio del error,  $\sqrt{CM\varepsilon}$ , y no linealidad intrínseca, NI) para los diferentes modelos estudiados

Modelo	$\sqrt{CM\varepsilon}$	NI
Logístico	1,002	0,0588
Gompertz	1,068	0,0692
Von Bertalanfy	1,240	0,0618
Brody	1,200	0,0618

La medida de NI de las parametrizaciones iniciales de los modelos Logístico (L1), Gompertz (G1), Von Bertalanfy y Brody resultó ser no significativa ( $> 0,05$ ). Estos valores miden el grado de no-linealidad en un modelo y si no es significativo indica que el grado de curvatura de la curva solución es lo suficientemente despreciable como para considerar que se aleja de la forma lineal, es decir, puede considerarse que el modelo se comporta como un modelo lineal. Por lo tanto, desde este punto de vista los cuatro modelos son adecuados.

Con respecto a los criterios clásicos de selección los cuatro modelos mostraron un adecuado comportamiento de los residuales y en todos los casos los parámetros estimados fueron altamente significativos de acuerdo con la prueba t ( $p < 0,001$ ). De la magnitud del error residual que se observa en la Tabla 3.5, resulta que los modelos que mejor ajustaron fueron el Logístico y Gompertz. Por lo que se continuó con el estudio de los parámetros correspondiente a las distintas parametrizaciones sólo en estos dos modelos.

En relación con los criterios clásicos de selección que permitieron evaluar la bondad del ajuste de los modelos, en ambos huertos y temporadas las estimaciones de  $\sigma$  arrojaron el menor valor para el modelo Logístico, seguido de Gompertz, Von Bertalanfy y Brody. Pero lo determinante en la elección de un modelo no lineal son las medidas de NI y EP.

En la Tabla 3.6 se presentan las medidas de EP de cada reparametrización correspondientes a los modelos Logístico y Gompertz.

Se puede observar que los valores de EP estuvieron muy por encima de los valores de NI. La presencia de un EP superior a la NI hace referencia a que el EP domina la no linealidad de modelos, resultando una ventaja porque el EP puede disminuirse con una adecuada reparametrización.

Tabla 3.6. Efecto parámetro para cada fruto (por combinación temporada-huerto) en las reparametrizaciones de los modelos Logístico (L) y Gompertz (G)

Temporada y huerto	Fruto	L1	L2	L3	L4	L5	L6	L7	G1	G2	G3	G4	
2010/11	1	0,4553	0,5659*	2,3534*	0,4437	0,4043	2,2049*	2,0842*	0,6084*	2,1776*	0,9043*	2,1390*	
	1.1	2	0,4746	0,4856	1,6819*	0,4651	0,4123	1,5276*	1,4619*	0,6319*	1,4821*	0,8899*	1,4563*
		3	0,4524	0,4405	1,2458*	0,4444	0,3895	1,1085*	1,0689*	0,6022*	1,0569*	0,8570*	1,0355*
		4	0,41	0,4086	1,0132*	0,4031	0,3558	0,9008*	0,8101*	0,4910	0,8285*	0,7665*	0,8122*
		5	0,5121	0,5226	1,5716*	0,5025	0,4487	1,4244*	1,2988*	0,6438*	1,3700*	1,0055*	1,3461*
		6	0,2294	0,3539	1,0841*	0,2238	0,2049	1,0142*	0,9248*	0,2923	0,9799*	0,4719	0,9667*
		7	0,3404	0,4236	1,2630*	0,333	0,2979	1,1554*	1,0846*	0,4449	1,0927*	0,6462*	1,0747*
		8	0,2659	0,4667	2,0168*	0,2567	0,2433	1,9375*	1,7237*	0,3373	2,0208*	0,7156*	1,9993*
2010/11	9	0,3321	1,6254*	1,6378*	0,3224	0,2882	1,5140*	1,4277*	0,4218	1,3128*	0,5398	1,2954*	
	3.1	10	0,3743	0,7057*	0,7113*	0,3695	0,2959	0,5772*	0,6637*	0,5299	0,5077	0,2416	0,4998
		11	0,5161	1,8260*	1,8404*	0,5065	0,4463	1,6700*	0,5098	0,6744*	1,5219*	0,9407*	1,4968*
		12	0,4037	1,3071*	1,3172*	0,3968	0,3426	1,1759*	1,1140*	0,6306*	1,1795*	0,7757*	1,5171*
		13	0,1778	0,6955*	0,7005*	0,1741	0,1489	0,6284*	0,6224*	0,2325	0,4888	0,0748	0,4795
		14	0,2982	1,1394*	1,1477*	0,2927	0,2503	1,0296*	1,0132*	0,4230	0,8903*	0,1702	0,8737*
		15	0,2313	0,8611*	0,8677*	0,2269	0,199	0,7873*	0,7294*	0,3372	0,7779*	0,4404	0,7650*
		16	0,1714	0,6898*	0,6898*	0,1679	0,149	0,6319*	0,5728*	0,2766	0,7009*	0,3885	0,6954*
2015/16	17	0,3822	1,3542*	1,3659*	0,3741	0,322	1,2157*	1,1701*	0,5503	1,1960*	0,6942*	1,1701*	
	1.4	18	0,4566	1,7423*	1,7575*	0,0333	0,3889	1,5818*	1,5512*	0,6546*	1,5823*	0,8666*	1,5512*
		19	0,5448	1,1725*	1,1830*	0,3961	0,338	1,0341*	1,0802*	0,8214*	1,1072*	0,8179*	1,0802*
		20	0,433	1,4561*	1,4885*	0,3569	0,3071	1,3350*	1,3141*	0,5749*	1,3419*	0,36	1,3141*
		21	0,6376	1,4997*	1,5129*	0,4239	0,3678	1,3501*	1,2421*	0,8848*	1,2686*	0,7647*	1,2421*
		22	0,3335	1,2854*	1,2965*	0,3261	0,2824	1,1636*	1,1080*	0,4755	1,1310*	0,6017*	1,1080*
		23	0,2397	1,1703*	1,1801*	0,2336	0,2052	1,0804*	1,0676*	0,3571	1,0862*	0,3133	1,0676*
		24	0,2739	1,2730*	1,2839*	0,267	0,2373	1,1800*	1,0835*	0,3542	1,1014*	0,501	1,0835*
2015/16	25	0,6384*	0,9085*	0,9085*	0,6303*	0,5309	0,7603*	0,8748*	0,7845*	0,7009*	1,0010*	0,7059*	
	3.2	26	0,41	0,521	0,521	0,4062	0,3289	0,4199	0,5509	0,6648*	0,5001	0,7215*	0,5419
		27	0,3192	0,5311	0,5311	0,3147	0,2748	0,4661	0,4556	0,465	0,5047	0,6934*	0,4981
		28	0,3764	0,6298*	0,6298*	0,3712	0,3186	0,54	0,5593	0,5682	0,5794	0,7562*	0,5733
		29	0,9443*	0,6387*	0,6783*	0,9360*	0,7123*	0,8076*	1,0527*	1,4782*	1,0822*	1,5500*	1,0310*
		30	0,6118*	0,7395*	0,7395*	0,6050*	0,4741	0,5757	0,8238*	1,0192*	0,7250*	1,0111*	0,8398*
		31	1,1088*	0,8027*	0,8027*	1,0987*	0,8822*	0,9464*	1,2002*	1,6069*	1,2141*	2,1020*	1,4005*
		32	0,425	0,8107*	0,8107*	0,4193	0,3566	0,6960*	0,7204*	0,6144*	0,6818*	0,7371*	0,6712*
Promedio		0,4306	0,9079	1,1823	0,3975	0,3501	1,0763	1,0301	0,6079	0,6079	0,7288	1,0728	

\* Indica significancia para  $\alpha=0,05$

De todas las parametrizaciones analizadas se podría considerar como las más satisfactorias L1, L4 y L5 por cuanto en ellas entre 28 y 30, de los 32 frutos presentaron

un comportamiento lineal, en las demás reparametrizaciones, solamente en el 50% de los frutos se logró este buen comportamiento de los parámetros.

Los valores de la Tabla 3.6 permiten deducir que la reparametrización más adecuada entre las estudiadas para los modelos Logístico y Gompertz, fue L5 por presentar menor valor medio de EP, aunque los modelos L1 y L4 se consideran adecuados para describir el patrón de crecimiento. A su vez la parametrización L5 presenta un valor de medida de no-linealidad efecto parámetro medio menor que L1 (0,4306) y L4 (0,3975), y es más homogénea, con una dispersión de 0,1528 (0,2021 para L1 y 0,2094 para L4).

En ambos huertos y temporadas los valores de EP para este L5 fueron menores al valor crítico, en 30 frutos de los 32 analizados, estos resultados concuerdan con los valores hallado para frutos de naranjo Valencia late por Avanza *et al.* (2008) donde el 93,3% de los frutos obtuvieron un menor valor de EP en esta parametrización.

Fernández *et al.* (2014) destacan que la elección de la parametrización afecta al modelo y por lo tanto influye en la fiabilidad e inferencias sobre los parámetros estimados. Los valores EP fueron más bajos que la NI para el crecimiento de fruta pequi o nuez souari (*Caryocar brasiliense*) en L6 y G4. Además, mencionan que las parametrizaciones en los modelos L6 y G3 para describir el crecimiento de frutas pequi, son los que tienen mayor error estándar para la estimación del parámetro  $\beta$ , resaltando la necesidad de un mayor esfuerzo en el cálculo computacional dado que estas estimaciones necesitan más iteraciones para lograr la convergencia de la curva solución.

Entre los modelos estudiados, el Logístico en su quinta reparametrización es considerado el más adecuado para describir el patrón de crecimiento de frutos de tangor Murcott, por lo que fue seleccionado y en la Tabla 3.7, se presenta la estimación de los parámetros para los distintos huertos y temporadas.

Tabla 3.7. Estimación de los parámetros  $\alpha$ ,  $\beta$  y  $\gamma$  del modelo logístico en su 5<sup>a</sup> reparametrización por temporada y huerto

Huerto	Temporada	$\alpha$	$\beta$	$\gamma$	Asíntota	Intercepto	Tasa crecimiento
1.1	2010/11	0,015751*	-2,14323*	0,97857*	63,4850	7,5153	0,0217
1.4	2015/16	0,013157*	-2,11467*	0,98127*	76,0052	7,4722	0,0189
3.1	2010/11	0,014114*	-1,93113*	0,97809*	70,8516	6,2854	0,0222
3.2	2015/16	0,013956*	-2,94956*	0,98480*	71,6538	15,0787	0,0153

\* asíntota: asíntota superior ( $1/\alpha$ ), intercepto  $1/(\alpha + e^\beta)$ , TC: tasa de crecimiento ( $-\ln\gamma$ ).

\* Indica significancia a  $\alpha=0,05$

La interpretación de los parámetros en el modelo L5 según Ratkowsky (1983) es la siguiente:

$\alpha$ : parámetro relacionado con la asíntota superior, en esta parametrización es la inversa ( $1/\alpha$ ).

$\beta$ : parámetro que relaciona la asíntota superior e inferior, describe el valor de Y correspondiente a  $X = 0$  (intercepto). Esta parametrización responde la función:  $[1/(\alpha + e^\beta)]$ .

$\gamma$ : parámetro relacionado con la tasa de crecimiento (TC) desde los valores iniciales (magnitud  $\beta$ ) hasta los valores finales (magnitud  $\alpha$ ), en esta parametrización esta dado por la función  $-\ln\gamma$ .

Las estimaciones de  $\alpha$ ,  $\beta$  y  $\gamma$  presentaron interacción entre temporada y huerto ( $p < 0,001$ ) (Figura 3.3).

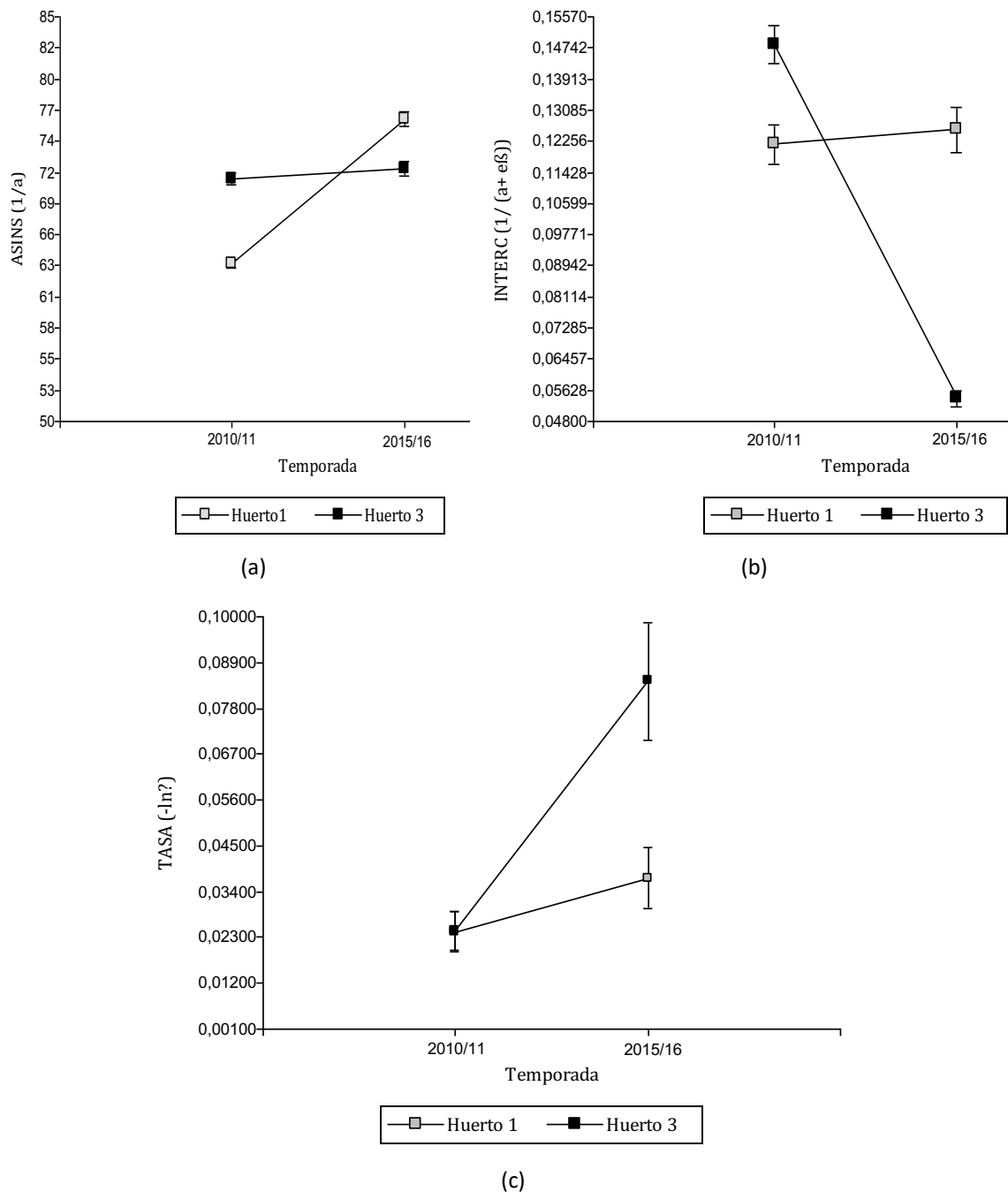


Figura 3.5. Interacción de los factores temporada y huerto sobre los parámetros del modelo Logístico L5:  
a. Asíntota superior ( $1/\alpha$ ),  
b. Intercepto  $1 / (\alpha + e^\beta)$  y  
c. Tasa de crecimiento ( $-\ln\gamma$ )

El crecimiento de frutos de tanger Murcott responde a un modelo logístico en su quinta reparametrización (L5), donde los parámetros  $\alpha$ ,  $\beta$  y  $\gamma$  presentaron interacción entre huerto y temporada (Figura 3.5) En el huerto 1.4 la asíntota superior ( $1/\alpha$ ), fue mayor en la temporada 2015/16, mientras que el intercepto ( $1/(\alpha + e^\beta)$ ), presentó diferencias entre temporadas, y la tasa de crecimiento ( $-\ln\gamma$ ) tuvo un comportamiento



similar a la asíntota superior. Para el huerto 3.1 y 3.2 no hubo variación de la asíntota superior en ambas temporadas, el intercepto fue mayor en la temporada 2010/11 y la tasa de crecimiento en la 2015/16.

Por lo tanto, el crecimiento del fruto está influenciado por los factores climáticos (temperaturas, precipitaciones), asociados a una temporada en particular y características del huerto (suelo, edad de la planta, si posee riego o no).

### 3.4.2. Modelización del crecimiento de frutos con modelos no lineales mixtos

Una de las aplicaciones más frecuentes de los modelos no lineales mixtos es la que permite cuantificar distintas fuentes de variabilidad aleatorias, además de la variabilidad del error. Estos modelos conocidos como efectos aleatorios proporcionan estimaciones de componentes de varianzas y son de amplio uso en mejoramiento genético y ensayos multiambientales. Otra aplicación es cuando se cuenta con cierta estructura de correlación entre las observaciones. Cuando se tiene más de una observación por unidad experimental o unidades experimentales dichas observaciones no son independientes. En otros casos, la falta de independencia entre las observaciones viene dada a lo largo del tiempo (Perelman, *et al.*, 2019).

Con los datos obtenidos del crecimiento de frutos de los huertos y temporadas descritos en la Tabla 3.1, se realizó el ajuste del modelo logístico seleccionado (L5), utilizando los valores iniciales obtenidos en el apartado anterior, contemplando efectos aleatorios mediante la aplicación de modelos mixtos.

Se obtuvieron entre 11 y 22 mediciones anuales repetidas del diámetro ecuatorial, para cada huerto y temporada, por lo que el conjunto de datos para este estudio es de tipo longitudinal.

Los modelos no lineales mixtos se ajustaron con el método de máxima verosimilitud restringida, utilizando la librería de *nlm* (*Nonlinear Mixed-Effects Models*) del paquete R (R Development Core Team<sup>11</sup>).

Se evaluó el ajuste de siete posibilidades de adición de efectos aleatorios, es decir se probó un modelo sin efectos aleatorios, denominado "modelo no lineal de efectos fijos"; tres modelos con adición de un solo efecto aleatorio; tres modelos con adición de dos efectos aleatorios; y un modelo con adición de efecto aleatorio en los tres parámetros. (Tabla 3.8).

Tabla 3.8. Modelos de efecto fijo y efectos aleatorios propuestos para evaluar el patrón crecimiento de tangor Murcott

Efectos	Parámetros	Modelo (L5)
Fijo	$\alpha\beta\gamma$	$y = \frac{1}{\alpha + \exp(\beta)\gamma^x}$
Aleatorio	$\alpha$	$y = \frac{1}{(\alpha + a_1) + \exp(\beta)\gamma^x}$
	$\beta$	$y = \frac{1}{(\alpha) + \exp(\beta + b_2)\gamma^x}$
	$\gamma$	$y = \frac{1}{(\alpha) + \exp(\beta)(\gamma + g_3)^x}$
	$\alpha\beta$	$= \frac{1}{(\alpha + a_1) + \exp(\beta + b_2)\gamma^x}$
	$\alpha\gamma$	$= \frac{1}{(\alpha + a_1) + \exp(\beta)(\gamma + g_3)^x}$
	$\beta\gamma$	$\frac{1}{\alpha + \exp(\beta + b_2)(\gamma + g_3)^x}$
	$\alpha\beta\gamma$	$\frac{1}{(\alpha + a_1) + \exp(\beta + b_2)\gamma(\gamma + g_3)^x}$

y= diámetro ecuatorial (mm), x = DDPF,  $1/\alpha$ : asíntota superior,  $1/(\alpha + e^\beta)$ : intercepto,  $-\ln\gamma$ : TC: tasa de crecimiento y  $\mu_1, \mu_2, \mu_3$  son los efectos aleatorios.

Los modelos se compararon utilizando el criterio de información de Akaike (AIC) (Akaike, 1974) y el criterio de información bayesiano (BIC) (Schwarz, 1978), para la elección del modelo que mejor se ajustaba.

<sup>11</sup> Disponible en: <https://www.r-project.org>

En los modelos que contienen efectos aleatorios, la matriz R del modelo mixto es la matriz tradicionalmente utilizada en el ajuste de curvas de crecimiento (independencia) y las covarianzas entre observaciones son introducidas vía la matriz G del modelo mixto subyacente. Existen muchas situaciones en la que es posible que algunos de los parámetros de una curva (no lineal) de crecimiento varíen de individuo a individuo y se puede reflejar la variabilidad mediante términos aleatorios en el modelo. Lo que cambia es que la función no lineal expresa la esperanza (media) condicional de la variable respuestas dado el o los efectos aleatorios (Balzarini, *et al.*, 2005).

Al modelo seleccionado previamente, que corresponde a un modelo no-lineal de efectos fijos, se incorporaron efectos aleatorios asociados a cada parámetro, dado que la observación se realizó sobre un mismo fruto (sujeto) a lo largo del tiempo y con el fin de establecer y/o explicar si la relación del diámetro ecuatorial se encuentra asociado a cada fruto en cada fecha referida a DDPF.

Los valores obtenidos con cada modelo para los criterios de bondad de ajuste se presentan en la Tabla 3.9, en todos los casos los modelos no lineales mixtos tuvieron un mejor ajuste que los que se asumen con efectos fijos (independientes), presentando los menores valores de AIC.

Tabla 3.9. Valores de AIC de los modelos con efectos fijos y aleatorios asociados a los parámetros  $\alpha$ ,  $\beta$  y  $\gamma$  y temporada y huerto

Temporada	Huerto	Parámetros							
		Efecto Fijo		Efectos Aleatorios					
		$\alpha\beta\gamma$	$a_1$	$b_2$	$g_3$	$a_1b_2$	$a_1g_3$	$b_2g_3$	$a_1b_2g_3$
2010/11	1.1	9967	6974	9299	8700	6033	5980	8026	<b>5836</b>
2014/15	1.2	18854	14889	18068	17344	13604	13534	16356	<b>13407</b>
	1.3	19998	16228	18666	17623	15237	15201	16394	<b>15165</b>
2015/16	1.4	15552	11935	14441	13580	10487	10721	12979	<b>10341</b>
	1.5	15696	13798	14313	13882	12597	12768	13608	<b>12577</b>
2010/11	2.1	18948	15903	16452	15847	15233	15507	15275	<b>15164</b>
	2011/12	18949	15906	16455	15850	15237	15511	15279	<b>15168</b>
	2013/14	15085	13567	13109	12934	12451	12474	12551	<b>12437</b>
2010/11	3.1	17355	13245	16317	15512	12603	12572	14345	<b>12419</b>
	2013/14	10844	9849	10148	9881	9447	9455	9762	<b>9428</b>
	2015/16	12327	9832	11091	10303	8423	8824	9271	<b>8246</b>
2011/12	4.1	17355	14142	15658	14560	12180	12593	13295	<b>12128</b>

Dentro de los modelos mixtos, aquellos que presentan los tres parámetros aleatorios fueron los que mostraron un mejor ajuste en función de la temporada y huerto, por lo que constituyen una estrategia para modelar la covarianza entre las observaciones del diámetro de los frutos.

Las estimaciones de los parámetros del modelo no lineal mixto con los tres parámetros ( $\alpha$ ,  $\beta$  y  $\gamma$ ) del modelo no lineal mixto por temporada y huerto se observan en la Tabla 3.10.

Tabla 3.10. Parámetros estimados para el modelo no lineal mixto, por temporada y huerto.

Temporada	Huerto	$\alpha + a_1$	$\beta + b_2$	$\gamma + g_3$
2010/11	1.1	0,0157	-2,1335	0,9784
2014/15	1.2	0,0163	-1,3693	0,9746
	1.3	0,0150	-1,7000	0,9785
2015/16	1.4	0,0144	-1,9106	0,9799
	1.5	0,0130	-2,0959	0,9813
2010/11	2.1	0,0115	-2,9109	0,9892
2011/12		0,0115	-2,9112	0,9892
2013/14		0,0138	-1,9628	0,9856
2010/11	3.1	0,0143	-1,9463	0,9780
2013/14		0,0139	-1,9611	0,9825
2015/16	3.2	0,0142	-2,8992	0,9850
2011/12	4.1	0,0142	-2,2748	0,9817

Las estimaciones de los parámetros aleatorios mostraron una escasa variación entre temporadas y huertos. Aunque las estimaciones del parámetro ( $\beta + b_2$ ) se diferenciaron en mayor medida respecto de las temporadas y huertos.

A su vez, el hecho de que ( $\alpha + a_2$ ) y ( $\beta + b_2$ ), sean aleatorios, resulta de interés desde el punto de vista agronómico, dado que según las características ambientales (combinación campaña-huerto) que el tamaño de los frutos se modifica (tanto en su inicio -  $\beta$  - como el tamaño final -  $\alpha$  -), por el contrario, la tasa de crecimiento ( $\gamma$ ) está influenciada por la temperatura durante cada etapa de desarrollo y de la humedad del suelo durante la fase III. Las temperaturas medias más elevadas proporcionan las mayores tasas de crecimiento de la fruta y las temperaturas medias bajas dan lugar a tasas de crecimiento más bajas, sin embargo, la humedad edáfica adecuada por lluvia o

riego mejora significativamente el tamaño de la fruta durante la fase III de crecimiento (Davies & Albrigo, 1994).

Kaack & Pedersen (2010) realizaron un análisis de regresión múltiple, con la selección hacia delante de las variables climáticas o sus interacciones, y el crecimiento de los frutos y mostraron que el diámetro de la fruta de la manzana estaba significativamente afectado por dos factores climáticos: los grados-día y el potencial de evaporación. Zadrabec *et al.* (2014), compararon 4 cultivares de manzana en donde el patrón de crecimiento fue exponencial negativo, estableciendo que un aumento lineal del diámetro del fruto al principio del crecimiento de este va seguido de un aumento algo más lento en las etapas finales del crecimiento del fruto. La asíntota superior y el diámetro del fruto observado (medido) fue significativamente mayor en comparación con otros cultivares y en la variedad Gala, varían significativamente en el diámetro final del fruto a lo largo de los años, dependiendo en gran medida de las temperaturas, la ubicación y otros factores ambientales. Resultados similares fueron reportados por Giménez & Tassile (2015).

Por lo establecido anteriormente la introducción de coeficientes de efecto aleatorios en la modelación de curvas de crecimiento de tangor Murcott, permite modelar la covarianza entre observaciones provenientes de un mismo fruto, introduciendo flexibilidad al modelo como para contemplar la variabilidad existente entre individuos de acuerdo con lo establecido por Giménez & Tassile (2015) y Godoy *et al.* (2008), en diferentes aplicaciones de curvas de crecimiento.

A raíz de esos resultados, se procedió a realizar estimaciones de componentes de varianzas a nivel de temporada y huerto mediante un modelo lineal generalizado mixto, utilizando la librería *nlmer (Fitting Nonlinear Mixe-Effect Models)* del software R; siendo el predictor lineal Normal y la función de enlace Identidad.

$$y_{ikj} = \mu + u_i + \tau_j + \varepsilon_{ij}$$

Donde:  $y_{ikj}$  es la respuesta de las estimaciones de componentes de varianzas para los parámetros  $(\alpha + a_1)$ ,  $(\beta + b_1)$  y  $(\gamma + g_1)$  del modelo L5, expresada en porcentaje (%);  $u_i$  es el efecto de la  $i$ -ésima temporada;  $\tau_j$  es el efecto del  $i$ -ésimo huerto; y  $\varepsilon_{ij}$  es el

error aleatorio de la i-ésima temporada en la j-ésima temporada en el k-ésimo parámetro del modelo.

En la Tabla 3.11, se detalla la estimación de las componentes de varianza y desvío estándares de los efectos aleatorios para cada parámetro del modelo logístico en su 5<sup>ta</sup> reparametrización.

Tabla 3.11. Estimación de componentes de varianza (Var%) y desvío estándar (SD) de los efectos aleatorios: Huerto y Temporada.

Efectos aleatorios	$(\alpha + a_1)$		$(\beta + b_1)$		$(\gamma + g_1)$	
	Var%	SD	Var%	SD	Var%	SD
Huerto	43,53	0,11	21,32	0,27	0	0
Temporada	5,00	0,04	25,2	0,29	0	0
Residual	51,47	0,12	53,48	0,43	100	0,19

Se puede observar en la Tabla 3.12., que para el parámetro  $(\alpha + a_1)$ : que corresponde a la asíntota superior, la mayor variabilidad se debe al efecto del huerto (43,53%), es decir que este parámetro se ve influenciado por las características que presentan los huertos y solo un 5% se debe al efecto de temporada. Para el parámetro  $(\beta + b_1)$ : que es el intercepto con el eje y, la mayor variabilidad se debe al efecto de temporada (25,2%) y luego por el huerto (21,32%). Y por último en lo que respecta al parámetro  $(\gamma + g_1)$  que se relaciona con la velocidad de crecimiento de la curva desde un valor inicial ( $\beta$ ) a un valor final ( $\alpha$ ), no presenta variación en los efectos temporada y huerto.

### 3.5. RESUMEN DE RESULTADOS

Se ajustaron modelos de regresión no lineal de tipo sigmoideo: Logístico, Gompertz, Brody, Von Bertalanfy, Weibull, Morgan Mercer Flodin, Richards, y sus respectivas reparametrizaciones. Como criterios de selección del modelo, se consideraron las magnitudes de las medidas de no linealidad (NI y EP) y las estimaciones de la varianza residual para los distintos modelos y parametrizaciones evaluadas. La combinación datos-modelos -parametrización adecuada para describir el crecimiento de frutos de tangor Murcott fue la quinta parametrización del modelo Logístico (L5), donde  $1/\alpha$  es la asíntota superior,  $1/(\alpha + e^\beta)$  es el intercepto y  $-\ln\gamma$  corresponde a la velocidad o tasa de crecimiento de la curva desde un valor inicial ( $\beta$ ) a un valor final ( $\alpha$ ).

La introducción de efectos aleatorios permitió modelar la covarianza entre observaciones provenientes de un mismo fruto, generando flexibilidad en el modelo logístico que permite contemplar la variabilidad existente entre individuos (frutos). Las estimaciones de las componentes de varianzas permitieron establecer que  $(\alpha + a_1)$ , que corresponde al tamaño máximo de los frutos o asíntota superior, está influenciado por las condiciones del huerto, en cambio  $(\beta + b_1)$ , que corresponde al tamaño inicial de los frutos o intercepto, presenta mayor variabilidad a nivel de temporada, y el parámetro  $(\gamma + g_1)$ , que se relaciona con la velocidad de crecimiento (tasa de crecimiento) no presentó variabilidad a nivel temporada ni huerto.

Esto indica que los tamaños finales alcanzados por los frutos ( $y$ ) dependen de las condiciones ambientales particulares donde se desarrollan, indicando que el crecimiento de los frutos está influenciado por las condiciones particulares correspondientes al año (temporada) y al huerto, razón por la cual se estimó una curva de comportamiento medio de crecimiento en función de los días después de plena flor ( $x$ ):

$$y = \frac{1}{0,0140 + e^{(-2,5866) * 0,9831x}}$$

### 3.6. BIBLIOGRAFÍA

1. Andrade, F.; Cirilo, A.; Uhart, S.; Otegui, M. 1996. Ecofisiología del cultivo de maíz. Dekalbpres. 15-39 p.
2. Andrade, L. R.; Cirillo, M. A.; Beijo, L. A., 2014. Proposal of a bootstrap procedure using measures of influence in non-linear regression models with outliers. Acta Sci-Technol, v. 36, n. 1, p. 93-97, 2014.
3. Agustí, M. 2000. Citricultura. Ed. Mundi Prensa. Madrid, España. 416 pp.
4. Agustí, M.; Martínez Fuentes, A.; Mesero, C.; Juan, M.; Almela V. 2003. Cuajado y desarrollo de frutos cítricos. Universidad Politécnica Valencia. Generalitat Valenciana. Valencia. España. 80 pp.
5. Agustí, M.; Mesejo, C.; Mesero, C., Martínez-Fuentes, A.; Zaragoza, S.; Primo Millo, E. 2020. El tamaño del fruto en los cítricos. Levante Agrícola. Revista Internacional de Cítricos. 451, 57-57.
6. Akaike, H. 1974. [A new look at the statistical model identification](#). Journal Article. 19 (6).
7. Amancio, A.L.L.; Silva, J.H.V.; Fernandes, J.B.K.; Sakomura, N.K.; Cruz, G.R.B. 2014. Use of mathematical models in the study of bodily growth in GIFT strain Nile tilapia. Revista Ciência Agronômica, 45, (2): 257-266.
8. Avanza, M.M.; Bramardi, S.J.; Mazza, S.M. 2008. Statistical models to describe the fruit growth pattern in sweet orange 'Valencia late'. Spanish Journal Agricultural. Res. 6 (4): 577-585.
9. AvanzM. 2011. Desarrollo de una propuesta metodológica para la predicción de cosecha de naranjo dulce (*Citrus sinensis* L. Osbeck) var. Valencia late en la provincia de Corrientes, Argentina. Tesis Doctoral de la Universidad Nacional del Nordeste, Área Recursos Naturales.
10. Balzarini, M.; Macchiavelli, R.; Casanoves, F. 2005. Aplicaciones de modelos mixtos en agricultura y forestería. Notas de Clase. Curso Internacional Aplicaciones de Modelos Mixtos en Agricultura y Foresteria. CATIE. Turrialba, Costa Rica. Mimeo. 189 pp.



11. Bates, D.M.; Watts, D.G. 1980. Relative curvature measures of nonlinearity: with discussion. *Journal of the Royal Statistical Society: Series B (Methodological)*. 42 (1): 1-25.
12. Bergh, O. 1982. Prediction of apple fruit size: a promising model. *Agroplanta*, 14: 43-46.
13. Beale, E.M.L.(1960). Confidence regions in non-linear estimation (with Discussion). *J. R. Stat. Soc., Series B* 22, 41-88.
14. Bramardi, S.D. 1995. Modelos de predicción de tamaños de fruta a la cosecha en base a mediciones sucesivas durante el período de crecimiento. Peras cvs. William's y Packman's Triumph en la región del Alto Valle de Río Negro y Neuquén. Tesis para optar el grado de Magister Scientiae- Área Biometría. Escuela para graduados Convenio Facultad de Agronomía- UBA e Instituto Nacional de Tecnología Agropecuaria. 169 pp.
15. Bramardi, S.J. Castro, H.R. Zanelli, M.L. 1997. Fruit Growth Pattern of pear cv. Bartlett and Packham's Triumph to improve hand thinning. *Acta Horticulturae*. (425): 283-293
16. Burnham, Kenneth P., David R. Anderson. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach*. 2° ed. New York: Springer-Verlag. (2).
17. Cañizares Chacin, A. E.; Bonafine, O.; Méndez Natera, J. R.; Laverde, D; Puesme, R. 2012. Crecimiento, desarrollo y calidad del fruto de Lima 'Tahití' (*Citrus latifolia* Tanaka) en el Estado Monagas, Venezuela. *Revista Científica UDO Agrícola* 12(3): 535-541.
18. Carneiro, A. P. S.; Muniz, J. A.; Carneiro, P. L. S.; Malhado, C. H. M.; Martins Filho, R.; Silva, F. F. 2014. Identidade de modelos não lineares para comparar curvas de crescimento de bovinos da raça Tabapuã. *Pesqui Agropecu Bras*, v. 49, n. 1, p. 57-62.
19. Carvajal, V.; Aristizaba, M.; Vallejo, A. 2012. Caracterización del crecimiento del fruto de la gulupa (*Passiflora adules* f. esdules Sims). *Agron. (Colombia)*. 20(1):77-88.
20. Cordeiro, G. M.; Prudente, A. A.; Demétrio, C. G. B. 2009. Uma revisão dos modelos normais não-lineares. *Rev Bras de Biometria*, v. 27, n. 3, p. 360-393.

21. Davidian, Marie & Giltinan, David. (2003). Davidian M, Giltinan DM. Nonlinear models for repeated measurements: an overview and update. *Journal of Agricultural, Biological, and Environmental Statistics*. 8. 387-419.
22. Davies, F. S. ; Albrigo L.G. 1994. *Citricos*. Editorial Acribia. Zaragoza, España. p283
23. Draper N. R.; Smith, H. 1981. *Applied regression analysis*. 2<sup>nd</sup> Edition. Wiley, NY. 709 pp.
24. Fernández, T.J.; Pereira, A.A.; Muniz, J.A.; Savian, T.V. 2014. Seleção de modelos não lineares para a descrição das curvas de crescimento do fruto do cafeeiro. *Coffee Science*, 9 (2): 207-215.
25. Gariglio, N.F.; Weber, M.E.; Castro, D.; Micheloud, N.G.. 2012. Influence of the Environmental Conditions, the Variety, and Different Cultural Practices on the Phenology of Peach in the Central Area of Santa Fe (Argentina). IN: Xiaoyang Zhang (Ed). *Phenology and Climate Change*. Ed. InTech. Croatia. 25 p.
26. Giménez, G.; Tassile, V. 2015. Using Non-linear mixed models and artificial neural network in the fitting growth pattern in pears cv. 'Williams' to predict final sizes at harvest. Conference. XV Conferencia Española y V Encuentro Iberoamericano. Bilbao, 23-25 Septiembre.
27. Godoy, C.; Monterrubianesi, G.; Tognetti, J. 2008. Analysis of highbush blueberry (*Vaccinium corymbosum* L.) fruit growth with exponential mixed models. *Scientia Horticulturae*. 115: 368-376.
28. Gravina, A. 1999. Ciclo fenológico – reproductivo en *Citrus*. Bases Fisiológicas y de Manejo. Universidad de la República, Facultad de Agronomía. 55 p.
29. Hossein-Zadeh, N. 2017. Modelling growth curve in Moghani sheep: Comparison of non-linear mixed growth models and estimation of genetic relationship between growth curve parameters. *The Journal of Agricultural Science*, 155 (7), 1150-1159.
30. Ibiapina Neto, V.; Viera Barbosa, F.J.; Guimarães Campelo, J.E.; Lindenberg R.G. 2020. Non-linear mixed models in the study of growth of naturalized chickens. *Revista Brasileira de Zootecnia (RBZ)*. 19. <https://doi.org/10.37496/rbz4920190201>
31. Iglesias, D.J.; M. Cercos; Colmenero Flores, J.M.; Naranjo, M.; Ríos, G.; Carrera, E.; Ruiz-Ribero, O.; Lliso, I.; Morillon, R.; Tadeo, F.R.; Talón. M. 2007. Physiology of *Citrus* fruiting. Review. *Brazilian Journal Plant Physiology*. 19 (4): 333-362.

32. Kaack, K. & Pedersen, Hanne. 2010) Prediction of Diameter, Weight and Quality of Apple Fruit (*Malus domestica* Borkh.) cv. 'Elstar' using Climatic Variables and their Interactions. *European Journal of Horticultural Science*. 75. 60-70.
33. Montgomery, D.C., Peck, E.A., Vining, G.G. 2004. *Introducción Al Análisis De Regresión Lineal*. Compañía Editorial Continental. México. 567pp.
34. Morgan, P.H.; Mercer, L.P.; Flodin, N.W. 1975. General model for nutritional responses of higher organism. *Proceedings of the National Academy of Sciences USA*. 72:4327-4331.
35. Palacios, J. 2013. *Citricultura*. Editor Talleres Gráficos ALFA BETA. S. A. Argentina. ISBN: 9789-8743-266. 518 pp.
36. Pereira, A.R.; Angelocci, L.R.; Sentelhas, P.C. 2002. *Agrometeorología. Fundamentos y aplicaciones prácticas*. Ed. Agropecuaria. Guaíba, Brasil. 472 pp.
37. Pereira, A. A.; Morais, A. R.; Scalco, M. S.; Fernández, T. J. 2014. Descrição do crescimento vegetativo do cafeeiro cultivar Rubi MG 1192, utilizando modelos de regressão. *Coffee Science*, v. 9, n. 2, p. 236-271
38. Perelman, S.; Garibaldi, L.; Tognetti, P. 2019. *Experimentación y Modelos Estadísticos*. 1ª Edición. Ciudad Autónoma de Buenos Aires: Editorial. Facultad de Agronomía. 348 p. ISBN: 978-987-3738-22-7.
39. Pimentel, C.; Bernacchi, C.; Long, S. 2007. Limitations to photosynthesis at different temperatures in the leaves of *Citrus limon*. *Brazilian Journal Plant Physiology*. 19 (2), 141-147.
40. Pinheiro, J.C.; Bates, D.D. 2000. *Mixed-Efeccts Models in S and S-PLUS*. Springer Verlag New York. Inc. 523 pp.
41. Prado, T.K.L.; Savian, T.V.; Muniz, J.A. 2013. Ajuste dos modelos Gompertz e Logístico a os dados de crescimento de frutos de coqueiro anão verde. *Ciencia Rural*. 43 (5): 803-809.
42. Ratkowsky, D.A. 1983. *Nonlinear regression modeling: a unified practical approach*, Marcel Dekker Inc., NY, USA, 276 pp.
43. Ratkowsky, D.A. 1989. *Handbook of nonlinear regression models*. Marcel Dekker, Inc., NY, 241 pp.
44. Rebolledo-Roa, A. 2017. *Fisiología de la floración y fructificación en los cítricos*. Corpoica. Palmira, Colombia. Capítulo 3. 89-106 pp Disponible:

<https://docplayer.es/33073723-Fisiologiade-la-floracion-y-fructificacion-en-los-citricos.html>.

45. Regazzi A.J.; Silva C.H.O. 2004. Teste para verificar a igualdade de parâmetros e a identidade e de modelos de regressão não-linear. I. Dados no delineamento interinamente casualizado. *Revista Matemática y Estadística* 22 (3), 33-45.
46. R Development Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Austria. ISBN 3-900051-07-0, URL: <https://www.r-project.org/>.
47. Richards, F.J. 1959. A flexible growth function for empirical use. *Journal Experimental Biology*. 10: 290-300.
48. SAS University Edition. 2019. **SAS 9.3 Programming Documentation**. United States.
49. Schabenberger, O.; Pierce, F.J. 2002. *Contemporary Statistical Models for the Plant and Soil Sciences* (1st ed.). CRC Press. <https://doi.org/10.1201/9781420040197>.
50. Schwartz, G. 1978. Estimating the dimension of a model. *The Annals of Statistics*, 6(2): 461-464.
51. Sousa, I.F.; Neto, J.E.K.; Muniz, J.A.; Guimarães, R.M.; Savian, T.V.; Muniz, F.R. 2014. Fitting non line auto regressive models to describe coffee seed germination. *Ciencia Rural*, 44 (11): 2016-2021.
52. Sozzi, G.O. 2007. Capítulo 20: Fisiología de la maduración de los frutos de especies leñosas. En: G.O. Sozzi (Ed.). *Árboles Frutales: Ecofisiología, Cultivo y Aprovechamiento*. Ed. Buenos Aires, Argentina. 669-686p.
53. Stegmann, G.; Jacobucci, R.; Harring, J.; Grimm, K. 2017. Nonlinear Mixed-Effects Modeling Programs in R. *Structural Equation Modeling: A Multidisciplinary Journal*. 25.
54. Taiz, L.; Zeiger, E. *Fisiología vegetal*. Porto Alegre: Artmed Editora, 2004. 719p.
55. Tassile, V.; Álvarez, O.; Bramardi, S. 2002. Aplicación de modelos no lineales mixtos en la descripción de frutos de pepita. V Congreso Latinoamericano de Sociedades Estadísticas. Buenos Aires, Argentina. 133 p.
56. Terra, M. F.; Muniz, J. A.; Savian, T. V. 2010. Ajuste dos modelos Logístico e Gompertz aos dados de crescimento de frutos da tamareira-anã (*Phoenix roebelenii* O'BRIEN). *Magistra*, 22( 1): 1-7.

57. Zadavec P, Veberic R, Stampar F, Schmitzer V, Eler K 2014. Fruit growth patterns of four apple cultivars using nonlinear growth models. *European Journal of Horticultural Science*, 79(2): 52-59.

## CAPITULO IV

### TABLAS DE CRECIMIENTO DE FRUTOS

#### 4.1. INTRODUCCIÓN

Los diferentes mercados consumidores tienen preferencias por frutos de distinto tamaño, ejemplo de ello son las exigencias de tamaños grandes de los consumidores estadounidenses y la preferencia por los calibres pequeños que manifiesta el mercado canadiense. Debido a esto, el conocimiento anticipado de los calibres de los frutos a producir puede constituir una herramienta útil para la negociación, ayuda a programar las cosechas, teniendo en cuenta la necesidad de mano de obra, el espacio de almacenamiento y el aprovisionamiento de materiales de embalaje del tamaño apropiado. Esta variable define, en gran medida, el mercado de destino de comercialización de la fruta y también su precio. Para las condiciones de cultivo de cítricos en nuestro país, y considerando las altas exigencias en calidad de los mercados internacionales, resulta altamente útil pronosticar, con anterioridad a la cosecha, la distribución de tamaños que tendrán las frutas a producir (Ortega *et al.*, 2002).

La predicción se define como una estimación a futuro, posibles resultados que se espera ocurran si las condiciones de producción permanecen, al menos aproximadamente iguales a aquellas en que se construyeron los modelos para la estimación. Existen modelos de predicción de tamaño medio de los frutos a cosecha propuestos para diferentes especies y variedades, en los que se utilizan factores como mediciones del tamaño de los frutos durante un determinado período de tiempo, días post floración, temperaturas durante períodos específicos de crecimiento y carga de los árboles (Franco y Gravina, 2000).

En regiones con fuerte impronta exportadora la producción cítrica se caracteriza por los altos estándares de calidad en la fruta. La madurez, el gusto, el aroma y la

aparición externa son los principales factores que inciden en la decisión de compra, seguidos por el valor nutritivo, el precio y la ausencia de residuos, entre otros. Dentro de la apariencia externa se incluye el tamaño de los frutos, que es una característica de cada variedad. La distribución de los tamaños de los frutos individuales en la cosecha, como muchas características biológicas, suele ser simétrica, con cierto sesgo varietal. En esta distribución del tamaño de los frutos, se presentan tamaños extremos, que por ser muy pequeños o muy grandes, no coinciden con los estándares de preferencia del consumidor y en consecuencia son rechazados por los mercados, obteniéndose de ellos muy bajos precios e incluso la imposibilidad de ser comercializados (Otero *et al.*, 2005).

El tamaño de los frutos ha llegado a ser tan importante como el rendimiento en la determinación de las ganancias, y un mayor beneficio económico es usualmente obtenido a través del aumento de frutos en calibres vendibles a expensas de la reducción del rendimiento general del cultivo en cítricos (Agustí *et al.*, 2020). En el caso de la producción cítrica de Valencia, España, cuando se efectúan raleos para mantener los tamaños de frutos dentro de los calibres entre 48 mm y 55 mm la pérdida de rendimientos es tan grande que no es viable económicamente, por lo que han desarrollado otras prácticas tendientes al aumento del tamaño de los frutos, como por ejemplo el uso de reguladores del crecimiento de síntesis.

En el diario oficial de la Unión Europea, las normas de comercialización general (Parte A) en cítricos <sup>12</sup>, establece que el tamaño de los frutos está determinado por los diámetros máximos de su sección ecuatorial o por el número de frutos. En la sección III: Disposiciones relativas al calibrado de dicha norma establecen:

- Calibre mínimo: en mandarinas, 35 mm.
- Homogeneidad:
  - a) Para garantizar la homogeneidad de calibre, el intervalo de calibres de los productos de un mismo envase depende del diámetro del fruto más pequeño, de acuerdo con la indicación del envase, y no podrá superar: 10 mm, si el diámetro del fruto más pequeño es < 60 mm; 15 mm, si el diámetro del fruto

---

<sup>12</sup> [https://www.ailimpo.com/documentos/2.-Anexo\\_Normas\\_Comercializacion\\_Citricos\\_Reglamento\\_543-2011.pdf](https://www.ailimpo.com/documentos/2.-Anexo_Normas_Comercializacion_Citricos_Reglamento_543-2011.pdf)

más pequeño es  $\geq 60$  mm, pero  $< 80$  mm; 20 mm, si el diámetro del fruto más pequeño es  $\geq 80$  mm pero  $< 110$  mm; y no está limitada en el caso de los frutos con diámetro  $\geq 110$  mm.

- b) Cuando se apliquen códigos de calibre, deberán respetarse los códigos e intervalos de tamaño de frutos.

Determinar relativamente pronto en la temporada de cultivo, el tamaño esperado de fruta y la proporción de fruta de calidad para la exportación al momento de la cosecha, es esencial para que los productores citrícolas puedan realizar una adecuada gestión de la producción. La predicción temprana del tamaño de fruta a obtener ayuda a planificar las operaciones de cosecha y a administrar las estrategias de comercialización con mucha antelación. Esto también permite a los productores eliminar tempranamente los frutos pequeños que no alcanzarán el tamaño deseado para el mercado elegido, evitando así la competencia en el árbol y mejorando los tamaños de los frutos que permanecen.

Para algunas especies se han encontrado modelos que son capaces de describir apropiadamente la distribución de los distintos tamaños (calibres) de fruta. Ejemplos de ello son los modelos planteados para manzana por Ortega *et al.* (1997, 1998).

Mayorano *et al.* (2006) presentaron un modelo computacional para la estimación de crecimiento de frutas en manzanas Granny Smith del Valle del Río Negro (Argentina), basado en cadenas de Markov, que fue validado con éxito. El modelo está orientado al soporte de decisiones comerciales y logísticas, para así realizar pronósticos de las distribuciones de tamaños de una variedad, utiliza correlaciones entre el crecimiento de la fruta y la temperatura y requiere de una base de datos histórica. Además, discretizaron el tamaño de las frutas en 10 rangos definidos para los 40 y 80 días después de plena floración (DDPF), donde establecieron rangos menores de acuerdo con las dispersiones típicas encontradas durante el crecimiento de la variedad y los rangos mayores los determinaron de acuerdo con las clasificaciones definidas por los estándares comerciales de empaque.

Marine *et al.* (2019), midieron diámetros de manzanas 60 días después de la floración, sobre muestras que representaban entre el 7 y el 30% de los frutos presentes



en los árboles y utilizaron estos diámetros para estimar el peso de los frutos en la cosecha mediante ecuaciones de regresión. Al momento de la cosecha, todos los frutos fueron pesados y las distribuciones de los pesos se compararon con las distribuciones de pesos estimadas tempranamente, encontrándose diferencias entre ambas distribuciones para todos los árboles estudiados. Al asignar los frutos a tamaños comerciales adecuados o cajas de recuento (número de frutas/19,05 kg) se comprobó la coincidencia de los máximos en los recuentos de cajas entre las poblaciones estimadas y reales, pero en las poblaciones estimadas muy pocos frutos correspondían a cajas de tamaños pequeños o grandes. Utilizando las estimaciones de la temporada, los productores y empacadores de manzanas comerciales pueden predecir con bastante exactitud el porcentaje de la cosecha que corresponderá a cajas de tamaño máximo, pero una estimación más precisa de la distribución del tamaño de la fruta requerirá probablemente la medición del 50% de los frutos de cada árbol.

Khurshid & Braysher (2009) estudiaron el crecimiento del diámetro de naranjas Washington Navel en la fase II (crecimiento celular), entre enero y julio, durante tres temporadas de cultivo, en Australia. Los frutos fueron seleccionados al azar alrededor de la copa del árbol y se marcaron para medir su diámetro a intervalos quincenales hasta la cosecha. Desarrollaron un modelo que fue capaz de predecir el tamaño final de los frutos y su distribución porcentual, en una fase temprana de desarrollo de los frutos, en los meses de enero-febrero (5-6 meses antes de la cosecha).

Ariza Flores *et. al.* (2010), en el estado de Morelo, México determinaron la calidad física y química de frutos de naranjo dulce Valencia en la cosecha. Describieron los contenidos de sólidos solubles y acidez como suficientes, pero los tamaños de los frutos, con diámetros polares promedio entre 74,2 y 82,8 mm fueron considerados pequeños para la comercialización internacional y estuvieron en mayor proporción entre los calibres 3 (81-92 mm) y 4 (77-88 mm).

En San Pablo (Fundecitrus, 2019) desarrollaron, para naranjo dulce, un modelo de regresión que considera como variable dependiente el tamaño final del frutos (frutos por caja en la cosecha) y variables independientes: el número de frutos por árbol, el tamaño inicial de los frutos, la suma de la producción de la primera y segunda floración en relación a la producción total y las precipitaciones acumuladas de mayo a julio

durante 11 temporadas. El modelo obtenido presentó un  $R^2 = 0,92$ , lo que indica que las variables independientes estudiadas explican conjuntamente el 92% de la variación de los calibres finales de los frutos (frutos por caja en cosecha). La comparación entre el tamaño final de los frutos estimados por este modelo y el observado en las últimas once cosechas presentó un error medio de  $\pm 2\%$ .

Otero *et al.* (2005) ajustaron una curva de referencia que describe el crecimiento promedio de frutos en naranjas Valencia late y Washington Navel, mandarinas Satsuma, Owari y Ellendale y pudo detectar cierta variación entre años para una misma variedad, que posteriormente fue explicada por las temperaturas media y máxima, las precipitaciones y, fundamentalmente por el número de frutos por planta. Mediciones periódicas de los diámetros de frutos le permitieron determinar que la distribución de los tamaños de frutos tiene una forma simétrica, acampanada, cuando esos frutos son cosechados y clasificados por tamaño de empaque.

Los objetivos de este Capítulo han sido, a partir del modelo de crecimiento descrito en el Capítulo III, definir los parámetros específicos para cada uno de los calibres comerciales establecidos por la Unión Europea y desarrollar tablas de crecimiento para frutos de tangor Murcott correspondientes a dichos calibres. Estudiar la relación entre el diámetro de los frutos y el peso, y así transformar los diámetros de frutos a cosecha en pesos de frutos, teniendo en cuenta los tamaños en base a las cajas que se emplean en Argentina para la comercialización de mandarinas. Por último, definir un procedimiento que permita aplicar las tablas de crecimiento desarrolladas para el raleo de frutos en tangor Murcott.

#### 4.1.1. Raleo de frutos

El número de frutos en el árbol es quizás el componente de rendimiento que más influye en el tamaño final de los frutos y en la producción total por árbol, independientemente del marco de plantación, no solo en cultivares cítricos sino en otras especies también. A su vez el número de frutos al momento de la cosecha depende de varios factores: la intensidad y tipo de floración, la intensidad de la caída de estructuras

florales y frutos jóvenes al principio de su período de crecimiento (finales de noviembre, principios de diciembre) y la caída de frutos casi en estado de maduración, previo a la cosecha. El comportamiento de la planta durante estas etapas está altamente asociado al cultivar, al estado sanitario, a las prácticas de manejo y a las condiciones ambientales que predominen en su transcurso. La regulación anual del número de frutos en la planta (carga) en función de las brotaciones y follaje, pasa a ser uno de los temas prioritarios en la obtención de fruta dentro de los calibres de mayor valor comercial, buscando un equilibrio permanente entre la minimización de los porcentajes de la producción de fruta en tamaños pequeños y la maximización de la producción por hectárea. Objetivos que muchas veces se enmascaran entre sí y que dependen de las exigencias cambiantes del mercado y de la estrategia de cada empresa productora, donde los altos rendimientos totales no significan necesariamente altos rendimientos exportables (Otero *et al.*, 2005).

El raleo o aclareo es la remoción de frutos en etapas tempranas con la finalidad de concentrar los esfuerzos productivos en una menor cantidad de frutos que asegure tamaños con alto valor comercial. Es una práctica común en la citricultura y más específicamente utilizada en variedades de mandarina. En Argentina y países limítrofes, el raleo manual es una práctica llevada a cabo por empresas citrícolas de exportación y productores de mercado nacional con gran éxito económico (Otero, 2004).

Para maximizar la producción de los tamaños de fruto de mayor valor comercial, la decisión respecto a la intensidad del raleo no es una tarea fácil, dado que en el desarrollo potencial de los frutos se encuentran involucrados diversos factores en diferentes momentos. Además de los factores intrínsecos al árbol, existen otros que en muchas regiones productivas deben de ser supuestos o pronosticados, como es el caso de las condiciones climáticas, que suelen tener una influencia importante en el desarrollo de los frutos con posterioridad a una intensidad de raleo seleccionada.

Existen básicamente dos grandes formas de expresar y medir la intensidad del raleo de frutos (Otero, 2004):

- 1) La reducción del número de frutos como un porcentaje del número presente en el árbol en un momento dado, por ejemplo, ralear el 20%, 35% o 50% de los frutos totales.

2) La reducción del número de frutos en función de la relación entre los frutos y hojas presentes en el árbol en ese momento, llevando la relación fruto-hoja a valores deseados.

En la determinación de la intensidad de raleo, existe un equilibrio permanente que debe de ser considerado: la reducción del porcentaje de fruta pequeña (no comercializable) y la producción total de fruta comercializable. Tomando como ejemplo, la mandarina Okitsu, en años de alta productividad, con un raleo de frutos del 75%, se logran altos porcentajes de fruta comercializable, pero los máximos rendimientos comercializables (kilos/planta) se logran con un raleo del 50% de los frutos (Otero, 2004).

La práctica del raleo manual de frutos es principalmente empleada en mandarinas, común, Satsuma, Abana, Montenegrina y Murcott. Se recomienda hacerlo lo antes posible en el desarrollo del fruto, luego de la caída fisiológica de fines de noviembre - principios de diciembre (hemisferio sur). El tangor Murcott es conocido por su alternancia productiva, cuando se dejan demasiados frutos en la planta, la planta se extenua en términos de carbohidratos y reservas en general. El raleo manual ha sido la opción más frecuentemente usada por los productores, fundamentalmente porque tiene frutos de muy pequeño calibre al final de la caída fisiológica, y en consecuencia se hace un raleo manual más tardío, porque es más fácil de este modo reconocer y remover el frutito. Esta práctica si bien disminuye la competencia entre los frutos y estimula el crecimiento individual, es más importante para la acumulación de reservas en los brotes y en los árboles, evitando la entrada de la planta en una profunda alternancia (Otero, 2004).

El aclareo puede realizarse en la floración o durante el desarrollo temprano de los frutos y a diferentes intensidades, para optimizar las respuestas del crecimiento de los frutos. Las respuestas al aclareo pueden ser específicas para cada cultivar y lugar. Sutton *et al.* (2020), estudiaron estrategias de aclareo en el sudeste de Estados Unidos, una de las principales regiones productoras de melocotón, donde evaluaron el momento y la intensidad del aclareo en varios cultivares durante tres años. El aclareo en la floración o a los 21 días después de ésta mejoró el tamaño de los frutos en comparación con los árboles sin aclareo en Cary Mac y July Prince, pero redujo la producción de fruta (kg por

árbol), lo que sugiere que el aclareo por sí solo puede no ser una opción viable en esa región. La intensidad del aclareo, evaluada como espaciamientos de 15 y 20 cm entre frutos, no afectó de forma diferencial al peso de los frutos ni al rendimiento. El diámetro de los frutos disminuyó cuadráticamente con el aumento del número de frutos por árbol; el peso de los frutos disminuyó en respuesta al aumento del número de frutos por árbol; y el rendimiento por árbol disminuyó con el aumento del tamaño de los frutos.

Stander & Cronj (2016), estudiaron el efecto del aclareo manual en verano (enero) y otoño (abril), en árboles de mandarina Nadorcott (*Citrus reticulata*). Evaluaron los efectos de los tratamientos sobre el rendimiento total, la calidad y la distribución de tamaños de los frutos. El aclareo manual de verano redujo el número de frutos pequeños y aumentó la tasa de crecimiento inicial de los frutos predominantes, esto dio lugar a un mayor número de frutos de tamaño superior por árbol, sin reducir el rendimiento total ni la calidad de fruta, además se logró mejorar la distribución de los tamaños de los frutos y la ruptura de la alternancia en la carga durante un periodo de dos años. Estos resultados proporcionan a los productores de mandarinas tardías una alternativa a los aclareos químicos para gestionar la carga de cultivo y mejorar el tamaño de los frutos.

#### 4.1.2. Tablas de crecimiento

Las tablas de crecimiento son arreglos tabulares que se construyen sobre la base de un modelo matemático ajustado a las curvas de crecimiento de frutos de una determinada variedad y bajo condiciones particulares de cultivo. Una vez definidas las curvas de crecimiento de los distintos tamaños comerciales de frutos y para diferentes momentos (generalmente referidos a días después de la plena floración, DDPF), se vuelca esa información en tablas de doble entrada similares a la que se presenta a modo de ejemplo en la Tabla 4.1. Donde la primer columna hace referencia al tamaño comercial al momento de la cosecha según normas de la Unión Europea (TC: fruta 1xxx la de mayor tamaño y 5 la de menor tamaño) y las demás columnas indican el diámetro

(en mm) alcanzado por los frutos en diferentes momentos expresado en días después de plena floración (DDPF).

Tabla 4.1. Modelo de tabla de crecimiento

	DDPF													
TC	60	80	100	120	140	160	180	200	220	240	260	280	300	320
1xxx														
1xx														
1x	23,11			40,01										71,05
1														
2														
3														
4														
5														

Estas tablas constituyen una herramienta práctica, que permite pronosticar, en un momento determinado del crecimiento de los frutos y para cada tamaño específico que los frutos han alcanzado en ese momento, los tamaños que dichos frutos alcanzarán al momento de la cosecha. Con ello se logra mejorar las estrategias de cultivo y la planificación de comercialización, por cuanto se puede determinar en un momento dado que frutos no alcanzarán un tamaño adecuado para ser comercializados y la distribución de tamaños que se obtendrá al momento de la cosecha. Por ejemplo, si el raleo manual del monte se realiza 120 días posteriores a la plena floración y queremos dejar solamente los frutos que alcanzarán tamaño comercial 1x o mayor al momento de cosecha (320 DDPF), debemos eliminar todos los frutos con diámetro inferior a 40,01 milímetros en ese momento.

Bramardi *et al.* (2006) construyeron tablas para predecir el tamaño final de los frutos en función de su tamaño a un número determinado de días después de plena floración, que son utilizadas para mejorar la práctica de raleo en montes frutales de peral de las variedades Bartlett y Packhman's Triumph, y de manzanas Red Delicious y Granny Smith en el Alto Valle de Río Negro y Neuquén, Argentina.

Avanza (2011) estudió el crecimiento de frutos de naranjo dulce Valencia late en la provincia de Corrientes, Argentina. Para ello seleccionó entre 10 a 15 árboles al azar por huerto y dentro de cada árbol identificó 30 a 40 frutos, cubriendo el rango de tamaños desde pequeños a grandes, distribuidos por toda la copa del árbol y registró el diámetro ecuatorial con frecuencia semanal o quincenal durante el período de crecimiento de los frutos, entre 10 a 11 meses a partir del momento de plena floración hasta la fecha de cosecha. Encontró que la combinación modelo-datos-parametrización más adecuada para describir el patrón de crecimiento de los frutos de naranjo dulce Valencia late fue la quinta parametrización del modelo Logístico y confeccionó tablas de crecimiento de frutos por calibre comercial de un envase de 15 y 18 kg.

Khurshid *et al.* (2005) en Nueva Gales del Sur, Australia, desarrollaron un macro en Excell para predecir el diámetro final de frutos de cítricos, el cual es distribuido gratuitamente a los productores y asesores como ayuda en la toma de decisiones de gestión y desarrollo de huertos.

Tablas de este tipo para tangor Murcott no han sido desarrolladas en la región cítrica correntina y serán de gran utilidad para técnicos y productores, contribuyendo a establecer criterios para el raleo de frutos según el destino de comercialización, y una vez pasada la fecha de esta práctica de manejo, realizar predicciones de volúmenes de producción y distribución de calibres esperados a la cosecha.

#### 4.1.3. Estimación del peso de los frutos

La estimación del crecimiento de los frutos con información de campo puede ser realizada a partir de modelos estocásticos o mediante curvas de crecimiento. Los primeros se basan en el supuesto de que existe una relación entre el tamaño de los frutos (peso) y la carga de los árboles. Las previsiones de producción de frutales de pepita realizadas por el Centro Servizi Ortofrutticoli (CSO) para la Unión Europea, se basan en modelos estocásticos para la estimación del peso medio de frutos. En cambio, en Argentina para las mismas especies frutales, dicha relación (peso vs. carga), no se ha logrado establecer, por lo que se utilizan curvas de crecimiento para la modelización del

crecimiento de los frutos en función del tiempo y poder así realizar la estimación del peso de frutos al momento de cosecha (Alvarez *et al.*, 2001, Bramardi *et al.*, 2005).

Con el fin de realizar un seguimiento del crecimiento de los frutos en el tiempo y a los efectos de no emplear técnicas destructivas, la mejor opción para desarrollar curvas de crecimiento es el seguimiento del diámetro (ecuatorial o polar) de los frutos. No obstante, al momento de la comercialización, por lo general se requiere conocer el peso de los frutos, por lo que distinta metodología ha sido probada para establecer una relación entre el peso y el diámetro de estos a efectos de convertir los valores de diámetro observados en pesos estimados.

Diversos autores estudiaron la relación entre el diámetro y el peso en frutos, en manzano y peral (Bramardi, 1995), en naranjo dulce Valencia late (Avanza, 2011), en tangor Murcott, Otikzu y limonero Eureka (Bóbeda *et al.*, 2013) y todos encontraron que el modelo potencial que se presenta a continuación fue el que mejor ajusta esa relación y permite estimar el peso a partir del diámetro ecuatorial.

## **4.2. MATERIALES Y MÉTODOS**

### **4.2.1. Datos experimentales**

Para la definición de las curvas de crecimiento y la elaboración de las tablas de crecimiento para diferentes tamaños comerciales de frutos de tangor Murcott, se trabajó con información de campo obtenida en algunos de los huertos descritos en el Capítulo II, todos ubicados en la provincia de Corrientes y que se indican en la Tabla 4.2.



Tabla 4.2. Identificación de huertos, temporadas, fechas de plena floración, de cosecha y días desde plena floración a cosecha

Temporada	Huerto	Fecha plena flor	Cosecha	Días a cosecha
2010/11	1.1	26/01/2010	8 y 13/08/2011	320
2015/16	1.4	31/09/2015	8/10/2016	314
	1.6	31/09/2015	18/08/2016	323
2010/11	2.1	17/09/2010	07/08/2011	326
2011/12		23/09/2012	07/07/2012	287
2013/14		30/08/2013	01-02/07/2014	305
2010/11	3.1	23/09/2010	10-11/08/2011	322
2013/14		18/08/2013	25/09/2014	305
2015/16	3.2	26/09/2015	24/09/2016	362
<i>Días promedio a cosecha</i>				325

En el perímetro de cada lote se descartó un borde compuesto por las tres primeras filas y en el resto se seleccionaron 10 plantas sobre las dos diagonales de este, mediante un muestreo sistemático con arranque aleatorio. En cada árbol se identificaron 30 frutos, distribuidos en toda la copa del árbol, a los que se realizó el seguimiento del diámetro ecuatorial (mm) empleando calibre digital, con una frecuencia de 10 a 15 días desde noviembre-diciembre de cada año (finalización de caída de frutas), hasta el momento de la cosecha. Se inició con un conjunto de 3900 frutos, de los cuales algunos se perdieron por caída y otros se descartaron por estar fuera del rango de diámetros aceptado por la Unión Europea, quedando un número final de 2176 frutos. Las mediciones fueron referidas a días desde la fecha de plena floración (DDPF), que corresponde al 80% de flores abiertas.

Las tablas se construyeron en función del tamaño comercial que rige en la Unión Europea expresada en milímetros y a través de la relación funcional peso-diámetro se estableció el rango de peso y diámetro de los frutos en cajas de 10 y 15 kgs según la cantidad de fruta contenida en los mismos (tamaño comercial expresado en número de frutos por caja). En la tabla 4.3 se presenta la caja utilizada para tangor Murcott en el mercado externo (con destino a la Unión Europea).

Tabla 4.3. Rangos de diámetro ecuatorial (mm) mínimo, máximo y medio, para cada tamaño comercial.

Tamaño Comercial	Diámetro (mm)		
	Mínimo	Máximo	Medio
1xxx	78	83	80,5
1xx	73	78	75,5
1x	68	73	70,5
1	63	68	65,5
2	58	63	60,5
3	54	58	56,0
4	50	54	52,0
5	46	50	48,0

Para la confección de las curvas de crecimiento se clasificaron los frutos dentro de 8 categorías de calibre comercial, establecidas por la Unión Europea en función del tamaño alcanzado al momento de la cosecha, cabe resaltar que dentro de estas categorías solamente se consideraron los calibres que utilizan los empaques de citrus en Argentina. Se categorizaron los frutos de acuerdo con el tamaño final alcanzado en cada temporada y huerto, y se estimaron para cada uno de esos grupos los parámetros ( $\alpha$ ,  $\beta$ ,  $\gamma$ ) del modelo de crecimiento sigmoideal seleccionado quinta reparametrización del modelo Logístico.

Dado que la variable respuesta a lo largo de las mediciones sucesivas fue el diámetro ecuatorial, para establecer la producción en kilos o toneladas, se debió establecer la relación funcional entre el diámetro ecuatorial de cada fruto y su peso, para lo cual, en cada uno de los árboles seleccionados se tomó una muestra de 50 frutos al momento de cosecha y se registró el diámetro ecuatorial (mm) con calibre digital y el peso (g) con balanza digital.

#### 4.2.2. Análisis Estadístico

Con el fin de definir las curvas de crecimiento, para cada tamaño comercial se ajustó el modelo Logístico en su quinta reparametrización (L5 según Ratkowski, 1983), que se

ha definido como mejor modelo para describir el crecimiento de frutos de tanger Murcott en el Capítulo III.

$$y = \frac{1}{\alpha + \exp(\beta)\gamma^x}$$

Definidas las curvas para cada tamaño comercial, se realizaron las estimaciones de diámetro de frutos, a intervalos de 10 días a partir de los 120 DDPF hasta pasado el momento promedio de cosecha (325 DDPF), construyendo las tablas de crecimiento para los diferentes tamaños comerciales de frutos.

Para evaluar la precisión del modelo, se calcularon los errores porcentuales (% Error) entre las diámetros estimados y los reales de la siguiente manera:

$$\% \text{ Error} = \frac{\sum_i |\hat{\mu}_i - y_i|}{\sum_i y_i} \cdot 100$$

Donde  $\hat{\mu}_i$  es la predicción es la predicción del diámetro del fruto para el *i-ésimo* DDPF y  $D_{real_i}$  es el diámetro de fruto medido a lo largo del tiempo DDPF.

Para establecer la relación entre el diámetro y el peso de los frutos al momento de la cosecha se recurrió al modelo potencial:

$$y = a * x^b$$

Donde: *y es el peso de frutos (g)*; a y b son parámetros del modelo; x es el diámetro ecuatorial (mm).

El cumplimiento de los supuestos se verificó mediante un análisis de residuales, la normalidad a través del gráfico Q-Qplot y el análisis de Correlación entre los cuantiles de una distribución normal y los observados en los residuales; y la homogeneidad de varianza mediante un gráfico de residuales versus predichos.

Tomando como referencia las cajas de 10 y 15 kg, se establecieron los rangos de pesos correspondientes a cada tamaño comercial y posteriormente, empleando el modelo potencial descripto, se transformaron los diámetros en pesos, por lo cual se procedió a un ajuste de *curvas medias*.

Los datos fueron analizados mediante el paquete estadístico SAS University Edition (SAS, 2019).

## 4.3. RESULTADOS Y DISCUSIÓN

### 4.3.1. Tablas de Crecimiento

De los 3.900 frutos que se marcaron inicialmente para el seguimiento del crecimiento, un poco más un tercio sufrió caída natural antes del momento de recolección (1.391 frutos). De los restantes, en el momento de la cosecha 34 frutos superaron el tamaño comercial máximo ( $1xxx = 83 \text{ mm}$ ) y 299 no alcanzaron el tamaño mínimo ( $5 = 46 \text{ mm}$ ), por lo que no fueron considerados en los análisis.

Si bien la primera fase no se ha podido observar completa por la fecha de inicio de las mediciones, se ha podido establecer que los frutos agrupados por tamaño comercial a cosecha en tangor Murcott, siguen una curva sigmoideal clásica al igual que los frutos individuales.

En la Tabla 4.4 se presentan las estimaciones de los parámetros para el modelo de crecimiento planteado en cada uno de los calibres comerciales establecidos de acuerdo con la normativa de la Unión Europea, y los valores correspondientes a la asíntota superior (tamaño final de los frutos), ordenada al origen (tamaño inicial) y tasa de crecimiento. Las estimaciones de los parámetros  $\alpha$ ,  $\beta$  y  $\gamma$  difirieron significativamente ( $p < 0,001$ ) entre huertos y temporadas.

Tabla 4.4. Estimaciones de los parámetros del modelo logístico L5, para cada tamaño comercial según Unión Europea.  $\alpha$ ,  $\beta$  y  $\gamma$  parámetros del modelo, asíntota superior, ordenada al origen (intercepto), tasa de crecimiento y número de frutos por rango de calibre

Tamaño Comercial	Parámetros estimados			Características de la curva			n
	$\alpha$	$\beta$	$\gamma$	Asíntota	Intercepto	Tasa Crecimiento	
1xxx	0,01181	-2,7953	0,9861	84,64	13,71	0,0139	114
1xx	0,01272	-2,6928	0,9850	78,58	12,43	0,0150	375
1x	0,01360	-2,5605	0,9838	73,49	11,00	0,0163	646
1	0,01466	-2,4381	0,9819	68,20	9,80	0,0182	542
2	0,01601	-2,2591	0,9816	62,44	8,30	0,0182	299
3	0,01756	-1,7722	0,9778	56,92	5,33	0,0224	120
4	0,01905	-1,4849	0,9745	52,48	4,07	0,0257	48
5	0,02067	-1,4798	0,9749	48,36	4,02	0,0253	29
							2173

El tamaño comercial 1xxx presenta la mayor asíntota superior (tamaño final de los frutos) y el mayor intercepto (tamaño inicial) y se observa que estos valores decrecen a medida que los calibres comerciales disminuyen, no obstante, las tasas de crecimiento tienen un comportamiento inverso y son menores en frutos grandes. Estos resultados indican que los frutos que presenten tamaños pequeños al principio de la temporada terminarán el periodo de crecimiento con tamaños pequeños, en términos relativos, independientemente de que la tasa de crecimiento de estos frutos sea superior a la de los frutos más grandes.

Estos resultados difieren de los encontrados por Bramardi (1995) en manzanas, quién estableció que las tasas de crecimiento era iguales para todos los tamaños comerciales; y por Zadabrec *et al.* (2013), quienes indicaron que el crecimiento de frutos en cultivares de manzanas de maduración temprana y tardía, tuvo un tasa de crecimiento uniforme, además establecieron que la principal influencia en el tamaño final del fruto dependía del cultivar y que existía grandes diferencias entre los años analizados.

Los resultados obtenidos en esta tesis coinciden con los de Khurshid & Braysher (2009), en relación con sus hallazgos de que los tamaños iniciales de los frutos eran diferentes en la temporadas evaluadas, pero no concuerdan en que la tasa de crecimiento a lo largo del período se mantenía constante. Coinciden con los de Koch *et al.* (1996) que establecieron para mandarinas Satsuma y Clementina un crecimiento

lineal después de la caída fisiológica del fruto, y que los frutos grandes continuaron siendo grandes y los pequeños, pequeños.

En razón de las diferencias encontradas en las estimaciones de los parámetros entre huertos y temporadas, es de interés contar con curvas que representen el comportamiento promedio de los frutos, es por ello que se consideraron curvas medias para cada calibre comercial, las que se presentan en la Figura 4.1.

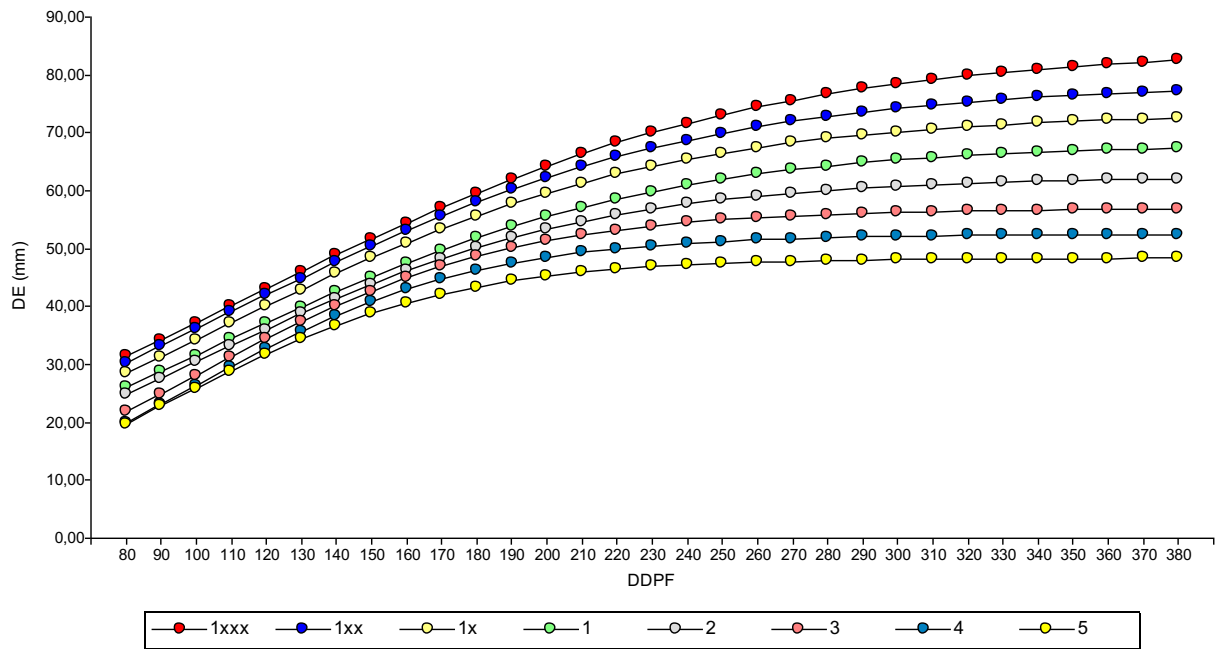


Figura 4.1. Curvas medias de crecimiento de frutos tanger Murcott por calibre comercial, según normas de comercialización de la Union Europea.

Debido a que el crecimiento de los frutos de tanger Murcott está influenciado por diversos factores ambientales y de manejo cultural y dado que las tablas de crecimiento del diámetro ecuatorial de frutos en función de los días después de plena floración desarrolladas en este trabajo fueron generadas a partir de información de diferentes huertos y temporadas, pero corresponden a comportamientos medios, para mejorar el grado de precisión en los pronósticos de diámetros de frutos a la cosecha o incorporar otras variedades de mandarina o de otras especies cítricas, es fundamental ajustar los modelos desarrollados en este estudio para cada variedad y zona.


Las tablas de crecimiento que se presentan en la Tabla 4.5 expresan el diámetro ecuatorial medio (crecimiento medio), que los frutos presentan a lo largo de los días después de la plena floración. Estos diámetros se encuentran referenciados a los 320

días después de plena floración, es decir al momento de cosecha promedio en Corrientes, Argentina, dado que tangor Murcott es una variedad cítrica tardía, aunque su comercialización depende de los precios del mercado interno y externo.

Estas tablas pueden utilizarse para predecir anticipadamente la distribución de tamaños que los frutos en un huerto tendrán al momento de la cosecha. Para ello bastará determinar los DDPF del momento que se desea realizar la estimación (120 por ejemplo), tomar una muestra de frutos, registrar sus diámetros (37 mm por ejemplo), entrar en la tabla en la columna que corresponda a la fecha (en DDPF), ubicar la fila que más se aproxima al diámetro de cada fruto (cuarta fila identificada con el calibre 1, por ejemplo) y desplazarse por esa fila hasta la columna correspondiente a los DDPF en que se tiene planificado cosechar, registrando el diámetro estimado de cada fruto (66,09 en este caso).

Tabla 4.5. Tabla de crecimiento de frutos de tanger Murcott por calibre comercial de acuerdo con las normas de comercialización de la Union Europea  
Incremento del diámetro ecuatorial (mm) en función de los días después de plena floración (DDPF), TC: tamaño comercial según Unión Europea

Período de Cosecha



DDPF																									
TC	120	130	140	150	160	170	180	190	200	210	220	230	240	250	260	270	280	290	300	310	320	330	340	350	360
1XXX	42,97	45,91	45,91	51,67	54,43	54,43	59,61	61,99	64,23	64,23	68,24	70,00	71,62	73,08	74,40	75,59	76,66	77,61	78,46	79,22	79,88	80,47	80,99	81,45	81,85
1XX	41,94	44,86	47,73	50,50	53,15	55,67	58,03	60,24	62,27	64,13	65,82	67,35	68,72	69,95	71,04	72,01	72,86	73,61	74,27	74,84	75,34	75,78	76,16	76,49	76,78
1X	40,01	42,89	45,70	48,40	50,97	53,38	55,63	57,70	59,58	61,30	62,83	64,21	65,43	66,50	67,45	68,28	69,00	69,63	70,17	70,65	71,05	71,40	71,70	71,96	72,19
1	37,09	39,83	42,49	45,05	47,48	49,76	51,87	53,81	55,58	57,17	58,60	59,87	61,00	61,98	62,85	63,60	64,25	64,82	65,31	65,73	66,09	66,40	66,66	66,89	67,09
2	36,02	38,74	41,36	43,82	46,11	48,21	50,11	51,81	53,32	54,65	55,81	56,82	57,68	58,42	59,06	59,60	60,05	60,44	60,76	61,04	61,27	61,46	61,62	61,76	61,87
3	34,35	37,32	40,09	42,62	44,88	46,87	48,60	50,07	51,31	52,34	53,20	53,91	54,49	54,96	55,34	55,65	55,90	56,10	56,27	56,40	56,50	56,59	56,65	56,71	56,75
4	32,76	35,71	38,42	40,82	42,92	44,72	46,22	47,47	48,49	49,32	49,98	50,51	50,93	51,27	51,53	51,74	51,90	52,03	52,13	52,21	52,27	52,32	52,35	52,38	52,41
5	31,71	34,36	36,74	38,83	40,63	42,13	43,39	44,41	45,24	45,90	46,43	46,85	46,85	47,44	47,64	47,80	47,93	48,02	48,10	48,16	48,20	48,24	48,26	48,29	48,30



La confiabilidad de las tablas predictivas de diámetro se basa en estudios previos como los de Gil-Albert (1992), quien indicó que el tamaño relativo de los frutos se conserva desde que estos cuajan hasta la cosecha, lo que es la base para la construcción de tablas de pronóstico de cosecha. Ortega-Farias *et al.* (2002) por otra parte, sugirieron que es posible construir una tabla de predicción de tamaños en manzana cv Granny Smith, en la zona de San Fernando, VI Región Chile, tomando como referencia la ecuación logística y grados días acumulados como variable regresora. Avanza (2011), a través de la ecuación logística en su 5° reparametrización referenció los diámetros ecuatoriales a calibres comerciales de frutos recolectados a la cosecha, e indicó que las tablas de crecimiento pueden utilizarse como una guía en general, pero se debe tener en cuenta la variabilidad existe entre huertos y temporadas.

Teniendo en cuenta los antecedentes mencionados y los resultados encontrados en este estudio, con los modelos desarrollados en cada rango de diámetro de frutos de tangor Murcott, fue posible elaborar una tabla de predicción de crecimiento en función de los días después de plena floración (DDPF). Lográndose porcentajes de error, expresados como cantidad relativa de fruta mal clasificada, que variaron entre -0,24 a 20,94% lo largo del periodo de crecimiento.

Para lograr una adecuada precisión de la distribución de calibres que se tendrá a cosecha, en el Capítulo V se determinará el momento adecuado de realizar las estimaciones.

En el Alto Valle<sup>13</sup>, Argentina, realizan las predicciones de la distribución de tamaños comerciales (referidos a un envase de 18 kg), para manzanas y peras en 10 categorías, pero como los calibres extremos son muy poco frecuentes, para reducir los errores de estimación se reduce a 4 categorías de tamaños comerciales: 80 o menos, 90-100, 110-120 y 135 o mayores.

En la Figura 4.2 se presenta la evolución de los errores en el tiempo, donde se ve claramente que a medida que se acerca la fecha de cosecha disminuyen los errores de estimación.

---

<sup>13</sup> <https://www.cafi.org.ar/wp-content/uploads/2015/01/Presentacion-2014.pdf>

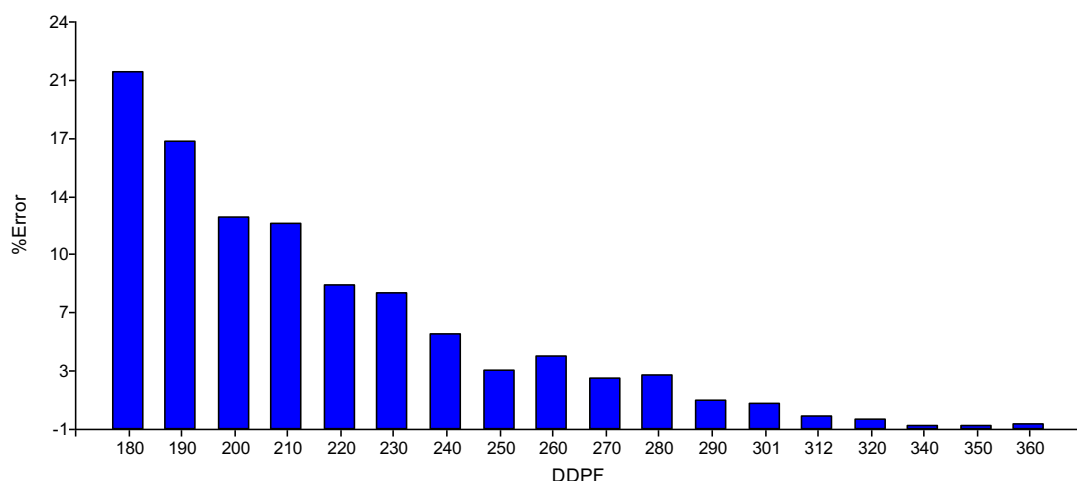


Figura 4.2. Porcentajes de fruta mal clasificada en su tamaño comercial de tangor Murcott en función de los días después de floración (DDPF)

### 4.3.2. Relación diámetro-peso

En algunas ocasiones resulta de interés conocer las distribuciones de tamaño de frutos expresadas en unidades de peso, pero si las mediciones de tamaño de los frutos se realizaron en función de los diámetros de estos, se requiere disponer de una herramienta que permita transformar los diámetros en peso, para lo que se evaluó la función que describiera de la mejor manera dicha relación. También esto es un requerimiento cuando se quiere hacer estimaciones de producción expresada en unidades de peso. Mediante el análisis de regresión no lineal, la relación entre los diámetros y los pesos de los frutos de tangor Murcott estudiados pudo ser ajustada a un modelo potencial, cuyos parámetros y sus estimaciones se describen en la Tabla 4.6, la representación gráfica de los valores observados y predichos según ese modelo se representan en la Figura 4.3.

Tabla 4.6. Estimación de los parámetros del modelo potencial, error estándar y p-valor.

Parámetros	Estimación	Error estándar	p-valor
a	0,00186	0,00011	<0,0001
b	2,66435	0,01402	<0,0001

Los parámetros estimados del modelo potencial resultaron altamente significativos ( $p < 0,0001$ ). Valores similares fueron obtenidos por Avanza (2011) para frutos de naranjo dulce Valencia late con parámetros estimados  $a = 0,00218$  y  $b = 2,67$ .

En la Figura 4.3 se presentan los valores observados y predichos en función del modelo potencial propuesto.

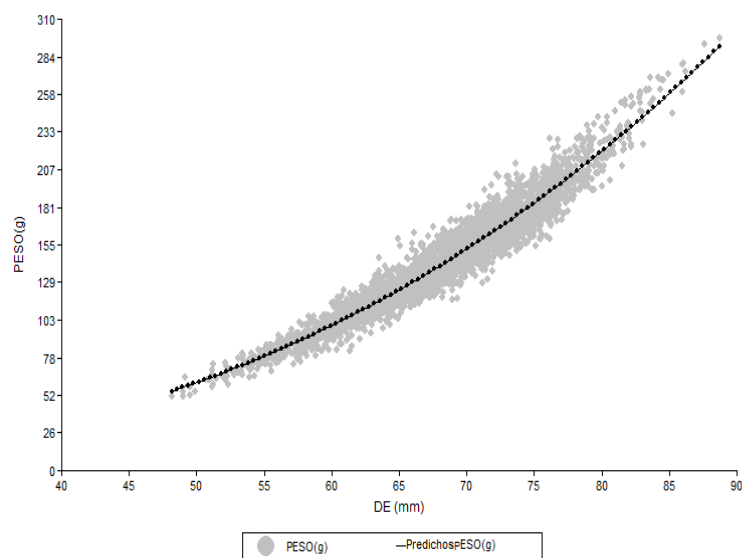


Figura 4.3. Valores observados y predichos del peso (g) en función del diámetro ecuatorial (mm) en frutos de tanger Murcott.

A fin de analizar los supuestos de normalidad y homogeneidad de varianza se confeccionaron gráficos Q-Qplot (Figura 4.4) y de residuales versus predichos (Figura 4.5). Del Q-Qplot se desprende un adecuado ajuste a la distribución normal ( $r = 0,999$ ) y la gráfica de los residuales versus predichos no indica algún patrón que evidencie heterogeneidad de varianza, por lo que ambos supuestos se consideran cumplidos.

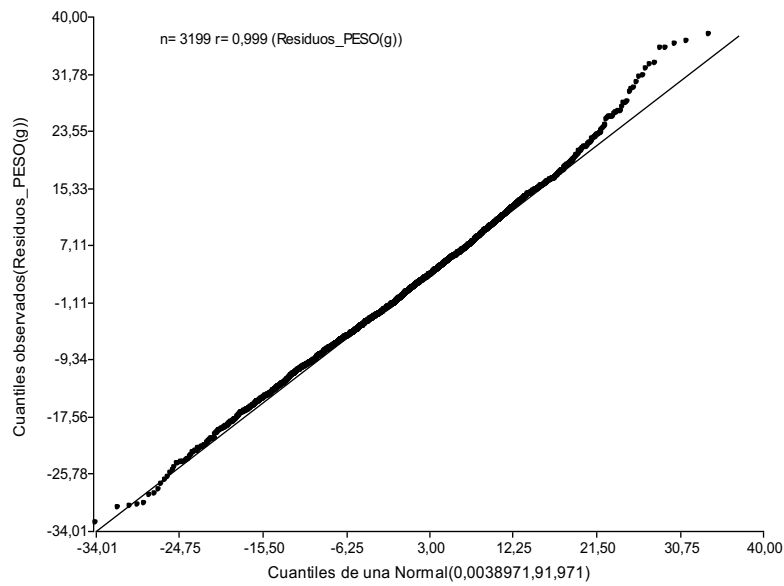


Figura 4.4. Ajuste a la distribución normal mediante Q-Qplot

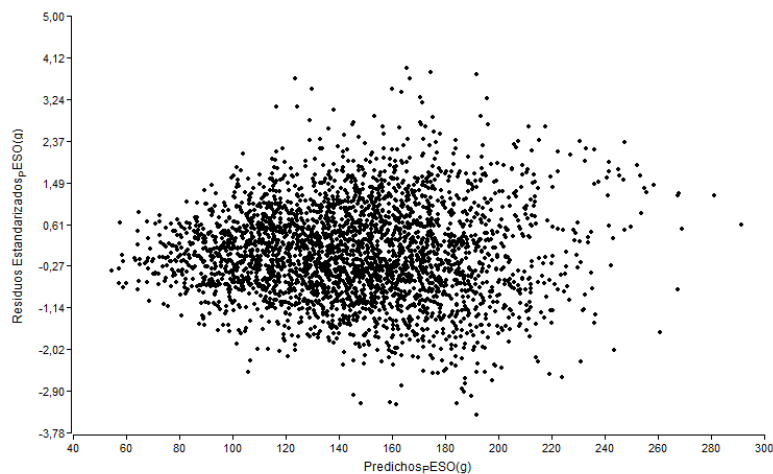


Figura 4.5. Gráficos de residuales en función a predichos del modelo potencial seleccionado para la relación peso (g)- diámetro (mm).

En un avance de la tesis, Bóveda *et al.* (2013) estudiaron la relación diámetro peso en frutos de tangor Murcott durante la temporada 2010/11 en una muestra de frutos de tangor Murcott y mandarina Okitsu, durante la temporada 2011/12 y limón Eureka, durante las temporadas 2010/11 y 2011/12, en huertos ubicados en las localidades de Saladas y Bella Vista, Corrientes, para lo que también ajustaron modelos potenciales con parámetros estimados  $a = 0,0037$  y  $b = 2,73$  para tangor Murcott,  $a = 0,00089$  y  $b = 2,83$  para mandarina Okitsu y  $a = 0,00018$  y  $b = 2,72$  para limón Eureka.

Las tablas de crecimiento, descritas en la Sección 4.3.1, expresan el diámetro ecuatorial medio de los frutos en función de los días después de floración. Estos diámetros se encuentran referenciados a tamaños comerciales de frutos recolectados al momento de cosecha (325 DDPF en promedio para la provincia de Corrientes).

En la Figura 4.6 se presenta los diámetros de los tamaños comerciales que rige la Unión Europea, con el fin de expresar los tamaños en unidades de peso, se utilizó el modelo potencial (relación diámetro-peso), estableciendo los pesos que corresponden a cada tamaño comercial. Posteriormente para cada caja de 10 y 15 kg (envases utilizados en los empaques de Argentina), considerando el número de frutos por caja, se establecieron los rangos de pesos correspondiente a cada caja y se determino los tamaños comerciales regidos por la Unión Europea de acuerdo con el tipo de caja a comercializar.

La tabla de crecimiento de frutos en función de los calibres comerciales presentada en este trabajo debe ser considerada como una guía general, debido a que el crecimiento de los frutos de tangor Murcott está influenciado por diversos factores internos y externos (ambientales y manejos culturales).

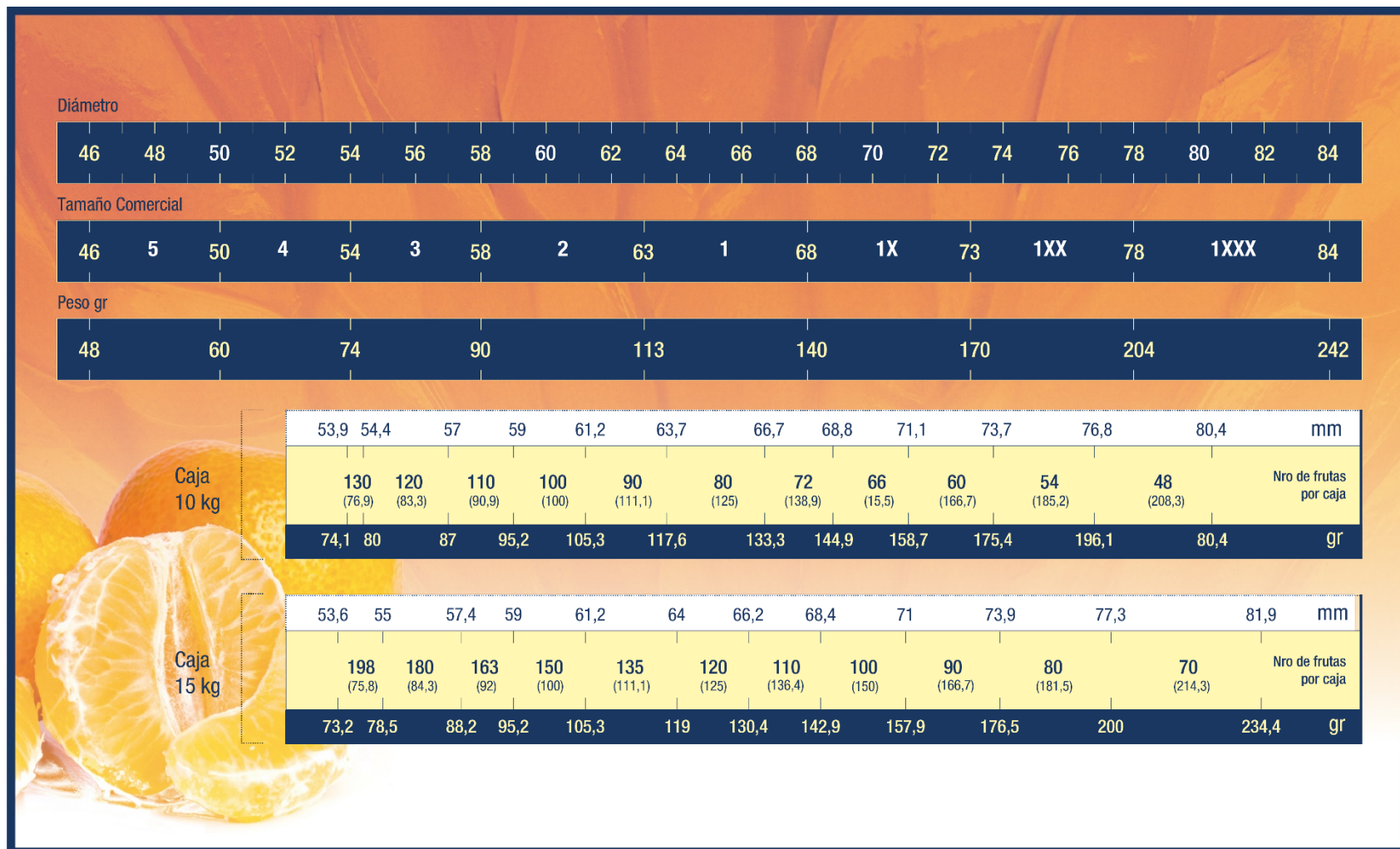


Figura 4.6. Relación diámetro-peso, tamaños comerciales, número de frutos por cajas de 10 y 15 kg para tangor Murcott.

### 4.3.3. Uso de tablas de crecimiento para raleo de frutos

Una de las técnicas más comunes en los huertos de mandarina es el raleo de frutos. Este raleo puede realizarse de diferentes maneras, en la región se utiliza principalmente el raleo manual y se realiza a finales de noviembre y principios de diciembre, entre los 100 a 130 días después de la plena floración (Otero, 2004).

La forma de realizar el raleo manual difiere para las distintas variedades y costumbres. Para realizar la remoción de frutos del árbol de manera temprana, en algunos casos se tienen en cuenta distancias mínimas entre frutos (se remueven frutos que se encuentran a menos de una distancia mínima establecida); en otros, se considera un determinado porcentaje de frutos (se remueve un porcentaje establecido de los frutos presentes); y en otros casos se define por los tamaños de los frutos (se remueven todos los frutos con tamaños menores o mayores a valores establecidos).

Teniendo en cuenta la tabla de crecimiento presentada anteriormente es posible realizar un raleo manual de frutos en tangor Murcott teniendo en cuenta sus tamaños, de manera tal que los frutos que permanezcan en la planta se correspondan a tamaños comerciales, con diámetros ecuatoriales comprendidos entre los calibres establecidos para comercialización a algún mercado de exportación como, por ejemplo, la Unión Europea.

Suponiendo que se desea identificar frutos que presentarán tamaños iguales o mayores al calibre 3 definido por la Unión Europea al momento de la cosecha, se deben seguir los siguientes pasos:

1. Ingresar por la columna que corresponde a los días después de plena floración del momento de cosecha promedio o, si ya se dispone de una fecha estimada de cosecha para el lote en cuestión, ingresar con esos DDPF. En este caso se considera el momento promedio de cosecha a los 320 DDPF, producto de contar los días que suceden a la plena floración (en el mes de septiembre) de un año en particular (2020) al mes de agosto del siguiente año (2021).

2. En dicha columna localizar el diámetro ecuatorial buscado (o aproximado). Si se desea cosechar frutos en la categoría 3, estos frutos tendrían al momento de la cosecha un diámetro promedio de 56,50 mm. Identificada la fila que lo contiene, se debe mover en esa fila hacia la izquierda, hasta la columna que corresponde al momento en que se va a realizar el raleo (entre 100 y 130 DDPF), por ejemplo, la de 120 DDPF.

3. En esa intersección se obtiene el diámetro ecuatorial que debe poseer un fruto en esa fecha (suponiendo 120 DDPF) que corresponde a 34,35 mm, para llegar a un diámetro de 56,50 mm al momento de cosecha. Frutos con diámetro menor al mencionado tendrán, al momento de la cosecha, tamaños inferiores a los establecidos para el calibre 3 y no podrán ser comercializados en esa categoría de la Unión Europea, por lo que deben ser removidos.

4. Una vez establecido el diámetro, se deberá utilizar la tabla de relación peso-diámetro según los tamaños comerciales Figura 4.6. Por ejemplo, para una caja de 10 kg y un diámetro medio de 56,5 mm corresponde a un rango de diámetros de frutos de 73-76 mm, por lo tanto los frutos tendrán un peso promedio de 185,2 g y la cantidad de frutos por caja será de 54.



#### 4.4. RESUMEN DE RESULTADOS

A partir del modelo de crecimiento sigmoideo L5 descrito en el Capítulo III se han podido definir, para los frutos de tangor Murcott, los parámetros específicos para cada uno de los calibres comerciales establecidos por la Unión Europea para comercialización de frutos de mandarinas.

El tamaño comercial 1xxx presenta la mayor asíntota superior (tamaño final de los frutos) y el mayor intercepto (tamaño inicial) y se observa que estos valores decrecen a medida que los calibres comerciales disminuyen, no obstante, las tasas de crecimiento tienen un comportamiento inverso y son menores en frutos grandes. Estos resultados indican que los frutos que presenten tamaños pequeños al principio de la temporada terminarán el periodo de crecimiento con tamaños pequeños en términos relativos, independientemente de que la tasa de crecimiento de estos frutos sea superior a la de los más grandes.

Debido a las diferencias halladas en las estimaciones de los parámetros entre huertos y temporadas, se han definido curvas que representen el comportamiento promedio de los frutos, es por ello que se consideraron curvas medias para cada calibre comercial.

Se han desarrollado tablas de crecimiento que expresan el diámetro ecuatorial medio (crecimiento medio), que los frutos presentan a lo largo de los días después de la plena floración. Estos diámetros se encuentran referenciados a los 320 días después de plena floración, es decir al momento de cosecha promedio en Corrientes, Argentina, dado que tangor Murcott es una variedad cítrica tardía, aunque su comercialización depende de los precios del mercado interno y externo.

Estas tablas pueden utilizarse para predecir anticipadamente la distribución de tamaños que los frutos en un huerto tendrán al momento de la cosecha. Para ello bastará determinar los días después de plena floración que corresponden al momento que se desea realizar la estimación, tomar una muestra de frutos, registrar sus diámetros, entrar en la tabla en la columna que corresponda a la fecha (en días después de plena floración), ubicar la fila que más se aproxima al diámetro de cada fruto y desplazarse por esa fila hasta la columna correspondiente al momento de cosecha

promedio (o al momento en que se tiene pensado cosechar si esa fecha ya está establecida), registrando el diámetro estimado de cada fruto.

En algunas ocasiones resulta de interés conocer las distribuciones de tamaño de frutos expresadas en unidades de peso, pero si las mediciones de tamaños de los frutos se realizaron en función de los diámetros de los mismos, se requiere disponer de una herramienta que permita transformar los diámetros en peso, para lo que se evaluó la función que describiera de la mejor manera dicha relación. Mediante el análisis de regresión no lineal, la relación entre los diámetros y los pesos de los frutos de tangor Murcott estudiados pudo ser ajustada a un modelo potencial.

A fin de expresar el tamaño en unidades de peso, mediante el modelo potencial descrito, se establecieron los rangos de pesos correspondiente a cada tamaño comercial y transformaron los diámetros en pesos, considerando el número de frutos para cajas de 10 y 15 kg, que son los envases utilizados en los empaques de Argentina.

Se ha definido y explicado un procedimiento que permite aplicar las tablas de crecimiento desarrolladas para el raleo de frutos en tangor Murcott.

## 4.5. BIBLIOGRAFÍA

1. Agustí, M; Mesejo, C; Mesero, C, Martínez-Fuentes, A; Zaragoza, S; Primo Millo, E. 2020. El tamaño del fruto en los cítricos. *Levante Agrícola. Revista Internacional de Cítricos*. 451: 57-57.
2. Álvarez, A; Bramardi, S; Srangaferro, S. 2001. Modelos de predicción anticipada del peso medio de los frutos en producción de frutos de pepita. *X Congreso Latino Americano de Biomatemática*. Campinas, Brasil.
3. Ariza Flores, R; Tejacal, I; Beltrán, M; Ambriz Cervantes, R; Lugo Alonso, A; Barrios Ayala, A; Barbosa Moreno, F. 2010. Calidad De Los Frutos De Naranja 'Valencia' En Morelos, México *Revista Iberoamericana de Tecnología Postcosecha*. Asociación Iberoamericana de Tecnología Postcosecha, S.C. Hermosillo, México. 11 (2): 148-153.
4. Avanza, MM. 2011. Desarrollo de una propuesta metodológica para la predicción de cosecha de naranjo dulce (*Citrus sinensis* L. Osbeck) var. Valencia late en la provincia de Corrientes, Argentina. Tesis Doctoral de la Universidad Nacional del Nordeste, Área Recursos Naturales.
5. Bóbeda, G; Giménez, L; Gaiad, J; Chabbal, M; Montoya, A; Mazza, S. 2013. Relación peso-diámetro de frutos en variedad de mandarina y limón de la provincia de Corrientes. *IV Encuentro Iberoamericano de Biometría y XXVIII Reunión Científica del GAB*. Mar del Plata, Buenos Aires. 3 p.
6. Bramardi, SJ. 1995. Modelos de predicción de tamaños de fruta a la cosecha en base a mediciones sucesivas durante el período de crecimiento. Peras cvs. William'S y Packman's Triumph en la región del Alto Valle de Río Negro y Neuquén. Tesis para optar el grado de Magister Scientiae- Área Biometría. Escuela para graduados Convenio Facultad de Agronomía-UBA e Instituto Nacional de Tecnología Agropecuaria. 169 pp.
7. Bramardi, S; Tassile, V; Reeb, P; De Bernandin, F. 2005. Comparación de métodos para la estimación anticipada del peso del fruto medio a cosecha. Curvas de crecimientos vs. Modelos estocásticos. *X Reunión Científica del Grupo Argentino de Biometría*, Corrientes, Argentina. 6 p.

8. Bramardi , S; Tassile, V; Reeb, P; Bernandin, F. 2006 . Tablas de raleo 1° ed. Editor: Area de Estadística, Facultad de Ciencias Agrarias, Universidad Nacional del Comahue. 24 p. 1. Frutos-Cultivos I. ISBN 987-05-1707-2.
9. Franco, J; Gaviña, A. 2000. Early Prediction of Fruit Size at Harvest for Valencia Orange Using Experimental Data. Proceedings of the International Society of Citriculture. 456-458.
10. Fundecitrus Brasil Tree Inventory and 2019-2020 Orange Crop Forecast for Sao Paulo and West-Southwest Minas Gerais Citrus Belt. 2019. 136 pp. Disponible: <https://www.fundecitrus.com.br/>.
11. Gil, A. 1992. La ecología del árbol frutal. 2, 237. Mundi-Prensa, Madrid, España.
12. Khurshid, T; Bevington, K; Falivene, S. 2005. Citrus fruit growth model. Department of Primary Industries, State of New South Wales.
13. Khurshid, T; Braysher, B. 2009. Early Fruit Size Prediction Model Using Cubic Smoothing Splines for 'Washington Navel' (Citrus Sinensis L. Osbeck) Oranges in Australia International Journal of Fruit Science. 9 (4): 394-408.
14. Koch, N; Theron, K; Rabe, E. 1996. Predicción del tamaño del fruto en mandarina 'Clementina' y 'Satsuma' y mejora del tamaño del fruto de mandarina 'Clementina' (*Citrus reticulata* Blanco) mediante 2,4-DP. Proceedings of the International Society of Citriculture. 2,539-542.
15. Marine, R; Schupp, J; Baugher, T; Crasswellerr, R. 2019 . Estimating Apple Fruit Size Distribution from Early-season Fruit Diameter Measurements. American Society for Horticultural Science (HortScience). 54 (11): 1947-1954.
16. Mayorano, F; Rubiales, A; Herrero, V; Clausse, A. 2006. Modelo Computacional para el pronóstico de tamaño de frutas. RIA. 35 (2): 143-162.
17. Ortega-Farias, S; Flores, L; Retamales, S. 1997. Elaboración de una tabla para predecir calibre de manzanas, variedad Red spur. Revista Frutícola. 18 (1): 1-5.
18. Ortega-Farias, S; Flores, L; Retamales, S. 1998. Modelo logístico para predecir el crecimiento en diámetro de las manzanas (variedad Granny Smith). Revista Frutícola, 19: 15-18.
19. Ortega-Farias, S; Flores, L; León, L. 2002. Elaboración de una tabla predictiva de diámetro de manzanas cv. Granny Smith usando los grados días acumulados. Agricultura Técnica. 62 (4): 624-632.

20. Otero, A. 2004. Raleo de frutos en mandarina Satsuma y otros cítricos. Instituto Nacional de Investigación Agropecuaria. Serie Técnica N°140. Uruguay.
21. Otero, A; Goñi, C; Carrau, F; Zefferino, E; Brennam, M. 2005. Previsión de productividad y calidad en citrus: herramientas para una gestión sustentable. *Revista INIA*, 4: 21-24.
22. Ratkowsky, DA. 1989. Handbook of nonlinear regression models. Marcel Dekker, Inc., NY, 241 pp.
23. SAS University Edition. 2019. SAS 9.3 Programming Documentation. United States.
24. Stander, O; Cronj, P. 2016. Reviewing the Commercial Potential of Hand Thinning in Citrus with a Cost-benefit Analysis of Summer Hand Thinning of Nadorcott 'Mandarin' American Society for Horticultural Science (HortScience). 26 (2): 206-212.
25. Sutton, M; Doyle, J; Chavez, D; Malladi, A. 2020. Optimizing Fruit-Thinning Strategies in Peach (*Prunus persica*) Production. *Horticulturae*. 6 (3): 41.
26. Zadavec, P; Veberic, R; Stampar, F; Eler, K; Schmitzer, V. 2013. Fruit size prediction of four apple cultivars: Accuracy and timing. *Scientia Horticulturae*. 160: 177-181.

## CAPITULO V

### ESTIMACIÓN DE CARGA BASADA EN INFORMACIÓN PARCELARIA

#### 5.1. INTRODUCCIÓN

Teniendo en cuenta el modelo de predicción de producción propuesto en el Capítulo I, en el presente Capítulo se estudia la estimación de la componente carga (número de frutos por árbol).

Determinar la carga de manera apropiada es fundamental para lograr una adecuada predicción del volumen de producción de frutos a cosecha, por lo que se han evaluado diferentes metodologías con el fin de determinar aquella que permita una estimación del número de frutos al momento de cosecha con suficiente anticipación y precisión satisfactorias (Avanza, 2011).

La evaluación de la carga es también esencial para tomar decisiones de gestión necesarias para conseguir un adecuado tamaño de los frutos. Proporciona una indicación del tipo de cosecha que se espera (muy abundante, media o escasa) y es particularmente útil cuando se espera una carga de cultivo muy abundante, dado que permite emprender estrategias para su reducción como el raleo o aclareo, químico o manual <sup>14</sup>.

Verreyne (2009) indicó que el número de frutos por árbol en el momento de la cosecha viene determinado tanto por la intensidad de la floración como por el cuaje de frutos alcanzado. Debido al aborto flores y pequeños frutos durante el período de caída fisiológica de frutos, la estimación de la carga para estimación de la producción sólo debe hacerse después de pasado dicho período. No obstante, entre los momentos de

---

<sup>14</sup> <https://www.agric.wa.gov.au/citrus/estimating-your-citrus-crop-load>

caída fisiológica de los frutos y la cosecha, en algunas variedades se produce una caída de frutos de gran proporción, denominada período de caída de frutos de verano-otoño, especialmente fuerte en el caso de los naranjos de ombligo, lo que constituye una fuente de error en las estimaciones.

La predicción de la carga es también esencial para gestionar el tamaño de los frutos. Aunque hay muchos factores que influyen en esta variable, el número de frutos por árbol desempeña un papel importante. El rendimiento de un huerto frutal viene determinado tanto por el número de frutos producidos por árbol como por el tamaño de éstos. Los árboles con mayor número de frutos dan lugar a frutos más pequeños y, si se identifican a principios de la temporada, se pueden llevar a cabo estrategias de reducción del número de frutos, como el aclareo. Esa reducción en esta fase no sólo puede dar lugar a frutos más grandes, sino que puede reducir la incidencia de la alternancia entre años (Racsko, 2006).

La carga en los cítricos viene determinada en gran medida por la intensidad y el tipo de floración, así como por los factores ambientales que afectan al cuajado de los frutos. En los cítricos, el número de frutos está inversamente relacionado con el tamaño de éstos, lo que significa que cuanto mayor sea el número de frutos, menor será su tamaño. La carga es la principal responsable de la variación anual del tamaño de los frutos (Otero *et al.*, 2005).

La estimación de la carga puede realizarse de diferentes maneras, las más empleadas son las mediciones parcelarias, que se detallan a continuación.

### **5.1.1. Mediciones parcelarias**

La estimación de carga a partir de información obtenida en los propios huertos es muy utilizada cuando se desean hacer previsiones a nivel local o parcelario. En general se realizan recuentos sobre una muestra de árboles preseleccionados, a través de diversa metodología, que tiene en cuenta la totalidad de la copa o secciones predefinidas de la misma (Anderson *et al.*, 2021)

### 5.1.1.1. Recuento total

Diferentes técnicas se han desarrollado y son utilizadas actualmente para la estimación de carga de distintas especies frutales. Verreyne (2009) propone el método del recuento total a través de dos mecanismos:

- ✓ Contar el número de frutos por árbol de uno o más árboles representativos por huerto o unidad aplicando una pegatina a cada fruto y posteriormente contar las pegatinas utilizadas por árbol.
- ✓ Eliminar totalmente de la fruta de uno o más árboles representativos después del período fisiológico de caída de la fruta y contar esos frutos. Si se determina, además, el diámetro de los frutos es posible calcular la distribución de tamaños, a partir de la cual es posible realizar una estimación más aproximada de la producción que se obtendrá.
- ✓ El recuento visual de los frutos del árbol, en general ayudados con un contador manual, por parte de relevadores experimentados. En este método es frecuente incurrir en errores por fruto oculto por lo que se pueden establecer empíricamente factores de corrección obtenidos de la relación entre frutos contados y frutos realmente presentes en el árbol.

En el programa de pronóstico de cosecha para frutales de pepita en Río Negro y Neuquén, Argentina<sup>15</sup> se utiliza actualmente el recuento visual de frutos, realizado según los años, entre 30 y 90 días previos al momento del inicio de cosecha según el cultivar. En todos los casos los valores de carga registrados por los relevadores son corregidos por un coeficiente de fruto oculto que actúa como divisor y que varía entre 0,85 y 0,90 de acuerdo con la variedad en cuestión.

Para el pronóstico anual de cosecha en frutos de pepita y carozo de la provincia de Mendoza, Argentina, se han evaluado diferentes metodologías de recuento de frutos

---

<sup>15</sup> <https://rionegro.gov.ar/download/archivos/00000253.pdf>



entre las que se encuentra el recuento total, que es empleado para la mayoría de las especies en el programa de pronóstico de cosecha (Gil, 1996, 2003).

El Departamento de Agricultura de los EUA (USDA) realiza estimaciones manuales de carga de frutas para pronósticos de cosecha en varios cultivos. Las previsiones de carga en almendros se basan en el muestreo de cultivos dos meses antes de la cosecha y han tenido, una tasa de error promedio del 7,3%, con un máximo del 13,8%. En el caso de las avellanas, basado en una encuesta de 180 huertos y dos árboles seleccionados al azar por huerto, la tasa de error promedio de las previsiones de la cosecha durante los últimos años fue del 8,1% (Olsen *et al.*, 2001).

En el caso de los huertos cítricos, Lacey (2019) recomienda conteos de frutos en una muestra de 60 árboles por bloque de cítricos, con muestreo de al menos dos lados del árbol, estimado el tiempo para este procedimiento en 60 min por bloque.

Avanza (2011), a fin de determinar la carga en árboles de naranjo Valencia late, empleó el recuento total del número de frutos por árbol mediante un contador manual, entre dos operarios, que corrigió posteriormente con un coeficiente por fruto oculto calculado para cada huerto.

#### *5.1.1.2. Recuento en secciones predefinidas de copa*

En la provincia de Mendoza, Argentina, para el pronóstico anual de cosecha en frutos de pepita y carozo, además del recuento total, se emplea el recuento parcial por rama primaria, el recuento a través del método del camino aleatorio desarrollado para predicción de cosechas en EUA por Huddleston (1978), y el método propuesto por Winter (1969), en el cual se estima la densidad de carga de un árbol a partir de una sección predefinida de copa de  $0,36 \text{ m}^2$ , y luego para el cálculo del número total de frutos participa el potencial del árbol, definido en base a sus dimensiones. No obstante, debido al manejo heterogéneo de la mayoría de las plantaciones, la carga por árbol presenta gran variabilidad y la distribución de los frutos en el árbol varía entre los diferentes sectores, lo que dificulta la definición de un método de estimación basado en el recuento de secciones de copa (Gil, 1996, 2003).

En cerezos, existen trabajos sobre la relación entre los índices de carga y la calidad de la fruta (peso, contenido de sólidos solubles y color). Tersoglio *et al.* (2004), en la provincia de Mendoza, Argentina, realizaron un estudio con los objetivos de establecer las relaciones cuantitativas entre los diferentes índices de carga y el tamaño de los frutos, e identificar el mejor modelo, relacionar el índice de carga con el color de los frutos y los sólidos solubles, analizar la distribución de los tamaños de frutos y estimar la fracción de frutos exportable para cada nivel de carga. Los resultados mostraron una importante correlación entre el índice de carga, el peso medio de los frutos y los contenidos de sólidos solubles; y una relación inversa entre la carga y la fracción de frutos de mayor tamaño.

En la producción australiana de mango, la mejor práctica comercial actual para la estimación de la carga de frutos de los huertos implica el recuento visual de frutos por árbol sobre el 5% de los árboles del huerto a través de un recuento de árboles frutales en transectos. En un ejemplo de producción de mango brasileño, el 5% de los árboles en el bloque se cuentan a lo largo de una línea de transecto, y el recuento se aumenta si la desviación estándar es mayor al 10% de la media (Anderson *et al.*, 2021).

Para frutos cítricos, de Rezende Triboni & Barbosa (2004) evaluaron un método para estimar el número de frutos por árbol, basándose en cuantificar de manera no destructiva el número de frutos en ramas terminales de tamaño fijo. Analizaron los datos de manera estratificada, utilizando como factores de estratificación: cultivares de naranja dulce y tres categorías de edades en diversos sitios del cinturón cítrico paulista en Brasil. Encontraron una correlación lineal significativa entre el número total de frutos cosechados y el número de frutos de las ramas terminales, arrojando valores de  $R^2$  entre 0,79 a 0,94, de acuerdo con las diferentes situaciones.

Reeb & Tassile, (2007), determinaron que la carga de frutos por hectárea esta expresada en unidades por planta o por hectárea, y es un parámetro que debe estimarse cada año pues su valor depende de las condiciones que caracterizan a la temporada. Esta debe calcularse para cada especie y variedad por separado, pero aun así cada combinación de estos factores debe estratificarse por sistema de conducción y edad del monte debido a la heterogeneidad de la variable. Para determinar la carga en manzanas y peras, utilizan el conteo de frutos en dos árboles por unidad muestral

(cuadro). El número de unidades muestrales dentro de cada estrato se determina en función de su variabilidad interna. Esto permite conocer la carga media por árbol, y complementariamente con los datos de distanciamiento entre árboles y entre filas, realizar la expansión a número de frutos por hectárea.

Rolim *et al.* (2008), evaluaron métodos de predicción del número de frutos de naranjo Valencia sobre dos portainjertos, a partir de diferentes estructuras reproductivas (botones florales y flores) en distintos sectores de la copa. Ajustaron funciones exponenciales para expresar la relación entre el número de estructuras reproductivas y el número de frutos, las que arrojaron  $R^2$  entre 0,62 y 0,85 y errores absolutos de predicción del 2 al 19 %.

La medición de la carga del cultivo a partir del recuento del número de frutos en un determinado volumen de la copa fue utilizada por Falivene & Hardy (2008) en el Estado de Nueva Gales del Sur, para medir la carga en cítricos mediante un marco de muestreo de 0,125 m<sup>3</sup> (0,5 m x 0,5 m x 0,5 m), que se coloca en la copa del árbol a una altura de 1,0-1,5 m del suelo y aconsejan repetir el procedimiento de muestreo al menos 40 veces en un bloque de 1 hectárea.

Dentro de las técnicas de recuento manual de frutos, aquellas que combinan el recuento de una sección de copa y lo extrapolan al volumen total de cada árbol, requieren de la estimación del volumen de la copa de los árboles y no existe un consenso en las ecuaciones que deben ser utilizadas para cada especie y variedad en cuestión. Avilán & Mazzi (1973) consideraron como copa útil del árbol la contenida en un cono truncado, al determinar que la producción que está ubicada en el ápice de la planta es despreciable y que cerca del 98% de ésta se ubica en la parte media. Turrel (1946), considera la forma de la copa de los árboles similar a una esfera, y la expresión que desarrolló fue utilizada para estimar el volumen de copa de naranjo Valencia sobre siete patrones en Venezuela (Monteverde, 1996), y de mandarino Arrayana en seis patrones en Colombia (Orduz-Rodríguez *et al.* 2006). Morín (1980), considera que la forma de la copa de los árboles cítricos es redondeada, pero otros autores utilizan una semiesfera alargada, como Zekri (2000) quien usó la fórmula de la mitad de una esfera alargada dada por Castle & Phillips (1980), para el cálculo del volumen de copa en la evaluación de once portainjertos para mandarino Nova, y Georgiou (2000) para la evaluación de

doce portainjertos para mandarina Clementina en Chipre. Otros autores como Bassal (2009), utilizan la ecuación dada por Wutscher & Hill (1995), con el fin de evaluar el crecimiento, cantidad y calidad de frutos de mandarina Marisol sobre cuatro portainjertos en Egipto. Turrell (1961) y Jahn (1979) utilizaron tablas para modelar la superficie de las copas de los cítricos, suponiendo la forma de un esferoide prolato, en donde la altura del árbol es mayor que su radio (considerando el radio del árbol igual en todas direcciones).

#### *5.1.1.3. Otras mediciones*

Best *et al.* (2008), analizaron ecuaciones de regresión lineal, con el fin de estimar la carga en plantaciones de manzano Royal Gala, a partir del índice foliar radiométrico, área transversal del tronco, área y volumen de canopia. El mejor ajuste ( $R^2= 0,51$ ) se obtuvo con la ecuación que permite estimar la carga a partir del índice de área foliar radiométrico. Sin embargo, se sugieren seguir las evaluaciones con un aumento del tamaño de muestras para lograr mejoras en la precisión de las estimaciones.

Velazquez-Martí & Cazco-Logroño (2016), desarrollaron un método para ciruelos (*Prunus domestica* L.) en Imbabura, Ecuador, utilizando una dendrometría adaptada. Definieron factores de forma y modelos de regresión para el cálculo del volumen de las ramas y caracterizaron la distribución de la biomasa leñosa en la copa del árbol en cada estrato. Obtuvieron el volumen total de los árboles mediante modelos de regresión a partir del diámetro y la altura de copa con  $R^2$  de 0,81.

Koirala *et al.* (2021) describieron una clasificación basada en visión artificial de imágenes de dosel para tres densidades de carga de cultivo, utilizando un modelo y una puntuación de silueta como primer paso para una estimación basada en visión artificial de la carga de frutos por árbol. Falivene & Hardy (2008), recomendaron un recuento de 20 fotogramas por hectárea, y señalaron que, para una representación más precisa y mejor, se podrían muestrear más unidades.

### 5.1.2. Momento de realizar las mediciones

El momento en el que se deben realizar las mediciones para la estimación de carga depende del objetivo de la estimación. Cuando se desea determinar la carga como una componente del pronóstico de producción, es de esperar que cuanto más cercano al momento de la cosecha se realicen los recuentos de frutos en los árboles seleccionados, mayor será la precisión del pronóstico obtenido, por lo que es necesario compatibilizar la necesaria anticipación para que el pronóstico tenga utilidad con la precisión requerida<sup>16</sup>.

Por otra parte, las mandarinas suelen tener floraciones muy intensas y alto cuajado de frutos, lo que va en desmedro del tamaño de éstos, por lo que es común la práctica de raleo o aclareo de frutos. Cuando el objetivo de la estimación de carga es obtener una estimación de la cantidad de frutos por árbol para decidir la intensidad del raleo, debe hacerse al momento en que se va a realizar esta práctica (ACGF, 2000).

Si se va a recurrir a un aclareo químico, suele realizarse cuando los frutos tienen un diámetro de unos 10-15 mm, sin embargo, el momento puede variar en función de la región, la variedad y las condiciones estacionales. El seguimiento y registro de la caída de los frutos ayudará a identificar con mayor precisión el momento correcto, una vez que la tasa de caída de frutos disminuya significativamente, se podrá medir la carga del cultivo. Si se opta por el aclareo manual, el mejor momento para medir la carga es el comienzo de la fase II de crecimiento del fruto, cuando el tamaño de los frutos es de unos 30-40 mm (ACGF, 2000).

### 5.1.3. Estudio de la variabilidad de la carga

La productividad de los cultivos de frutales de clima subtropicales, como es el caso de los cítricos, depende de la combinación de características del árbol frutal, condiciones del medio ambiente y de los aspectos agronómicos relativos a la implantación y manejo

---

<sup>16</sup> <https://www.agric.wa.gov.au/citrus/estimating-your-citrus-crop-load>

de los montes frutales. Se puede destacar la incidencia del cultivar, portainjerto y edad entre los factores relacionados al individuo, así como del suelo y clima entre aquellos correspondientes al medio ambiente. Por tratarse de cultivos perennes, el potencial productivo está influenciado por el diseño de plantación y las tareas culturales ejecutadas durante cada ciclo anual de crecimiento, tales como poda, raleo de frutos, fertilización y riego (Agustí, 2020). La combinación de estos factores determina que el nivel de producción de frutos de un monte frutal varíe dentro y entre explotaciones de una misma zona geográfica.

Reeb *et al.* (2003) estudiaron la variabilidad de la producción en manzana Red Delicious en los montes frutales del Alto Valle de Río Negro, con el objetivo de conocer la naturaleza de la variabilidad de los rendimientos en los montes frutales aplicando técnicas de modelos multifactoriales mixtos. La principal fuente de variación se debió al efecto establecimiento, seguida en importancia por año, edad y sistema de conducción. Se encontró un ajuste a la distribución Gamma con diferencia no significativa para los parámetros de forma y escala entre los factores analizados. La magnitud de la variabilidad observada entre establecimientos de un mismo estrato motiva el análisis de otros factores (tecnológicos, sociales, etc.), no considerados en este estudio, que permitan caracterizarlos y explicar su influencia en la productividad del cultivo.

Avanza *et al.* (2010), propusieron un modelo lineal mixto para el análisis de la carga a cosecha, en el cual los efectos correspondientes a árbol y fruto se consideraron aleatorios debido al proceso de selección y los efectos de huerto y temporada, fijos. En cada combinación huerto-temporada, los árboles seleccionados fueron diferentes, por lo que el factor árbol se anidó dentro de la interacción en el modelo. Las estimaciones de varianza de árboles y frutos por huerto y temporada no presentaron diferencias entre huertos, y si mostraron algunas diferencias entre temporadas. Consideran que, tanto para el factor árbol como para fruto, las variabilidades del modelo incluyen estas variaciones particulares, y proponen utilizarlas como estimaciones de varianza en la determinación de tamaños de muestra para determinar el número óptimo de árboles y de frutos para realizar mediciones de carga.

Bella & Fichet (2014) evaluaron la carga frutal en olivos durante dos temporadas, en arboles de 13 años. Estudiaron la floración, el área de la sección transversal del tronco,

el porcentaje de flores y cuajado de las mismas, la producción en kg y el número de frutos por árbol. Emplearon Modelos Lineales Generales y Mixtos para detectar si existían diferencias significativas entre la carga frutal y las variables evaluadas en ambas temporadas. Indican que la carga frutal se ve afectada por el desarrollo fenológico y un alto porcentaje de flores cuajadas no compensa la producción de la temporada.

Ojer *et al.* (2009), realizaron un estudio en duraznero para determinar la relación entre la carga frutal y la productividad. En el ensayo realizado en un monte comercial, los árboles fueron raleados 20 días después del inicio de endurecimiento del carozo. A momento de la cosecha se evaluó el tamaño de los frutos, la producción de los árboles y el valor de la producción. La carga frutal se vinculó, con un alto grado de correlación, con todas las variables analizadas y se observó una respuesta a la carga de frutos muy similar en las tres variedades en estudio. A medida que se incrementó la carga frutal, aumentó la producción total, aunque disminuyó el peso medio de los frutos.

Independientemente de las herramientas utilizadas para la estimación de carga, todas las técnicas citadas son específicas para cada combinación de especie-variedad y condiciones en las que han sido desarrolladas, debido a que los factores climáticos y de manejo del cultivo son los que determinan, en gran medida, la variabilidad de producción entre sitios y ciclos productivos, por lo tanto, los errores de predicción reportados en algunos casos se dan para cada situación en particular.

En el presente Capítulo se evaluó y comparó la precisión obtenida con diferentes métodos de estimación de carga de frutos a partir de mediciones parcelarias, en tangor Murcott y naranja dulce Valencia late y Salustiana a fin de ser utilizados en el modelo de predicción de producción.

## 5.2. MATERIALES Y MÉTODOS

Se trabajó sobre la base de características de los huertos, información morfológica de los árboles y número de frutos por árbol.

### 5.2.1. Datos Experimentales

La información de campo fue obtenida en algunos de los huertos descritos en el Capítulo II, los que se presentan en la Tabla 5.1.



Tabla 5.1. Huertos seleccionados para la estimación

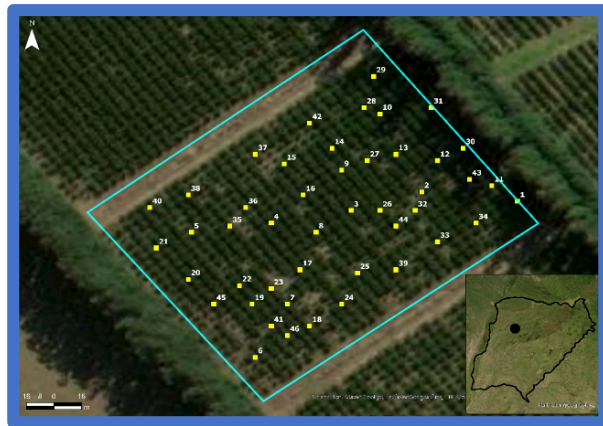
Temporada	Huerto	Especie y variedad	Fecha plena flor	Momento estimación	n
2010/11	1.1	Tangor Murcott	26-09-10	1-2	30
2014/15	1.2	Tangor Murcott	30-09-14	1	18
2014/15	1.3	Tangor Murcott	01-10-14	1	18
2015/16	1.4	Tangor Murcott	30-09-15	1-2	48
2015/16	1.5	Tangor Murcott	30-09-15	1-2	42
2010/11			17-09-10	1-2	37
2011/12	2.1	Tangor Murcott	18-08-13	1-2	37
2013/14			30-08-12	1-2	37
2010/11			23-09-10	1-2	32
2013/14	3.1	Tangor Murcott	18-08-13	1-2	28
2014/15			13-10-14	1-2	28
2015/16	3.2	Tangor Murcott	26-09-15	1-2	46
2011/12	4.1	Tangor Murcott	23-09-11	1	26
2007/08			08-10-07	1-2	26
2008/09	5.1	Tangor Murcott	01-10-08	1-2	25
2009/10			02-10-08	1-2	24
2007/08			04-10-07	1-2	19
2008/09	6.1	Tangor Murcott	28-09-08	1-2	17
2009/10			30-09-09	1-2	17
2007/08			02-10-07	1-2	42
2008/09	7.1	Tangor Murcott	30-09-08	1-2	36
2009/10			30-09-09	1-2	41
2007/08			05-10-07	1-2	31
2008/09	8.1	Tangor Murcott	01-10-08	1-2	31
2008/09			30-09-09	1-2	31
2005/06	1.1	Naranja dulce Valencia late	03-10-05	1	17
2005/06	2.1	Naranja dulce Valencia late	08-10-05	1	25
2006/07			25-09-06	1-2	31
2007/08	3.1	Naranja dulce	10-10-07	1-2	41
2008/09		Valencia late	08-10-08	1-2	47
2009/10			12-10-09	1-2	51
2006/07	4.1	Naranja dulce Valencia late	15-10-06	1-2	23
2007/08			09-10-07	1-2	31
2008/09	5.1	Naranja dulce	04-10-08	1-2	25
2009/10		Valencia late	05-10-08	1-2	25
2006/07			23-10-06	1-2	16
2007/08	6.1	Naranja dulce	09-10-07	1-2	16
2008/09		Valencia late	03-10-08	1-2	17
2009/10			05-10-08	1-2	17

Tabla 5.1. Continuación

Temporada	Huerto	Especie y variedad	Fecha plena flor	Momento estimación	n
2006/07			06-09-06	1-2	42
2007/08	6.2	Naranja dulce	08-10-07	1-2	44
2008/09		Valencia late	03-10-08	1-2	28
2009/10	6.2	Naranja dulce	05-10-08	1-2	30
		Valencia late			
2006/07			08-09-06	1-2	45
2007/08	6.3	Naranja dulce	08-10-07	1-2	45
2008/09		Valencia late	29-09-08	1-2	31
2009/10			02-10-09	1-2	32
2006/07			13-09-06	1-2	51
2007/08	7.1	Naranja dulce	06-10-07	1-2	53
2008/09		Valencia late	26-09-08	1-2	37
2009/10			03-10-09	1-2	37
2013/14	8.1	Naranja dulce Salustiana	15-09-13	1-2*	38
2013/14	9.1	Naranja dulce Salustiana	15-09-13	1-2*	50

\*Para Salustiana, la estimación 1 corresponde al mes de febrero y 2 al mes de abril, por tratarse de una variedad de ciclo intermedio y la cosecha se realizó en el mes de junio

En cada huerto se seleccionó entre el 2% y 3% de los árboles mediante un muestreo al azar sistemático con arranque aleatorio. A modo de ejemplo, en la Figura 5.1. se presenta la distribución de los árboles seleccionados en 3 de los huertos.



(a)



(b)



(c)

Figura 5.1. Plantas seleccionadas para el recuento de frutos.  
 (a) Huerto 1.5, (b) Huerto 1.6, (c) Huerto: 3.1

En los árboles seleccionados se registró la información que se consigna a continuación.

### 5.2.1.1. Dimensiones de los árboles

Para cada huerto, en los árboles seleccionados se midieron, con cinta métrica, las siguientes variables:

**Altura de árbol (h):** medida desde el piso hasta la altura máxima de copa.

**Diámetro mayor** ( $d_M$ ): medida entre los extremos de las ramas más largas en la dirección entre de las líneas.

**Diámetro menor** ( $d_m$ ): medida entre los extremos de las ramas más largas en la dirección dentro de las líneas.

Y se determinó,

**Diámetro promedio** ( $d$ ): calculado como el promedio entre  $d_M$  y  $d_m$ .

**Radio de copa** ( $r$ ): calculado como la mitad de  $d$ .

#### 5.2.1.2. Estimaciones de carga

En los meses de abril y junio se determinó el número de frutos por árbol mediante:

**Recuento total de frutos en cada árbol** (RT), realizado de manera simultánea por dos operarios que utilizaban un contador manual y repetían la operación hasta lograr una diferencia entre ambos menor o igual al 10% (Figura 5.2). Para corregir la carga estimada se consideró un factor denominado 'de fruto oculto', mediante un valor obtenido de manera experimental, para las distintas situaciones previstas en las fechas usuales de relevamiento a campo. Este valor se expresa como el cociente entre la carga estimada ( $x$ ) y la carga real a cosecha ( $y$ ), lo que permite determinar el coeficiente de fruto oculto ( $CFO = \frac{x}{y}$ ).



Figura 5.2. Determinación de la carga mediante el recuento total

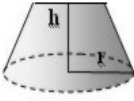
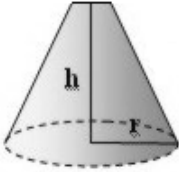
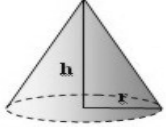
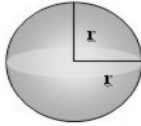
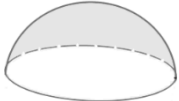
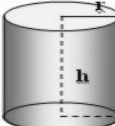
**Recuento en secciones predefinidas de copa**, se efectuó mediante el uso de un marco de muestreo de  $0,125 \text{ m}^3$  colocado en los 4 puntos cardinales (NumN - NumS - NumE - NumO) (Figura 5.3) y luego se determinó el promedio entre las cuatro mediciones para cada árbol (NumP). Posteriormente se extrapolaró el número de frutos promedio por marco de muestreo (NumP) al total de copa, para lo que se calcularon diferentes volúmenes de copa utilizando las expresiones que se presentan en la Tabla 5.3, obteniéndose las cargas estimadas para diferentes volúmenes de copa (VolM1, VolM2, VolM3, VolM4, VolM5 y VolM6).



Figura 5.3. Uso del marco de muestreo

Para calcular los volúmenes de VolM1, VolM2, VolM4 y VolM5 se aplican fórmulas de volumen de copa que han sido utilizadas por diferentes autores para árboles cítricos, y para VolM3 y VolM6 se aplican ecuaciones de volumen de diferentes cuerpos geométricos que podrían ajustarse a la forma de la copa que presentan los árboles en estudio.

Tabla 5.2. Fórmulas para el cálculo de volumen de copa correspondientes a cada método del marco contador (VolM1, VolM2, VolM3, VolM4, VolM5, VolM6) y sus respectivas referencias

Método	Fórmula	Expresión unificada	Esquema
VolM1 (Roose <i>et al.</i> ,1989)	$V = \frac{4}{6} \cdot h \cdot r^2$	$V = 0,1667 \cdot d^2 \cdot h$	
VolM2 (Wutscher & Hill, 1995)	$V = \left( \frac{d^2 \cdot h}{4} \right)$	$V = 0,2500 \cdot d^2 \cdot h$	
VolM3 (Cono)	$V = \pi \cdot r^2 \cdot \frac{h}{3}$	$V = 0,2618 \cdot d^2 \cdot h$	
VolM4 (Esfera o Turrel, 1946)	$V = \frac{4}{3} \cdot \pi \cdot r^3$	$V = 0,5236 \cdot d^2 \cdot h$	
VolM5 (Zekri, 2000)	$V = \frac{\pi}{6} \cdot h \cdot d_m \cdot d_M$	$V = 0,5230 \cdot d^2 \cdot h$	
VolM6 (Cilindro)	$V = \pi \cdot r^2 \cdot h$	$V = 0,7854 \cdot d^2 \cdot h$	

V= volumen de copa (m<sup>3</sup>), h= altura de copa (m), r= radio de copa (m), d= diámetro de copa (m), d<sub>M</sub>= diámetro mayor de copa (m) y d<sub>m</sub>=diámetro menor de copa (m).

### 5.2.1.3. Valor de comparación

En los huertos estudiados, en el momento de la cosecha (Agosto - Octubre), en cada árbol seleccionado se realizó el conteo de la carga real por árbol, luego de la remoción de toda la fruta, que se empleó como valor de comparación (Carga a cosecha).

## 5.2.2. Análisis de datos

### 5.2.2.1 Error de estimación

Para evaluar la precisión de cada método de estimación, se calcularon los errores porcentuales absolutos (% Error) por especie, entre la carga estimada y la real registrada en cosecha, mediante la expresión:

$$\% \text{ Error} = \frac{\sum_i |\hat{\mu}_i - y_i|}{\sum_i y_i} \cdot 100$$

Donde  $\hat{\mu}_i$  es la predicción de carga para el *i*-ésimo árbol e  $y_i$  es el valor de comparación, esto es la carga real del *i*-ésimo árbol.

Con la finalidad de determinar si existen diferencias entre los momentos de estimación (abril y junio), se utilizó la prueba de t-Student para muestras pareadas.

A los efectos de evaluar la relación entre la carga estimada por cada método y la cantidad de frutos real (cosecha), se calcularon coeficientes de correlación de Pearson (*r*).

### 5.2.2.2. Variabilidad de la carga estimada y la cosecha real

Para estimar la variabilidad de la carga a cosecha en tangor Murcott y naranjo dulce, de acuerdo con el método seleccionado y la carga real a cosecha se consideraron los efectos aleatorios: *temporada*, *huerto* y *árbol*, el modelo lineal asociado fue el siguiente:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \varepsilon_{ijk}$$

Donde:  $y_{ijk}$  es la variable respuesta correspondiente a la temporada *i*, al huerto *j* y al árbol *k*;  $\mu$  es la media general;  $\alpha_i$  es el efecto de la *i*-ésima temporada,  $i = 1, \dots, n$ ;  $\beta_j$  es el efecto del *i*-ésimo huerto,  $j = 1, \dots, n$ ;  $\tau_k$  es el efecto del *i*-ésimo huerto,  $k = 1, \dots, n$ ; y  $\varepsilon_{ij}$  es el error aleatorio de la *i*-ésima temporada en el *j*-ésimo huerto, en el *k*-ésimo árbol.

Los modelos no lineales mixtos se ajustaron con el método de máxima verosimilitud restringida, utilizando la librería de *lme4* (Linear Mixed-Effects Models using 'Eigen' and S4) del paquete R (R Development Core Team<sup>17</sup>).

## 5.3. RESULTADOS Y DISCUSIÓN

### 5.3.1. Porcentajes de error

En la Tabla 5.3 se presentan los coeficientes de correlación y los p-valores asociados entre la carga estimada por los diferentes métodos de recuento y el valor de comparación (carga real a cosecha) para las diferentes especies y momentos de estimación.

Tabla 5.3 Coeficientes de correlación de Pearson (r) entre los diferentes métodos de estimación de carga (medición en secciones predefinidas de copa: VolM1, VolM2, VolM3, VolM4, VolM5, VolM6; recuento total: RT) y la carga real por especie y momento de estimación (ME1: abril y ME2: junio)

Métodos	Tangor Murcott				Naranja dulce			
	ME1	p-valor	ME2	p-valor	ME1	p-valor	ME2	p-valor
<b>VolM1</b>	0,22	<0,0001	0,18	<0,0001	0,58	<0,0001	0,68	<0,0001
<b>VolM2</b>	0,28	<0,0001	0,27	<0,0001	0,63	<0,0001	0,69	<0,0001
<b>VolM3</b>	0,35	<0,0001	0,32	<0,0001	0,68	<0,0001	0,69	<0,0001
<b>VolM4</b>	0,08	0,0523	0,03	0,5017	0,20	<0,0001	0,62	<0,0001
<b>VolM5</b>	0,07	0,0685	0,02	0,5806	0,33	<0,0001	0,61	<0,0001
<b>VolM6</b>	0,12	0,0041	0,07	0,1102	0,45	<0,0001	0,65	<0,0001
<b>RT</b>	0,84	<0,0001	0,89	<0,0001	0,92	<0,0001	0,94	<0,0001

En naranja dulce los coeficientes de correlación de Pearson fueron significativos para todos los métodos de recuento, e indican que la carga estimada por los diferentes métodos se encuentra asociada de manera positiva con la carga real. En el caso de tangor Murcott, la carga real se asoció de manera significativa y positiva, solamente con

<sup>17</sup> Disponible en: <https://www.r-project.org>



la carga estimada por los métodos VolM1, VolM2, VolM3 y VolM6 (ME1) y RT. En el caso de los métodos VolM4, VolM5 y VolM6 (ME2) la relación no fue significativa.

La estimación VolM4, corresponde al volumen calculado mediante la fórmula de Turrel (1969). Esta forma de los árboles ha sido reportada por Chapararro-Zambrano *et al.* (2017) para mandarina Arrayana injertada sobre diferentes portainjertos. Y la estimación M5 a un volumen de copa asimilado a una semiesfera, según lo informado por Zekri (2000), en el cálculo del volumen de copa de mandarina Nova, y por Georgiou (2000) para la evaluación en mandarina Clementina.

Las asociaciones determinadas en esta tesis indicarían que estas formas de copa no serían apropiadas para tangor Murcott. Coincidiendo con Anderson *et al.* (2021) quienes indicaron que atributos de los árboles de mango como el volumen de la copa y la circunferencia del tronco estaban poco correlacionados con la carga de fruta ( $R^2 = 0,21$  y  $0,17$ , respectivamente).

El método de recuento (RT), presenta una fuerte correlación entre la carga estimada y la carga real, con coeficientes  $r = 0,83$  (ME1), y  $r = 0,89$  (ME2) para tangor Murcott y  $r = 0,92$  (ME1) y  $r = 0,94$  (ME2) para naranjos dulces, estos valores fueron similares a los hallados por Avanza (2011) en naranjo dulce Valencia late ( $r = 0,87$ ).

En la Tabla 5.4 se detalla el comportamiento de los diferentes métodos para estimar la carga en tangor Murcott y naranjo dulce.

Tabla 5.4. Porcentajes de error entre la carga estimada y la real para los diferentes métodos de estimación de carga (medición en secciones predefinidas de copa: VolM1, VolM2, VolM3, VolM4, VolM5, VolM6; recuento total: RT) por especie y momento de estimación (1: Abril y 2: Junio)

Especie	Momento estimación	VolM1	VolM2	VolM3	VolM4	VolM5	VolM6	RT
Tangor Murcott	1	78,23	58,88	63,10	37,27	37,48	42,50	7,29
	2	78,38	60,73	63,41	39,71	39,66	52,14	6,56
	General	78,31	59,81	63,26	38,49	38,57	47,32	6,93
Naranjo dulce	1	88,79	72,89	76,24	66,03	71,38	53,48	31,89
	2	90,15	76,89	77,94	75,74	76,06	60,27	20,97
	General	89,47	74,89	77,09	70,89	73,72	56,88	26,43

Los errores de estimación para los métodos de marco contador variaron entre 47,32 y 78,31 % en tangor Murcott y entre 56,88 y 89,47 % en naranjo dulce, presentando comportamientos dispares. Todos los porcentajes de error medio obtenidos con los

diferentes métodos fueron superiores a encontrados por Rolim *et al.* (2008), en la estimación de carga naranjo dulce.

Cuando se estimó la carga con el método de recuento total se obtuvieron porcentajes de error de 6,93% en tangor Murcott y de 26,43%, en naranjo dulce. En esta última especie en particular, los errores fueron mayores a los hallados por Avanza (2011) para naranjo Valencia late, quien obtuvo valores entre 3,37 y 10,21% según huertos y temporadas. Este método presentó mejor comportamiento para estimar la carga que los de recuento en secciones predefinidas de copa, coincidiendo con los resultados de Gil (2003).

Los métodos de recuento en secciones predefinidas de copa estudiados no tuvieron la capacidad de estimar la carga real con una adecuada precisión y el método de recuento total, fue el más preciso para ambas especies. Estos resultados concuerdan con los obtenidos en el programa de pronóstico de cosecha para frutales de pepita en Río Negro y Neuquén, Argentina, donde se utiliza el recuento total de frutos corregido por un coeficiente de fruto oculto que actúa como divisor y que varía entre 0,79 y 0,93 de acuerdo con la especie.

Dado que el método de recuento total presentó los porcentajes de error de estimación más bajos, se decidió seleccionar este método para la estimación de la carga y a los efectos de definir el mejor momento para realizar la estimación se procedió a realizar una prueba T ( $\alpha = 0,05$ ) entre los momentos de estimación por especie (Tabla 5.5).

Tabla 5.5. Prueba T pareada de los recuentos de frutos por árbol entre momentos de estimación (ME1 y ME2),  $\bar{X}_{dif}$ : promedio de las diferencias,  $Sx_{dif}$ : desvío estándar de la diferencia, p-valor, obtenidos mediante el método recuento total

Especie	ME1	ME2	$\bar{X}_{dif}$	$Sx_{dif}$	p-valor
tangor Murcott	368,36	407,21	-38,86	168,33	<0,0001
Naranjo dulce	324,44	379,99	-55,56	99,92	<0,0001

Se detectaron diferencias significativas (p-valor < 0,05) entre la carga determinada en abril y junio para tangor Murcott y naranjo dulce. En ambas especies el número estimado de frutos por árbol fue mayor en los recuentos de junio que en los de abril,

con un promedio de diferencia de 38,86 frutos y un desvío estándar de 168,83 frutos en tangor Murcott y de 55,56 frutos y un desvío estándar 99,92 frutos.

En función de que los menores porcentajes de error se obtuvieron con el método del recuento total y cuando las estimaciones se realizaron en el segundo momento, esta combinación método-momento fue seleccionada como método de estimación de la carga. A partir de ahí se continuó trabajando solamente con las estimaciones realizadas en el mes de abril para naranjo Salustiana y de junio para tangor Murcott y naranjo Valencia, con las estimación de carga mediante recuento total de frutos.

No obstante haber seleccionado el método de recuento y el momento para la estimación de la carga que presentaban los menores errores, éstos seguían siendo superiores a los hallados en la bibliografía, por lo que se procedió a corregir la carga estimada mediante el coeficiente de fruto oculto (CFO), calculado a partir de los datos experimentales y cuyos valores se presentan en la Tabla 5.6.

Tabla 5.6. Carga estimada, real y corregida por coeficiente de fruto oculto para tangor Murcott.

Temporada	Huerto	Carga estimada	Carga real	Carga estimada corregida
2010-11	1.1	547,69	622,38	639,78
2015-16	1.2	394,79	448,63	483,57
2015-16	1.5	315,92	359,00	383,67
2013-14	2.1	241,55	274,49	276,29
2010-11	3.1	747,25	849,15	876,00
2013-14		382,50	434,66	520,25
2014-15		309,33	351,51	541,35
2015-16	3.2	282,14	320,61	299,16

Para tangor Murcott se obtuvo un valor promedio de CFO = 0,88, superior al determinado por Avanza (2011) de 0,65 para naranjo Valencia late.

Al corregir la carga estimada mediante el CFO, la carga corregida presentó valores inferiores a los de la carga real, excepto el huerto 3.1 en la temporada 2014-15 que fue superior. Esta corrección es necesaria porque que el método de recuento total se realiza de manera visual entre dos operarios y se produce un error sistemático debido a que ciertos frutos no se visualizan.

El porcentaje de error promedio del método de recuento total fue de 9,85%, que es inferior a los errores entre 11 y 18% reportados por Marini (2011) para manzana, los

porcentajes de error variaron entre el 11% y el 19% y similar al 10% hallado por Anderson *et. al.* (2019) en huertos de mango.

### 5.3.2. Variabilidad de la carga estimada y real a cosecha

A continuación, se presenta el estudio de la variabilidad de la carga estimada para el momento 2 (Junio), considerando los efectos aleatorios de los factores *temporada*, *huerto* y *árbol*, por especie.

En la Tabla 5.7 se detallan los componentes de varianzas correspondientes a tangor Murcott y en la Tabla 5.8 los correspondientes a naranja dulce.

Tabla 5.7. Componentes de varianza para las estimaciones de carga en tangor Murcott

Componente	Varianza estimada	Variabilidad relativa al total (%)
Temporada	$\sigma^2 = 57,46^2 = 3.301,65$	4,95
Huerto	$\sigma^2 = 115,56^2 = 133.707,92$	19,96
Árbol	$\sigma^2 = 202,67^2 = 41.075,12$	61,62
Residuo	$\sigma^2 = 94,69^2 = 8.966,19$	13,45

Según se desprende de la Tabla 5.7, al factor árbol (que representa la variabilidad del número de frutos entre los diferentes árboles de un huerto comercial) le corresponde una componente de varianza estimada de 61,62%, superior a la variabilidad entre huertos (19,96%) y temporadas (4,95%). Estos resultados fueron similares a los reportados por Bóveda *et al.* (2014), en un estudio previo, en el que se evaluó la variabilidad del número de frutos, mediante el marco contador, dentro de cada árbol y entre árboles dentro del huerto, encontrando que la mayor variabilidad se debía al efecto árbol y representaba el 59,09%.

Tabla 5.8. Componentes de varianza para las estimaciones de carga en naranjo dulce

Componente	Varianza estimada	Variabilidad relativa al total (%)
Temporada	$\sigma^2 = 117,87^2 = 13.893,33$	21,16
Huerto	$\sigma^2 = 154,50^2 = 23.870,25$	36,36
Árbol	$\sigma^2 = 153,77^2 = 23.645,21$	36,02
Residuo	$\sigma^2 = 65,11^2 = 4.239,31$	6,45

En el caso de naranjo dulce, por el contrario, el factor huerto (que representa la variabilidad del número de frutos de los diferentes árboles dentro del huerto) arrojó una componente de varianza estimada de 36,36% similar a la variabilidad de los frutos dentro de los árboles (36,02%).

Estos resultados son similares a los reportados por Chapman *et al.* (1992), quienes detectaron que puede haber marcadas diferencias en la cantidad de fruta producida por un solo árbol durante diferentes períodos de fructificación. Por lo que se necesita una evaluación de dicha variación a lo largo del tiempo o entre árboles del mismo tamaño en diferentes áreas.

Al analizar la carga real a cosecha, las estimaciones de componentes de varianzas para ambas especies se detallan en las Tablas 5.9 y 5.10.

Tabla 5.9. Componentes de varianza de la carga real a cosecha para tangor Murcott

Componente	Varianza estimada	Variabilidad relativa al total (%)
Temporada	$\sigma^2 = 85,37^2 = 7.288,03$	3,96
Huerto	$\sigma^2 = 216,54^2 = 46.889,57$	25,50
Arbol	$\sigma^2 = 355,76 = 126.543,83$	68,83
Residuo	$\sigma^2 = 55,83^2 = 3.116,98$	1,69

Tabla 5.10. Componentes de varianza estimada de la carga real a cosecha para naranjo dulce

Componente	Varianza estimada	Variabilidad relativa al total (%)
Temporada	$\sigma^2 = 141,00^2 = 19.881,00$	12,41
Huerto	$\sigma^2 = 256,46^2 = 65.771,73$	41,07
Arbol	$\sigma^2 = 271,62^2 = 73.777,42$	46,07
Residuo	$\sigma^2 = 26,04^2 = 678,08$	0,42

Para ambas especies el factor árbol, que representa la variabilidad del número de frutos entre los diferentes árboles dentro de un huerto, obtuvo la mayor componente de varianza estimada (68,83% en tangor Murcott y 46,07% en naranjo dulce). Estos valores resultaron superiores a las variabilidades debidas a los factores huerto y temporada (25,50% y 3,96% respectivamente en tangor Murcott y de 41,07% y 12,46% respectivamente en naranjo dulce).

La alta variabilidad de la carga observada implica que, para una misma temporada, los huertos produzcan una cantidad de frutos por árbol muy distinta. Estos resultados difieren de los hallados por Reeb *et. al.* (2003) en montes frutales de manzano, quienes encontraron que la principal fuente de variación se debió al efecto establecimiento. No obstante, estos resultados resultan coherentes si se comparan los estudios realizados por Bramardi (1995) en perales y manzanos y Avanza (2011) en naranjo dulce, quienes determinaron los tamaños muestrales para estimación de carga en función de la variabilidad, que arrojaron valores de 4 árboles en peras y manzanos y entre 18 a 21 árboles en naranjo, de donde se puede desprender que la variabilidad de la carga en cítricos es superior a la de frutales de pepita como manzanos y perales.

## 5.4. RESUMEN DE RESULTADOS

Los métodos de determinación de carga que utilizan el marco contador presentaron asociaciones significativas con la carga real, pero comportamientos dispares respecto a los errores de estimación. Estos resultados indican baja precisión en las estimaciones mediante el método del marco contador, lo que no concuerda con los obtenidos por otros autores, por lo que se considera adecuado continuar con la investigación respecto a un mayor número de marcos por árbol y disposición de éstos en diferentes sectores de la copa.

El método seleccionado para ser utilizado para la estimación de carga en tangor Murcott y naranjo dulce, es el recuento total de frutos corregido por un coeficiente de fruto oculto de 0,88, que presenta la mayor asociación entre la carga estimada y la real a cosecha y menores errores de estimación con dos meses de anticipación a la cosecha, que en promedio se estableció en 325 después de plena floración para la región del NEA.

Al analizar las componentes de varianza de la carga estimada se encontró en tangor Murcott que la mayor variabilidad le corresponde al factor árbol (61,62%), superior a la variabilidad entre huertos (19,96%) y temporadas (4,95%). En el caso de naranjo dulce, por el contrario, el factor huerto arrojó una componente de varianza estimada de similar a la variabilidad de los frutos dentro de los árboles (36,36% y 36,02% respectivamente).

En el caso de la carga real a cosecha, para ambas especies el factor árbol obtuvo la mayor componente de varianza estimada (68,83% en tangor Murcott y 46,07% en naranjo dulce), con valores superiores a las variabilidades debidas a los factores huerto y temporada (25,50% y 3,96% respectivamente en tangor Murcott y de 41,07% y 12,46% respectivamente en naranjo dulce).

## 5.5. BIBLIOGRAFÍA

1. Agustí, M.; Mesejo, C.; Mesero, C., Martínez-Fuentes, A.; Zaragoza, S.; Primo Millo, E. 2020. El tamaño del fruto en los cítricos. Levante Agrícola. Revista Internacional de Cítricos. 451, 57-67.
2. Anderson, N.T.; Walsh, K.B.; Wulfsohn, D. 2021. Technologies for Forecasting Tree Fruit Load and Harvest Timing—From Ground, Sky and Time. *Agronomy*. 11(7):1409.
3. Australia Citrus Growers Federation (ACGF). Fruit Size Management Guide. P.2 2000 [en línea].  
<http://www.citrusaustralia.com.au/aspdev/resources/documents/FruitSizeManagementGuidePart2.pdf> [Consulta: Octubre, 2021].
4. Avanza M.M.; Bramardi, S.; Mazza, S.M. 2010. Optimal Sample size evaluate the pattern grow of Valencia late orange fruit. Revista Brasileira de Fruticulture. (32) 4.
5. Avanza, M.M. 2011. Desarrollo de una propuesta metodológica para la predicción de cosecha de naranjo dulce (*Citrus sinensis* L. Osbeck) var. Valencia late en la provincia de Corrientes, Argentina. Tesis Doctoral de la Universidad Nacional del Nordeste, Área Recursos Naturales.
6. Avilán, L.R., Mazzi, L. 1973. Estimación de cosecha en frutales. 15 pp. (Monografía).
7. Avilán L. Et Meneses L. 1979. Efecto de las propiedades físicas del suelo sobre la distribución de las raíces del mango (*Mangifera indica* L.). *Turriialba*, 29 (2) 117-122.
8. Bassal, M. 2009. Growth, yield and fruit quality of 'Marisol' clementine grown on four rootstocks in Egypt. *Scientia Horticulturae*. 119. 132-137.
9. Beya, V. M. & Fichel, T.M. 2014 Efecto de la carga frutal sobre la fenología y crecimiento de raíces en olivos (*Olea Europea* L) 'Frantoio' En la Región Metropolitana. Capítulo 1. Presentaciones orales. 3 p.
10. Best, S.; Salazar, F.; Bastías, R.; Leon, L. 2008. Crop Load Estimation Model to Optimize Yield – Quality Ratio in Apple Orchards, *Malus Domestica* Borkh, Var. Royal Gala. *Journal of Information Technology in Agriculture*, 3 (1): 11-18.



11. Bóbeda, G.R.R.; Giménez, L.; Mazza, S.M. 2014. Estimación preliminar de la variabilidad del número de frutos por árbol en tangor Murcott, en la localidad de Bella Vista, Corrientes. IV Simposio Nacional. *I Congreso Latinoamericano. Investigación y Desarrollo Tecnológico en Citrus*. 3-5 Noviembre. Salto, Uruguay.
12. Castle, W.S., Phillips, R.L., 1980. Performance of Marsh grapefruit and Valencia orange trees on 18 rootstocks in a closely spaced planting. *J. Am. Soc. Hort. Sci.* (105), 496:499
13. Chapman, C.A., Wingham, R., Hunt, K.D., Gebo, D.L., & Gardner, L. 1992. Estimators of Fruit Abundance of Tropical Trees. *24(4): 527-531*
14. Chaparro-Zambrano, H. N., Velásquez-Ramírez, H. A., & Ordúz-Rodríguez, J. O. (2017). Evaluation of Arrayarúa tangerine (*Citrus reticulata* Blanco) grafted onto different rootstocks in tropical lowlands of Colombian Orinoquia, 2005-2011 (second cycle). *Agronomía Colombiana*, 35(1), 29-34.
15. De Rezende Triboni, H., Barbosa, J. C. 2004. Estimativa do número de frutos por amostragem de parte da copa em laranjeiras. *Revista Brasileira de Fruticultura*, 26 (3): 454-458.
16. Falivene, S.; Hardy D.S. 2008., Assessing Citrus Crop Load Primefact 787. ISSN 1832-6668. 4p.
17. Georgiou, Androula. 2000. Performance of Nova mandarin on eleven rootstocks in Cyprus. *Scientia Horticulturae*. (84):115-126.
18. Gil, F. 1996. Bases para elaborar un pronóstico de cosecha en manzano (*Malus pumila* Mill) cv. Fuji. Factores que modifican el crecimiento del fruto. Tesis presentada para optar al grado de Magister en Ciencias Agropecuarias, Mención Producción Frutícola. Facultad de Ciencias Agrarias y Forestales, Universidad de Chile, 99 pp.
19. Gil, F. 2003. Curso de Capacitación NEA. Técnicas de estimación para cítricos. Instituto de Desarrollo Rural. Programa de Pronóstico de Cosecha Frutícola, 53 pp.
20. Huddleston, H.F. 1978. Sampling techniques for measuring and forecasting crop yields. *Economics, Statistics and Coop. Serv. U.S. Dept. of Agric.* 192 pp.
21. Jahn, O. L. 1979. Penetration of photosynthetically active radiational a measurement of canopy density of citrus trees. *Journal American Horticulture Scientia*. (104): 557-560.

22. Koirala, A., Walsh, K. B., Wang, Z., and McCarthy, G. (2019). Deep learning – Method overview and review of use for fruit detection and yield estimation. *Comput. Electron. Agric.* (162), 219–234.
23. Lacey, K. Department of Primary Industries and Regional Development. Estimating Your Citrus Crop. 2019. Available online: **¡Error! Referencia de hipervínculo no válida.** (Acceso: 10 de octubre 2021).
24. Marini, R. P. 2001. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked advertisement solely to indicate this fact. Estimating Mean Fruit Weight and Mean Fruit Value for Apple Trees: Comparison of Two Sampling Methods with the True Mean *J. AMER. SOC. HORT. SCI.* 126(4):503–510. 2001
25. Monteverde, E. 1996. Evaluación del Naranja ‘Valencia’ sobre siete patrones en Valles Altos de Carabobo-Yaracuy, Venezuela. *Agronomía Tropical*, 46 (4): 391-393.
26. Morín, C.H. 1980. Cultivos de cítricos. Lima. Stadium, S.A. 597 pp.
27. Ojer, M., G. Reginato y F. Vallejos. 2009. Manejo de la carga frutal y productividad en duraznos conserveros. *Revista de la Facultad de Ciencias Agrarias UNCuyo* 41(1): 65-76.
28. Olsen, J.; Goodwin, J. The methods and results of the Oregon Agricultural Statistics Service: Annual objective yield survey of Oregon hazelnut production. In *Proceedings of the VI International Congress on Hazelnut*, Davis, CA, USA, 15 July 2001; 533–538 pp.
29. Orduz-Rodríguez, J.O., Arango-Wiesner, L., Monroy, H., Fischer, G. 2006. Comportamiento de la mandarina Arrayana en seis patrones en suelos ácidos del piedemonte Llanero de Colombia. *Agronomía Colombiana*, 24 (2): 266-273.
30. Otero, A. Goñi, C., Carrau, F.; Zefferino, E.; Brennam, M. 2005. Previsión de productividad y calidad en citrus: herramientas para una gestión sustentable. *Revista INIA*, 4: 21-24.
31. Racskó, József. 2006. Crop Load, Fruit Thinning and their Effects on Fruit Quality of Apple (*Malus domestica* Borkh.). *Journal Of Agricultural Sciences.* (24).
32. Reeb, P., Bramardi, S. Álvarez, O. 2003. Estudio de la variabilidad de la producción de manzanas Red Delicious en los montes frutales del Alto Valle de Río Negro, Argentina. *Agro Sur*, 31(1): 21-26.

33. Reeb, P. & Tassile, V. 2007. Estudios Estadísticos. Pronósticos de Producción. (<https://docplayer.es/42642541-Pronosticos-de-produccion.html> )
34. Rolim, G. D. S., Ribeiro, R. V., Azevedo, F. A., Camargo, M. B. P., Machado, E. C. 2008. Previsão do número de frutos a partir da quantidade de estrutura reprodutiva em laranjeiras. *Revista Brasileira de Fruticultura*, 30 (1): 48-53.
35. Roose, M. L., Cole, D. A., Atkin, D., Kupper, D.R. 1989. Yield and tree size of four citrus cultivars on 21 rootstocks in California. *Journal of the American for Horticultural Science*, 114 (4): 678-684.
36. Tersoglio , M.L. Rivero; C. A.; Quiroga, M. 2004. Nuevo Índice de densidad de carga en Cerezas *prunus avium* cvs. 'Rainier'y 'Bing', su Relación con el tamaño, contenido de sólidos solubles y distribución de Calibres. Instituto Nacional de Tecnología Agropecuaria INTA. 100 (3):173- 186.
37. Turrel, F. M. 1946. Tables of surface and volumes of apheres and of prolate and oblates and spheroidal coefficients. Berkeley. Univ. California Press.
38. Turrel, M.F.; Garber, M. J.; Cooper, W.C; And Young, R. H. 1969. Growth Equations and Curves for Citrus Trees. Journal of agricultural ciencia published by California agriculture experiment station. Hilgardia. 30 (16): (429-445). W. C.
39. Velázquez-Martí, B. and Cazco-Logroño, c. (2017). structure analysis and biomass models for plum tree (*prunus domestica* l.) in Ecuador. *Experimental Agriculture*. (54):1-9.
40. Verreyne, S. 2009. Fruit size and crop load prediction for citrus. Citrus Research International, Department of Horticultural Science, University of Stellenbosch ([sv@cri.co.za](mailto:sv@cri.co.za)).
41. Wutscher, H. K., Hill, L. L. 1995. Performance of 'Hamlin' orange on 16 rootstocks in east-central Florida. *HortScience*, 30 (1): 41-43
42. Zekri, M. 2000. Citrus rootstocks affect scion nutrition, fruit quality, growth, yield and economical return. *Fruits*, (55): 231-239.

## CAPITULO VI

### ESTIMACION DE CARGA BASADA EN INFORMACIÓN PARCELARIA, CLIMÁTICA Y SATELITAL

#### 6.1. INTRODUCCIÓN

La agricultura implica altos niveles de riesgo en la producción y la toma de decisiones requiere considerar diferentes factores. Definir estrategias de gestión y programas de desarrollo demanda un conocimiento adecuado de las variables que afectan más directamente a la producción. Entender el comportamiento de dichas variables es tarea difícil debido a la complejidad de las relaciones entre ellas y a la cantidad de factores que intervienen. La producción de cítricos se convierte, además, en un reto especial debido a la importante variabilidad espacial y temporal presente en los huertos (Díaz *et al.* 2017).

En la producción de frutos en la región, las operaciones de recolección están escasamente mecanizadas y necesitan la contratación temporal de mano de obra, lo que exige una cuidadosa planificación (Torregrosa *et al.*, 2014). Además del momento de la cosecha, el reto más importante en dicha planificación es estimar el número de frutos que habrá que cosechar con la mayor antelación posible (Lado *et al.*, 2014). Las estimaciones tempranas de la producción también son útiles a nivel gubernamental para generar presupuestos, planificar estrategias ante desastres naturales y anticipar las demandas de transporte y almacenamiento (Lopresti *et al.*, 2015).

En la mayoría de los países productores de fruta se destinan grandes superficies a este tipo de cultivos con el objeto de satisfacer la demanda mundial, por lo que es necesario que la estimación del rendimiento de los huertos sea lo más precisa posible a fin de obtener un pronóstico adecuado de la producción total y el tamaño de frutos a cosechar. Esto permite que otras actividades, como por ejemplo la cosecha y la comercialización, puedan ser planificadas de manera eficaz por los agricultores.

Tradicionalmente, la estimación del rendimiento se ha llevado a cabo mediante recuento manual a campo, pero conduce a resultados de baja precisión, con altos costos, dado que requiere la observación de expertos y mayores tiempos para la estimación, por lo que es necesario un sistema inteligente de estimación del rendimiento que supere estos problemas.

En los huertos de cítricos, la producción se define principalmente por la cantidad y el tamaño de los frutos, y puede verse afectada por factores endógenos y exógenos. Los factores endógenos son, por ejemplo, las características genéticas de las especies o variedades y las cuestiones fisiológicas. Entre los factores exógenos destacan las condiciones ambientales y de cultivo, especialmente el riego y la fertilización (Agustí, 2020). La producción también viene determinada por la edad de los árboles, y por el hecho de que alcancen un volumen de producción comercial (> 50 kg/árbol) en edad adulta (> 7 años después del trasplante) (Orduz Rodríguez *et al.*, 2007).

Los modelos para estimar el rendimiento se basan en la ecuación que se describe a continuación:

$$\textit{Producción} = \textit{superficie} \times \textit{peso del fruto medio} \times \textit{carga}$$

La primera componente se refiere a la superficie implantada con la variedad sobre la cual se desea realizar el pronóstico, combinada con una densidad de plantación, con lo que se obtendrá una cantidad total de árboles en producción y generalmente se determina mediante un censo. La segunda componente, que concierne al tamaño o al peso que los frutos tendrán al momento de la cosecha, se estima a través del modelado de curvas de crecimiento de los frutos, descritas en el Capítulo III. La tercera componente corresponde a la estimación de carga, es decir la cantidad de frutos a producir, constituye un desafío debido a la alta variabilidad presente entre árboles de un mismo huerto, entre huertos y también entre temporadas, lo que genera errores de estimación importantes. Ya fue abordada en el Capítulo V mediante técnicas estadísticas clásicas y se retoma en este Capítulo empleando Aprendizaje Automático, junto con la incorporación de información de otra naturaleza, menos costosa, pero que puede contribuir a mejorar las estimaciones.

En este Capítulo se analiza información de campo, meteorológica y proveniente de sensores remotos aplicando diferentes métodos de Aprendizaje Automático, con la finalidad de encontrar modelos que permitan la estimación de la carga por árbol con los menores errores de estimación.

### 6.1.1. Estimación con información de campo

En Argentina, Reeb & Tassile (2007) establecieron para cultivos de manzanas y peras, que la carga de los frutos, expresada en unidades por planta o por hectárea, es un parámetro que debe estimarse cada año dado que su valor depende de las condiciones que caracterizan a la temporada. Debe determinarse para cada especie y variedad por separado, estratificando por sistema de conducción y edad del monte debido a la heterogeneidad existente en las plantaciones.

En el cultivo de olivo, el Instituto de Desarrollo Rural<sup>18</sup> en la provincia de Mendoza, utiliza un método de recuento para estimar la cantidad de frutos, que consiste en medir el perímetro del tronco, en una muestra de 4 ramas al azar medir la longitud de 25 brindillas y contar el número de aceitunas que contienen. A través de la determinación del número de frutos por metro de brindillas, la proporción de brindillas con y sin frutos y el área de sección transversal del tronco se estima número de frutos por planta.

Para estimar la carga en naranjo dulce Valencia late en condiciones de campo, Avanza (2011) evaluó diferentes métodos basados en el recuento total y del número de frutos en secciones predefinidas de copa. Este último método tiene el inconveniente que se requiere estimar el volumen total de la copa de los árboles para calcular la carga por árbol y no hay consenso en las ecuaciones que deben utilizarse para cada especie y variedad en cuestión. Con la finalidad de estimar el volumen, Morin (1980) especularon que la copa de los cítricos es redondeada, en naranjo Valencia Orduz Rodríguez *et al.* (2007) y en mandarino Arrayana Monteverde *et al.* (1996), consideraron la forma de la

---

<sup>18</sup> Disponible en: <https://www.idr.org.ar/>

copa como una esfera, y otros autores como Zekri (2000) y Georgiou (2000) utilizaron una semiesfera alargada en mandarinas Nova y Clementina.

Falivene & Hardy (2008) en el Estado de Nueva Gales del Sur, para medir la carga en cítricos utilizaron un marco de recuento de 0,125 m<sup>3</sup> (0,5 m x 0,5 m x 0,5 m), que se coloca en la copa del árbol a una altura de 1,0-1,5 m del suelo y aconsejan repetir el procedimiento de muestreo al menos 40 veces en un bloque de 1 hectárea.

### 6.1.2. Uso de sensores remotos

Actualmente la agrotecnología y la agricultura de precisión han surgido como nuevos campos científicos que utilizan enfoques con gran cantidad de datos para impulsar la productividad agrícola, al tiempo que minimizan su impacto medioambiental. Los datos generados en las operaciones agrícolas modernas proceden de una variedad de sensores diferentes que permiten comprender mejor el entorno operativo (una interacción de condiciones dinámicas de cultivo, suelo y clima) y la propia operación (datos de la maquinaria), lo que conduce a una toma de decisiones más precisa y rápida (Liakos *et al.*, 2018).

#### 6.1.2.1. Sensores remotos

Los sensores remotos permiten hacer estimaciones de producción de forma anticipada y con un costo menor que las técnicas tradicionales basadas en información de campo. Además de su objetividad, la observación por teledetección proporciona datos homogéneos que pueden ser registrados geográfica y temporalmente. Sin embargo, el uso de la teledetección no es tan eficaz para estimar el rendimiento de los cultivos como para estimar la superficie ocupada, donde la observación es más directa y es posible encontrar un método aplicable a diferentes cultivos, zonas y temporadas de crecimiento.

En las últimas décadas, debido a los avances en informática, robótica, inteligencia artificial y sensores remotos, se han desarrollado diferentes opciones para estimar el rendimiento de los cultivos dentro de la llamada agricultura de precisión. Estos métodos semiautomatizados utilizan imágenes y aplican técnicas de segmentación (Teixidó *et al.*, 2012). Diversos métodos de captura de imágenes han sido probados con éxito, como el caso de Wang *et al.* (2021) quienes desarrollaron una tecnología, basada en el aprendizaje profundo, para imágenes de teledetección óptica de alta resolución.

Se ha demostrado que la teledetección es un instrumento eficaz para vigilar las prácticas de cultivo. Debido a una gran variedad de sensores a bordo en un número cada vez mayor de satélites civiles, los valores espectrales y las propiedades temporales de la superficie terrestre resultantes de las prácticas humanas pueden ser capturadas y vigiladas a diferentes escalas espaciales y temporales (Begué, 2018).

Hewson (2014) establece que el sensoramiento remoto o teledetección, es el proceso mediante el cual se detecta la energía emitida o reflejada por un objeto a una longitud de onda determinada en el espectro electromagnético, sin estar en contacto directo con el objeto. El ojo humano, por ejemplo, detecta una fracción relativamente pequeña del espectro total de energía que emite el sol, esta es la fracción visible del espectro electromagnético. La cantidad y el tipo de energía que se detecta se suele registrar en formato digital donde el tipo se representa por el registro de la señal en el espectro y la cantidad por la fuerza de la señal.

Los sensores remotos aéreos generan interés ya que se pueden usar en regiones de gran tamaño o en países enteros, dependiendo del tipo de equipo, del enfoque de muestreo y de los recursos disponibles. Los sensores remotos satelitales son los más comunes gracias a la cobertura total y repetida que ofrecen una o más fuentes de datos satelitales, permitiendo así el monitoreo nacional de aplicaciones terrestres. Al mayor nivel, existen dos tipos generales de sensores remotos para monitorear la cobertura de la tierra: pasivos y activos (Hewson, 2014).

Debido a la gran variedad de sensores a bordo de los satélites civiles, los valores espectrales y las propiedades temporales de la superficie de la tierra resultantes de las prácticas humanas pueden ser capturados y monitoreados a diferentes escalas espaciales y temporales (Begué, 2018). De hecho, las imágenes satelitales de baja



resolución se han utilizado ampliamente para el seguimiento de los cultivos y la previsión del rendimiento durante más de 30 años y desempeñan un papel importante en un número cada vez mayor de sistemas operativos. La combinación de una alta frecuencia temporal con su amplia cobertura geográfica, generalmente asociada a un bajo costo por unidad de superficie, hace que estas imágenes sean una opción conveniente. Además, la continuidad de los sistemas existentes sigue siendo crucial para garantizar la disponibilidad de series temporales largas, tal y como necesitan la mayoría de los métodos de predicción del rendimiento utilizados en la actualidad (Rembold *et al.*, 2013).

El espectro radiómetro de imagen de resolución moderada (MODIS) tiene una resolución espacial y temporal que permite la vigilancia y el mapeo de la vegetación a nivel regional nivelado. Con mínimas variaciones asociadas a las influencias externas (atmósfera, vista y ángulos solares), los productos de índice de vegetación del MODIS proporcionan comparaciones espaciales y temporales consistentes de las condiciones de la vegetación (Lopesti *et al.*, 2015). Además, es uno de los más antiguos, por lo que permite combinar la información que proporciona con los datos históricos de las cosechas.

#### *6.1.2.2. Índices de vegetación*

Los indicadores de teledetección se utilizan en la agricultura para la vigilancia de las condiciones de los cultivos y la previsión rendimiento. Hay índices que se centran en la capacidad de captación de luz y las tasas de fotosíntesis, como el Índice de Área Foliar (LAI) y la Fracción de Absorción de Radiación Fotosintética Activa (FPAR). Entre los indicadores más ampliamente utilizados en agricultura se incluyen índices de vegetación como el Índice de Diferencia Normalizada de Vegetación (NDVI), que permite seguir el progreso de los cultivos y la evolución de la biomasa verde (Jurečka *et al.*, 2018). El NDVI está estrechamente relacionado con el vigor de la vegetación y ha sido reconocido por su capacidad para vigilar los cultivos y como estimador de los rendimientos desde principios de los años ochenta.

Lopresti *et al.* (2015) concluyen que la estimación del rendimiento por el NDVI, obtenido de manera temprana, rápida y económica, podría considerarse un complemento prometedor de las evaluaciones de rendimiento basadas en encuestas. En la India se ha comprobado que la estimación del rendimiento de los cultivos, mejora sustancialmente después de combinar los datos espectrales de los satélites con los datos de los estudios basados en experimentos de corte de cultivos (Parihar, 1987).

El NDVI es un indicador numérico que utiliza las bandas espectrales roja y cercana al infrarrojo y está altamente asociado con el contenido de vegetación. Los valores altos de NDVI corresponden a áreas que reflejan más en el espectro del infrarrojo cercano, que corresponde a una vegetación más densa y saludable<sup>19</sup>. La combinación de su formulación diferencial normalizada y el uso de las regiones de mayor absorción y reflectancia de la clorofila la hacen robusta en una amplia gama de condiciones. Sin embargo, puede saturarse en condiciones de vegetación densa cuando el LAI se vuelve alto.

Lopresti *et al.* (2015) determinaron que el NDVI permite estimar tempranamente el rendimiento del trigo y encontraron fuertes coeficientes de correlación entre la estimación del modelo empírico calibrado con los rendimientos oficiales. Arango *et al.* (2016 a) mostraron cómo mapear las tierras de cultivo utilizando imágenes del sensor MODIS y un conjunto de variables asociadas como las propiedades del suelo, las características biofísicas de los sitios de cultivo, los tratamientos culturales y la producción. Asimismo, Arango *et al.* (2016 b) desarrollan un método para detectar tierras cultivables utilizando imágenes del sensor MODIS, complementadas con un conjunto de diferentes índices de vegetación, todo ello procesado bajo algoritmos de aprendizaje automático.

---

<sup>19</sup> Disponible en: <https://www.geo.university/>

### 6.1.3. Aprendizaje automático

El Aprendizaje Automático (ML por su denominación en inglés) es un campo multidisciplinar relacionado con la Estadística, la Teoría de la Información, la Teoría de los Juegos y la Optimización. Puede verse, también, como una rama de la Informática y de la Inteligencia Artificial. Trata de utilizar los puntos fuertes y las capacidades especiales de las computadoras para complementar la inteligencia humana, lo que permite detectar patrones que la percepción humana no sería capaz de reconocer. A diferencia de la Estadística, en el ML las consideraciones algorítmicas juegan un papel importante y se desarrollan algoritmos para realizar las tareas de aprendizaje teniendo en cuenta su eficiencia computacional. Mientras la Estadística está a menudo interesada en el comportamiento asintótico, la teoría del Aprendizaje Automático se centra en los límites de las muestras finitas.

El campo del Aprendizaje Automático involucra la cuestión de cómo construir programas informáticos que mejoran automáticamente con la experiencia (Hastie *et al.*, 2017; James *et al.*, 2017). En los últimos años se han desarrollado muchas aplicaciones exitosas de ML, que incluyen programas de minería de datos y sistemas de filtrado de información, entre otros, que abarcan un vasto conjunto de herramientas para comprender los datos. A continuación, se describen algunos métodos de ML, poniendo énfasis en los utilizados en este Capítulo.

La mayoría de los problemas de aprendizaje puede ser clasificados en dos categorías: supervisados o no supervisados. En esta tesis vamos a centrarnos en el caso de los aprendizajes supervisados, en los que para cada observación de la variable predictora ( $x_i$ ,  $i = 1, \dots, n$ ) existe una respuesta asociada  $y_i$ , el desafío consiste en encontrar un modelo que relacione la variable respuesta con la predictora, con el objetivo de predecir con la mayor precisión la respuesta para observaciones futuras (en problemas de predicción) o comprender mejor la relación entre la variable respuesta y la predictora (en problemas de inferencia). Muchas técnicas clásicas de aprendizaje, como la Regresión Lineal y la Regresión Logística, así como enfoques más modernos como las Máquinas de Vectores Soporte, operan en el dominio del aprendizaje supervisado. Dentro del aprendizaje supervisado se pueden distinguir dos situaciones, la

Clasificación, cuando se trata de variables categóricas y la Regresión, cuando las variables sean cuantitativas (James *et al.*, 2017).

Todos los métodos pertenecientes a esta clase comparten el hecho de que se realizan en dos etapas:

✓ *Etapa de calibración o entrenamiento*, durante la cual se desarrollan algoritmos de resolución de acuerdo con las restricciones matemáticas de cada método, en base a la muestra de objetos de estudio previamente clasificados.

✓ *Etapa de validación*, donde se pone a prueba el modelo con muestras problemas pertenecientes a clases o grupos conocidos y se calcula la tasa de error de clasificación.

#### *6.1.3.1. Métodos basados en árboles de decisión*

Los árboles de decisión son técnicas de ML que permiten la construcción de modelos predictivos basados en su clasificación según ciertas características o propiedades, o en la regresión mediante la relación entre distintas variables para predecir el valor de otra (James *et al.*, 2017).

Son estructuras formadas por ramas y nodos de distintos tipos. Los nodos internos representan cada una de las características o propiedades a considerar para tomar una decisión, las ramas representan la decisión en función de una determinada condición y los nodos finales u hojas representan el resultado de la decisión.

En la Figura 6.1, se representa un árbol de decisión para el archivo *Iris*, construido a partir de información de 4 variables morfológicas de las flores (longitud de sépalo -SepalLen-, ancho sépalo -SepalWid-, longitud del pétalo -PetalLen- y ancho de pétalo -PetalWid-), en 3 variedades del género *Iris* (*setosa*, *versicolor* y *virginica*). El árbol parte del conjunto formado por las 150 unidades bajo estudio y va abriendo en ramas a partir de nodos según valores definidos para las diferentes variables, hasta llegar a los nodos finales u hojas del árbol.

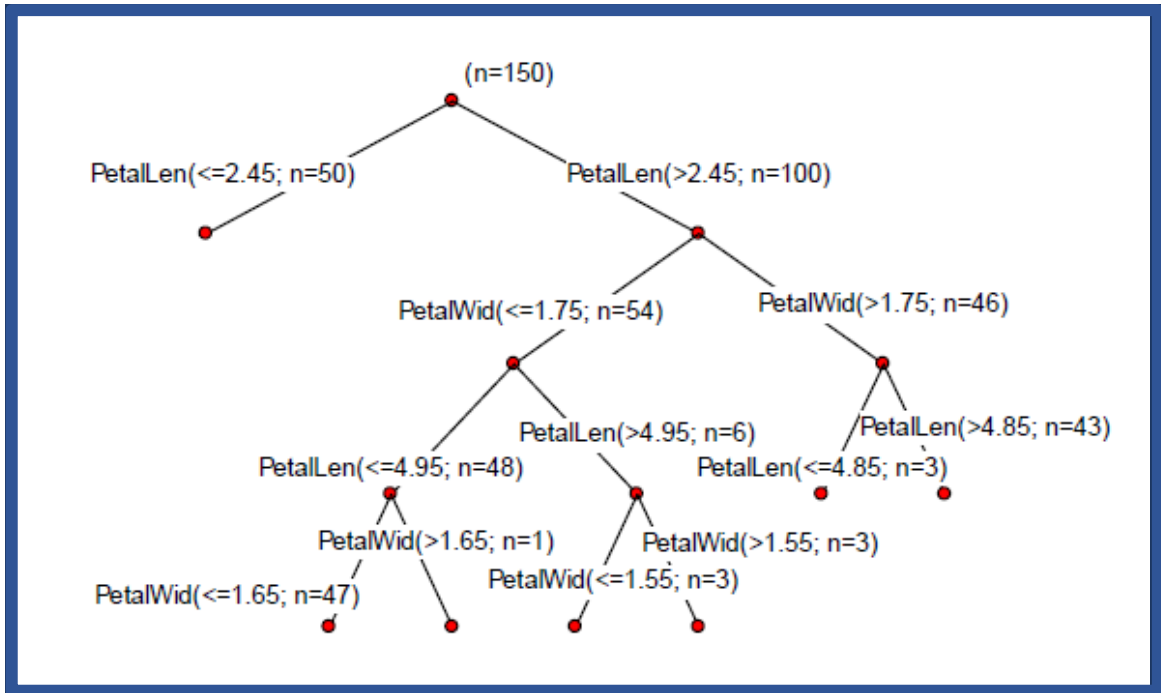


Figura 6.1. Ejemplo de un Árbol de decisión. Archivo *Iris*. (Balzarini et al., 2008)

La primera separación se realiza en base a valores de largo de pétalo, en menores e iguales a 2,45 (50 ejemplos) y mayores a 2,45 (100 ejemplos). De una de estas ramas a su vez se separan, en base al ancho de pétalo, los ejemplos menores e iguales a 1,75 (54 ejemplos) y los mayores de 1,75 (46 ejemplos) y así el proceso continúa. La medida de calidad de los nodos, el forma en la que se ramifica, la técnica de poda y la condición de parada son las características fundamentales que determinan la construcción del árbol.

A continuación, se detallan algunas de las técnicas basadas en árboles de decisión.

#### 6.1.3.1.1. Árboles de Regresión

Los árboles de regresión (RT por su denominación en inglés) son un tipo de árboles de predicción que se aplican cuando la variable respuesta es continua. En el entrenamiento de un RT, las observaciones se van distribuyendo por bifurcaciones (nodos) generando la estructura del árbol hasta alcanzar un nodo terminal. Cuando se quiere predecir una nueva observación, se recorre el árbol acorde al valor de sus predictores hasta alcanzar uno de los nodos terminales. La predicción del árbol es la

media de la variable respuesta de las observaciones de entrenamiento que están en ese mismo nodo terminal.

Son diversos los algoritmos utilizados para la construcción de árboles de regresión, algunos de ellos se presentan a continuación.

### **CART (*Classification and Regression Trees*)**

El método consta de dos componentes, un árbol con nodos terminales y un vector de  $\vec{\theta}_i = (\theta_1, \theta_2, \dots, \theta_b)$ ; donde  $\theta_i$  se asocia al nodo terminal.

Dado un conjunto de muestras, CART identifica una variable de entrada y un punto de ruptura, antes de dividir las muestras en dos nodos. Partiendo de todo el conjunto de muestras de entrenamiento disponibles (nodo raíz), se realiza una partición binaria recursiva para cada nodo hasta que no sea posible ninguna otra división o se satisfaga un determinado criterio de terminación. El mecanismo de CART pretende producir no un árbol, sino una secuencia de árboles anidados y podados; el árbol óptimo se identifica evaluando el rendimiento predictivo de cada árbol de la secuencia de poda en datos de prueba independientes (Steinberg, 2009). El valor estimado de una hoja es constante en CART y los valores de los nodos terminales se reducen a la media de las observaciones en esa región (Loh, 2011).

### **M5-Prime**

Es un método que construye árboles de regresión que producen una clasificación, basada en funciones lineales a trozos (Wang & Witten, 1997). Para ello, el espacio se divide en un conjunto de regiones. Además, el valor predicho se ajusta dentro de cada región utilizando un modelo lineal. Suponiendo un conjunto de entrenamiento con ejemplos, cada uno definido por su valor en un conjunto de atributos (discretos o continuos) y un objetivo continuo, el método construye un modelo que relaciona los valores objetivo de los ejemplos de entrenamiento con los valores de las variables que definen el ejemplo.

M5-Prime selecciona la división que maximiza la reducción de error esperada. Una vez construido el árbol, se calcula un modelo lineal multivariante para los ejemplos de cada nodo del árbol con técnicas de regresión estándar y utilizando sólo los atributos a los que se refieren las pruebas o los modelos lineales en algún lugar del subárbol bajo este nodo. Las principales características de este método son la construcción del árbol inicial, la poda del árbol y el suavizado.

En la construcción del árbol inicial el criterio de división se basa en tratar la desviación estándar de los valores de clase que llegan a un nodo como una medida del error en ese nodo, y en calcular la reducción esperada del error como resultado de probar cada atributo en ese nodo. La reducción de la desviación estándar (SDR) se calcula mediante la fórmula:

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} \times sd(T_i)$$

Donde T es el conjunto de ejemplos que llegan al nodo y  $T_i$  son los conjuntos que resultan de dividir el nodo según el atributo elegido.

El procedimiento de poda del árbol hace uso de una estimación del error esperado que se experimentará en cada nodo para los datos de prueba. En primer lugar, se promedia la diferencia absoluta entre el valor predicho y el valor real de la clase para cada uno de los ejemplos de entrenamiento que llegan a ese nodo. Esta media subestimaré el error esperado para los casos no vistos, y para compensar, se multiplica por el siguiente factor:

$$\frac{n + v}{n - v}$$

Donde n es el número de ejemplos de entrenamiento que llegan al nodo y  $v$  es el número de parámetros del modelo que representa el valor de la clase en ese nodo.

M5 calcula un modelo lineal para cada nodo interior del árbol no podado. El modelo se calcula mediante una regresión estándar, utilizando sólo los atributos que se comprueban en el subárbol situado debajo de este nodo. El modelo lineal resultante se simplifica eliminando términos para minimizar el error estimado calculado utilizando el factor de multiplicación anterior, eliminar un término disminuye el factor de

multiplicación, lo que puede ser suficiente para compensar el inevitable aumento del error medio en los ejemplos de entrenamiento. Los términos se eliminan uno a uno, con avidez, siempre que la estimación del error disminuya. Por último, una vez que se ha establecido un modelo lineal para cada nodo interior, se poda el árbol desde las hojas, siempre que el error estimado esperado disminuya.

La última etapa consiste en utilizar un proceso de suavizado para compensar las discontinuidades bruscas que inevitablemente se producirán entre los modelos lineales adyacentes en las hojas del árbol podado, especialmente en el caso de algunos modelos construidos a partir de un pequeño número de instancias de entrenamiento. El procedimiento de suavizado descrito por Quinlan (1992) utiliza primero el modelo de la hoja para calcular el valor predicho, y luego filtra ese valor a lo largo del camino de vuelta a la raíz, suavizándolo en cada nodo combinándolo con el valor predicho por el modelo lineal para ese nodo. El cálculo es

$$p' = \frac{nq + kr}{n + k}$$

Donde  $p$  es la predicción,  $q$  es el valor predicho por el modelo en este nodo,  $n$  es el número de instancias de entrenamiento que llegan al nodo de abajo y  $k$  es una constante.

#### *6.1.3.1.2. Árboles de inferencia condicional (CTREE)*

Los algoritmos de partición recursiva son casos especiales de un algoritmo simple de dos etapas, primero la partición de las observaciones por separación univariante de forma recursiva y segundo el ajuste de un modelo constante en cada celda de la partición resultante. Realiza una búsqueda exhaustiva sobre todas las posibles particiones maximizando una medida de información de impureza del nodo seleccionando, la covariable que muestra la mejor partición. Este enfoque tiene dos problemas fundamentales, el sobreajuste y un sesgo de selección hacia covariables con muchas divisiones posibles (Horton, 2006). Para cada nodo, se cuantifica la asociación entre cada característica de entrada independiente y la variable de salida, utilizando la



prueba de permutación y la corrección de pruebas múltiples. Si la asociación más fuerte supera un umbral estadístico, se realiza una división binaria en esa variable de entrada correspondiente; de lo contrario, el nodo actual es un nodo terminal.

#### *6.1.3.1.3. Métodos de ensamble: Bosques Aleatorios*

La técnica de Bosques Aleatorios (RF por su denominación en inglés) funciona bajo el ensamble de algoritmos de bajo aprendizaje (árboles de decisión, en este caso) para mejorar el porcentaje de acierto de la técnica global. Los algoritmos de ensamble o conjunto, como los RF, tienden a ser más robustos a los cambios en el conjunto de datos. Se puede decir que es un método robusto al ruido, esto implica que pequeños cambios en el conjunto de datos de entrenamiento, tendrán un impacto bajo o ínfimo si es que lo tuviese.

Si se emplea un solo árbol de decisión es posible que se presente un sobreajuste de los datos, pero mediante la conjunción de muchos árboles usando diferentes variables, el ajuste del modelo a los datos resulta diferente. La aleatoriedad introducida por el método se presenta en la selección de observaciones y variables. Es esta aleatoriedad que permite que el método sea robusto al ruido, a los datos extremos y al sobreajuste, cuando se lo compara con un solo árbol de decisión. La aleatoriedad también tiene beneficios en cuanto al costo computacional. En resumen, el modelo de los bosques aleatorios es una buena elección por varios motivos. A menudo, muy poco preprocesamiento de los datos es necesario realizar, los datos no necesitan ser normalizados y es bastante robusta a los datos extremos. La necesidad de selección de variables se evita ya que el algoritmo lo realiza durante el análisis. Como muchos árboles se construyen a dos niveles de aleatoriedad (observaciones y variables), cada árbol es un modelo independiente y el modelo resultante tiende a no sobreajustar el conjunto de datos de entrenamiento (Williams, 2011).

### *6.1.3.2. Métodos vagos, K-vecinos más cercanos*

Son métodos de aprendizaje basados en instancias que predicen el valor de la variable de salida basándose en los valores de sus vecinos más cercanos en el conjunto de entrenamiento. El fundamento de este método supone que todas las características utilizadas para describir el conjunto de datos son relevantes y que es probable que los ejemplos cercanos tengan el mismo valor de salida. Básicamente, calcula las distancias (euclidianas u otras) entre el ejemplo a clasificar y cada ejemplo de entrenamiento, seleccionando como vecinos los  $K$  más cercanos al ejemplo a clasificar. Como se basa en las distancias, el método de los K-Vecinos más Cercanos (KNN) es bastante sensible a la escala de deslizamiento. Por otro lado, el parámetro  $k$  no se conoce de antemano y es necesario fijarlo experimentalmente, KNN puede ser útil cuando la interpretabilidad no es un requisito para modelar la solución de un problema de predicción (Diaz, 2021).

### *6.1.3.3. Redes Neuronales Artificiales*

Las Redes Neuronales Artificiales (ANN por su denominación en inglés) se inspiran en las redes neuronales biológicas y se han aplicado en una gran variedad de problemas de clasificación y regresión (Cook, 2020).

Las ANN son un sistema de procesamiento de información que tiene propiedades inspiradas en las redes neuronales biológicas: el procesamiento de información ocurre en muchos elementos simples llamados neuronas. Las señales son transferidas entre neuronas a través de enlaces de conexión. Cada conexión tiene un peso asociado representando la sinapsis. Cada neurona aplica una función de activación a su entrada de red para determinar su salida, representando la generación de potenciales de acción. En la siguiente Figura se presenta la estructura básica de una red neuronal.

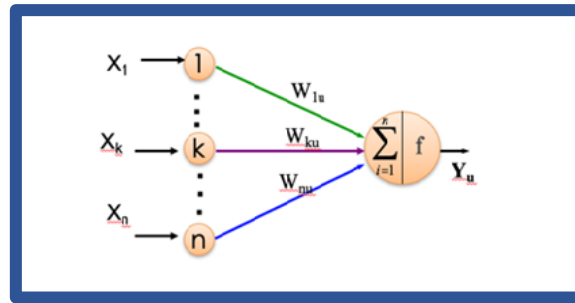


Figura 6.2. Representación gráfica de una red neuronal

Las ANN utilizan las neuronas como unidades mínimas de procesamiento conectadas entre sí, para producir modelos complejos. Existen muchos tipos de redes neuronales, en función de las siguientes características:

- Función de activación, que transforma las señales de entrada combinadas de una neurona en una sola señal de salida para ser transmitida a través de la red.
- La topología de la red que describe el número de neuronas en el modelo, el número de capas y la forma en que la que están conectadas.
- El algoritmo de entrenamiento con el que se aprenden los pesos de las conexiones, en función de los cuales se inhibe o excita una señal.

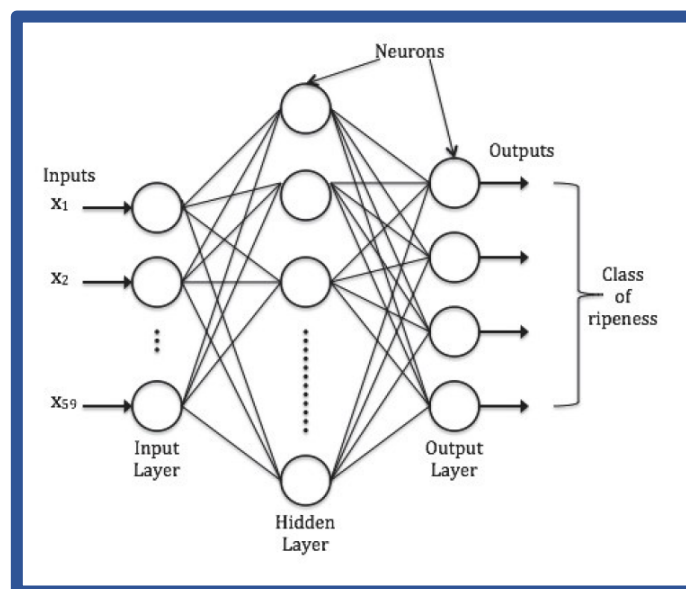


Figura 6.3. Estructura de una red neuronal multicapa (Norasyikin *et al.*, 2012)

La primera capa de la red neuronal se conoce como capa de entrada o *input layer* y recibe los datos en bruto, es decir los valores predictores. La capa intermedia (conocida como capa oculta o *hidden layer*), recibe los valores de la capa de entrada, ponderada por los pesos. La última capa llamada *ouput layer*, combina los valores que salen de la capa intermedia para generar la predicción.

Los perceptrones multicapa (MLP) son redes sin bucles cuyas neuronas están dispuestas en capas, en las que cada neurona proporciona entrada sólo a las neuronas de la siguiente capa de la secuencia. La primera capa contiene una neurona asociada a cada característica del conjunto de datos. La última capa contiene las neuronas de salida. El número de capas ocultas no se establece de antemano y se suele determinar experimentalmente. Durante el proceso de aprendizaje, las conexiones entre dos neuronas situadas en capas consecutivas se ajustan para minimizar el error de predicción. Las ANN resuelven problemas de clasificación bastante complejos en muchos campos diferentes. El principal inconveniente de este enfoque es la interpretabilidad.

#### *6.1.3.4. Máquinas de Vectores Soporte*

Las máquinas de vectores soporte (SVM) son un enfoque para la clasificación que se desarrolló en la comunidad de Ciencias de la Computación en la década de 1990 y que ha crecido en popularidad desde entonces, demostrando funcionar bien en una variedad de ajustes. Originalmente resolvían problemas de clasificación binaria, aunque las implementaciones actuales permiten resolver problemas multiclase y de regresión.

Este método resuelve un problema de optimización cuya solución es un hiperplano de separación de los datos en dos clases si el conjunto de datos a tratar es linealmente separable, como los datos que se muestran en la Figura 6.4.

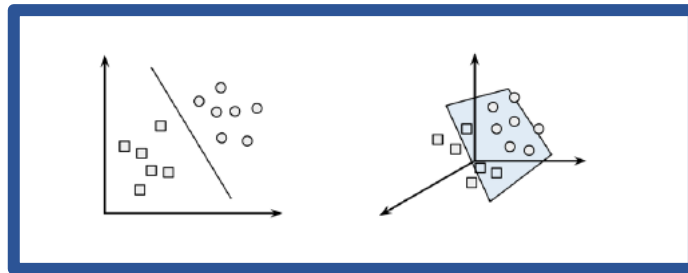


Figura 6.4. Ejemplo de frontera de separación para un problema bidimensional (izquierda) o tridimensional (derecha).

La Figura 6.4., muestra cómo los ejemplos están perfectamente separados según su clase (círculos o cuadrados) por una línea recta (hiperplano en un espacio de dimensión 2) o una superficie (un hiperplano en un espacio de dimensión 3).

SVM decide, entre todas las posibilidades factibles, aquella que se corresponde con el Hiperplano de Margen Máximo, que es el que crea la mayor separación entre las dos clases. La Figura 6.5 muestra un ejemplo de hiperplanos de separación para un problema bidimensional.

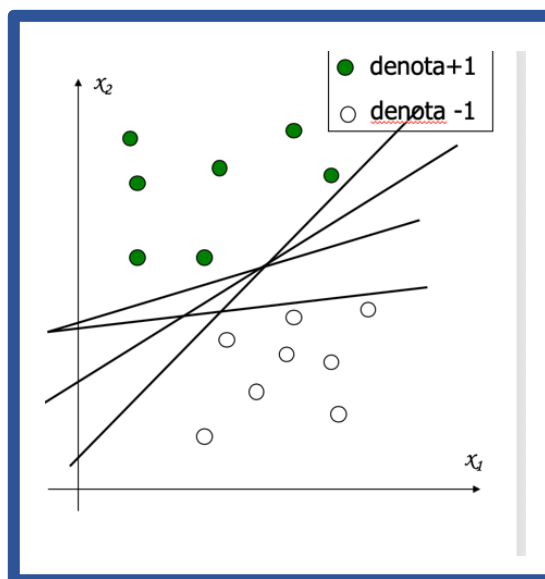


Figura 6.5. Diferentes separadores lineales para un conjunto de datos

Cualquiera de las líneas que separan los círculos blancos de los verdes clasificaría correctamente todos los datos, es probable que la línea que conduce a la mayor separación generalice mejor los datos futuros. El margen máximo mejorará la posibilidad de que, a pesar de exista ruido, los ejemplos se clasifiquen mejor.

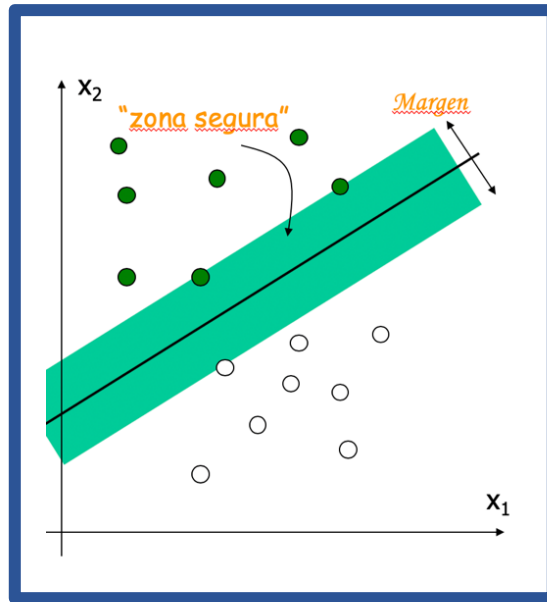


Figura 6.6. Margen máximo de separación para un conjunto de datos

Los vectores soporte son los puntos de cada clase que son los más cercanos al margen máximo. Cada clase debe tener al menos un vector de soporte, pero es posible tener más de uno. De hecho, solo se utilizan los vectores soporte para determinar el margen máximo. Y esta es la característica fundamental de las SVM dado que estos vectores soporte generan el modelo de clasificación.

El esquema de funcionamiento anterior permite resolver cualquier tipo de problema introduciendo funciones *kernel* que transforman el espacio de entrada, posiblemente no linealmente separable, en un espacio de características donde el conjunto sí es linealmente separable. La solución que proporciona SVM es un subconjunto del conjunto de entrenamiento que se denomina conjunto de vectores soporte.

Los vectores proporcionan una forma muy compacta de almacenar un modelo de clasificación, incluso si el número de características es extremadamente grande (Hastie *et al.*, 2017; James *et al.*, 2017).

En muchas aplicaciones reales, las relaciones entre variables son no-lineales. Una característica fundamental del algoritmo SVM es su habilidad para mapear el problema en dimensiones más grandes usando un proceso conocido como el truco del kernel (Figura 6.7). Empleando esta técnica una relación no-lineal puede aparecer como lineal.

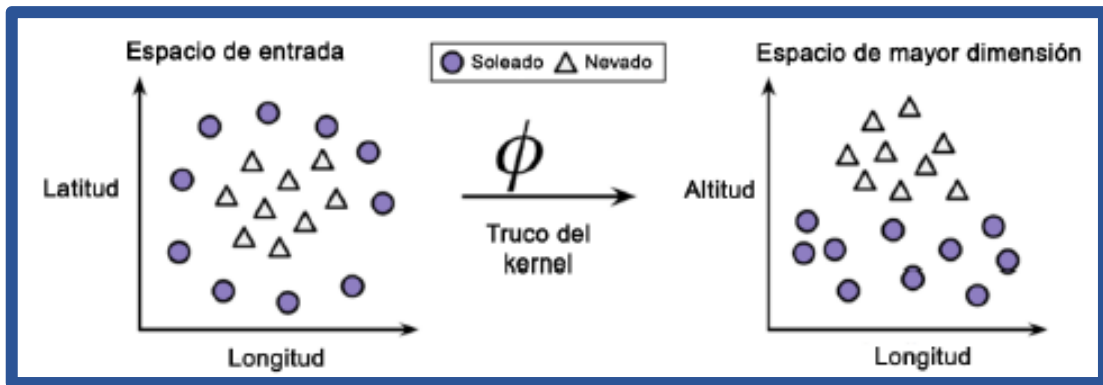


Figura 6.7. El truco de Kernel  
Fuente: Lantz, 2015

En la primera imagen se observa un grupo de observaciones que muestra una disposición no lineal en un gráfico de dos dimensiones, cuando se grafican la longitud y la latitud. Sin embargo, en la figura de la derecha, luego de aplicar el truco del kernel, se pueden observar los datos con respecto a otra variable que es la altitud. Se observa que ahora las clases son linealmente separables, porque se observan los datos desde una nueva perspectiva (Lantz, 2015).

La técnica de SVM con kernels no lineales agrega dimensiones adicionales a los datos para crear una separación en este sentido. En esencia, el truco del kernel implica un proceso de nuevos atributos que expresan la relación matemática entre los atributos medidos. En principio, la altitud puede ser expresada matemáticamente como una interacción entre la latitud y la longitud. Esto permite a la técnica de SVM aprender nuevos conceptos que no fueron medidos explícitamente en el conjunto de datos originales (Lantz, 2015).

Uno de los principales inconvenientes de las SVM es que la fase de entrenamiento es muy costosa, debido a su complejidad computacional, lo que se traduce en que es un algoritmo inaplicable en grandes conjuntos de datos. Sin embargo, este problema se está solucionando con éxito. Una posibilidad es dividir un gran problema de optimización en una serie de problemas más pequeños, en los cuales solamente se usan un par de variables cuidadosamente seleccionadas de modo que la optimización se pueda realizar recientemente. El proceso se repite hasta que se resuelven satisfactoriamente todos los problemas de optimización en los que se descompuso el problema original.

#### 6.1.4. Selección de características

Las técnicas selección de características permiten reducir el tamaño del conjunto de datos a priori, eligiendo las variables más relevantes para el caso de estudio, a partir de las disponibles (Liu & Motoda, 1998; Liu, 2005). Sin embargo, en algunos casos, debido a las características del problema, es preferible no realizar esta selección a priori ya que puede generar clasificadores que dependan en gran medida de las características seleccionadas, por lo que una mala reducción puede provocar una pérdida en el poder discriminante del clasificador.

Por otro lado, algunos métodos supervisados presentan el inconveniente de su baja o nula interpretabilidad, de modo que, aun obteniendo resultados de clasificación altamente satisfactorios, no son usables dado que no proporcionan información relativa a las variables que están influyendo en la toma de decisiones. En estos casos es deseable desarrollar un método en el que, aunque sea a posteriori, se puedan identificar las variables relevantes que han intervenido en el proceso de clasificación.

La identificación de características se emplea con alguna de las siguientes finalidades (Jain *et al.*, 2000):

- ✓ Mejorar tanto el proceso como el modelo de un problema.
- ✓ Mejorar la comprensión del proceso que genera los datos.
- ✓ Mejorar la eficiencia del predictor (costo de implementación y ejecución).

Cuando estas características se identifican a posteriori, es decir cuando el proceso de aprendizaje ha finalizado, su finalidad es comprender el modelo.

Para identificar estas características a posteriori asumimos que si el rendimiento de un método empeora cuando entrenamos con todas las variables menos una, entonces esa variable es importante para la construcción del modelo. De ese modo, cuanto más se degrade el aprendizaje (más error) al eliminar una variable, más relevante será esta. De este hecho se puede inducir un orden sobre las variables atendiendo a este criterio. Este proceso, para evitar sesgos, se repite  $n$  veces (típicamente 50), de modo que



finalmente tenemos  $n$  rankings de variables que es preciso agregar, para obtener la importancia final de las variables. El procedimiento general se detalla a continuación:

*Para cada repetición ( $j=1\dots n_r$ )*  
*Para cada variable  $f_i$*   
*Entrenar el método usando todas las variables salvo  $f_i$*   
*Evaluar el error asociado a este entrenamiento  $e_i$*   
*Ordenar las variables  $f_i$  según  $e_i$  (en orden decreciente), obteniendo un ranking de variables  $r_j$ .*  
*Agregar el conjunto de rankings  $\{r_j, j=1, \dots, 50\}$*

Para realizar el último paso de este procedimiento empleamos reglas de clasificación por puntuación que asignan una puntuación a cada candidato en función de su posición en un ranking. Las reglas de clasificación de puntuación convexas otorgan mayores recompensas a los candidatos en mejores posiciones. Una opción que puede resultar adecuada para esta selección es el Método de Recuento de Borda, un proceso electoral propuesto por el matemático francés Jean-Charles Borda en 1770, según el cual los candidatos son ordenados de acuerdo con las preferencias de cada elector.

El recuento de Borda es una regla de clasificación convexa bastante sencilla que es una de las funciones de elección social más populares debido a su simplicidad. Trata de elegir a los candidatos ampliamente aceptables, en lugar de los preferidos por la mayoría. Por esta razón, se utiliza aquí para encontrar el ganador de consenso. Para cada clasificación de un conjunto de clasificación, un candidato  $c_i$  es recompensado con tantos puntos como el número de candidatos que están clasificados en una posición peor que  $c_i$ . Los puntos obtenidos por cada candidato en cada clasificación se suman para obtener una puntuación final para ese candidato. A continuación, los candidatos se clasifican en orden decreciente, de modo que el candidato con la mayor puntuación es el ganador.

### 6.1.5. Antecedentes

Recientemente, la estimación del rendimiento mediante sistemas basados en Inteligencia Artificial, como por ejemplo el Aprendizaje Automático, está produciendo resultados prometedores en condiciones de campo (Gongal *et al.*, 2015; Kamilaris & Prenafeta-Boldu, 2018).

Varios autores emplearon con éxito técnicas de Aprendizaje Automático para estimar variables de interés en la agricultura. Taghizadeh-Mehrjardi *et al.* (2020) utilizaron SVM y RF, sobre un conjunto de variables (geomorfología, atributos del terreno y datos de teledetección) con el fin de predecir las clases de idoneidad del terreno para el trigo y la cebada de secano.

Refonaa *et al.* (2021) compararon un modelo de Regresión Lineal con un RF para estimar el rendimiento del maíz. Han *et al.* (2019) exploraron con una precisión aceptable algoritmos de regresión y de aprendizaje automático (regresión lineal, SVM, ANN y RF) para modelizar la biomasa de maíz sobre el suelo, utilizando datos de teledetección. Maya Gopal & Bhargavi (2019) evalúan las características más necesarias para la predicción precisa del rendimiento de los cultivos, utilizando RF, ANN, SVM y KNN. Abbas *et al.* (2020) estudiaron el potencial de los algoritmos de aprendizaje automático para la predicción del rendimiento de los tubérculos de patata utilizando datos sobre los componentes horizontales y verticales de la conductividad eléctrica del suelo, el contenido de humedad del suelo, la pendiente del campo, el pH del suelo, la materia orgánica del suelo y el NDVI.

El ML es una de las técnicas más utilizadas para la toma de decisiones relacionadas con la estimación del rendimiento, la gestión del suelo o la gestión de enfermedades en las plantas, entre otras. En la última década, se han realizado muchos trabajos para la estimación de rendimiento de frutos utilizando técnicas de ML debido a su prometedora capacidad (Yamamoto *et al.*, 2014; Dorj *et al.*, 2017).

Qureshi *et al.* (2017) propusieron un método para la detección de frutos en mango mediante el análisis de imágenes de copas de árboles de mango. Los autores identificaron los píxeles correspondientes a sectores con y sin frutas aplicando un

conjunto de filtros en la imagen de entrada y analizaron los límites de las frutas como una elipse. Esos datos se compararon con los algoritmos de ML como KNN y SVM.

Díaz *et al.* (2017) estudiaron la eficacia de las técnicas de ML para predecir la producción de los huertos cítricos e identificaron las variables que afectaban a la misma. Analizaron 964 huertos de limón, mandarina y naranja provenientes de Corrientes, Argentina. La producción se predijo mediante M5-Prime, un constructor de árboles de regresión de modelos que produce una clasificación basada en funciones lineales. Para todas las especies estudiadas, la variable más informativa fue la edad de los árboles; en los huertos de mandarinas y naranjas, la edad fue seguida por las distancias entre y dentro de filas; el riego también afectó a la producción de las mandarinas. Obtuvieron un coeficiente  $r = 0,80$  y un error medio absoluto de 0,1, lo cual indica que el M5-Prime es un método adecuado para clasificar los huertos de cítricos según la producción y, además, permite identificar las variables más informativas que afectan a la producción por árbol.

Debido a que los métodos de estimación de carga en los huertos cítricos utilizados actualmente en la región arrojan errores elevados, el objetivo del presente Capítulo fue encontrar procedimientos accesibles y de bajo costo, combinando diferentes selecciones de variables y técnicas de análisis de datos.

## 6.2. MATERIALES Y MÉTODOS

El trabajo se realizó en dos instancias, que se diferencian por la información de entrada, las variables utilizadas y las técnicas de análisis empleadas. La primera instancia se empleó únicamente información de campo y en la segunda instancia se agregó información climática y proveniente de sensores remotos.

## 6.2.1. Datos Experimentales

### 6.2.1.1. Huertos

La información de campo fue obtenida en algunos de los huertos descritos en el Capítulo II, los que se presentan en las Tablas 6.1 y 6.2, para cada instancia de trabajo.

Tabla 6.1. Huertos seleccionados para la instancia 1

Temporada	Huerto	Especie y variedad	Fecha plena flor	Momento estimación	n
2010/11	1.1	Tangor Murcott	23-09-10	1-2	40
2007/08			08-10-07	1-2	26
2008/09	5.1	Tangor Murcott	01-10-08	1-2	25
2009/10			02-10-08	1-2	24
2007/08			04-10-07	1-2	19
2008/09	6.1	Tangor Murcott	28-09-8	1-2	1-2
2009/10			30-09-09	1-2	17
2007/08			02-10-07	1-2	42
2008/09	7.1	Tangor Murcott	30-09-08	1-2	36
2009/10			30-09-09	1-2	41
2007/08			05-10-07	1-2	31
2008/09	8.1	Tangor Murcott	01-10-07	1-2	31
2008/09			30-09-09	1-2	31
2005/06	1.1	Naranja dulce Valencia late	03-10-05	1	17
2005/06	2.1	Naranja dulce Valencia late	08-10-05	1	25
2006/07			25-09-06	1-2	31
2007/08	3.1	Naranja dulce	10-10-07	1-2	41
2008/09		Valencia late	08-10-08	1-2	47
2009/10			12-10-09	1-2	51
2006/07	4.1	Naranja dulce Valencia late	15-10-06	1-2	23
2007/08			09-10-07	1-2	31
2008/09	5.1	Naranja dulce	04-10-08	1-2	25
2009/10		Valencia late	05-10-08	1-2	25
2006/07			23-10-06	1-2	16
2007/08	6.1	Naranja dulce	09-10-07	1-2	16
2008/09		Valencia late	03-10-08	1-2	17
2009/10			05-10-08	1-2	17
2006/07			06-09-06	1-2	42
2007/08	6.2	Naranja dulce	08-10-07	1-2	44
2008/09		Valencia late	03-10-08	1-2	28
2009/10			05-10-08	1-2	30
2006/07			08-09-06	1-2	45
2007/08	6.3	Naranja dulce	08-10-07	1-2	45
2008/09		Valencia late	29-09-08	1-2	31
2009/10			02-10-09	1-2	32
2006/07			13-09-06	1-2	51
2007/08	7.1	Naranja dulce	06-10-07	1-2	53
2008/09		Valencia late	26-09-08	1-2	37
2009/10			03-10-09	1-2	37
2013/14	8.1	Naranja dulce Salustiana	15-09-13	1-2*	38
2013/14	9.1	Naranja dulce Salustiana	15-09-13	1-2*	50

\* En Salustiana corresponde, estimación 1 febrero y 2 abril

Tabla 6.2. Huertos seleccionados para la instancia 2

Temporada	Huerto	Especie y variedad	Fecha plena flor	Momento estimación	n
2010/11	1.1	Tangor Murcott	26-09-10	1-2	30
2014/15	1.2	Tangor Murcott	30-09-14	1	18
2014/15	1.3	Tangor Murcott	01-10-14	1	18
2015/16	1.4	Tangor Murcott	30-09-15	1-2	48
2015/16	1.5	Tangor Murcott	30-09-15	1-2	42
2010/11			17-09-10	1-2	37
2011/12	2.1	Tangor Murcott	18-08-13	1-2	37
2013/14			30-08-12	1-2	37
2010/11			23-09-10	1-2	32
2013/14	3.1	Tangor Murcott	18-08-13	1-2	28
2014/15			13-10-14	1-2	28
2015/16	3.2	Tangor Murcott	26-09-15	1-2	46
2011/12	4.1	Tangor Murcott	23-09-11	1	26
2007/08			08-10-07	1-2	26
2008/09	5.1	Tangor Murcott	01-10-08	1-2	25
2009/10			02-10-08	1-2	24
2007/08			04-10-07	1-2	19
2008/09	6.1	Tangor Murcott	28-09-08	1-2	17
2009/10			30-09-09	1-2	17
2007/08			02-10-07	1-2	42
2008/09	7.1	Tangor Murcott	30-09-08	1-2	36
2009/10			30-09-09	1-2	41
2007/08			05-10-07	1-2	31
2008/09	8.1	Tangor Murcott	01-10-08	1-2	31
2008/09			30-09-09	1-2	31
2005/06	1.1	Naranja dulce Valencia late	03-10-05	1	17
2005/06	2.1	Naranja dulce Valencia late	08-10-05	1	25
2006/07			25-09-06	1-2	31
2007/08	3.1	Naranja dulce Valencia late	10-10-07	1-2	41
2008/09			08-10-08	1-2	47
2009/10			12-10-09	1-2	51
2006/07	4.1	Naranja dulce Valencia late	15-10-06	1-2	23
2007/08			09-10-07	1-2	31
2008/09	5.1	Naranja dulce Valencia late	04-10-08	1-2	25
2009/10			05-10-08	1-2	25
2006/07			23-10-06	1-2	16
2007/08	6.1	Naranja dulce Valencia late	09-10-07	1-2	16
2008/09			03-10-08	1-2	17
2009/10			05-10-08	1-2	17
2006/07			06-09-06	1-2	42
2007/08	6.2	Naranja dulce Valencia late	08-10-07	1-2	44
2008/09			03-10-08	1-2	28
2009/10			05-10-08	1-2	30
2006/07			08-09-06	1-2	45
2007/08	6.3	Naranja dulce Valencia late	08-10-07	1-2	45
2008/09			29-09-08	1-2	31
2009/10			02-10-09	1-2	32
2006/07			13-09-06	1-2	51
2007/08	7.1	Naranja dulce Valencia late	06-10-07	1-2	53
2008/09			26-09-08	1-2	37
2009/10			03-10-09	1-2	37
2013/14	8.1	Naranja dulce Salustiana	15-09-13	1-2*	38
2013/14	9.1	Naranja dulce Salustiana	15-09-13	1-2*	50

\* En Salustiana corresponde, estimación 1 febrero y 2 abril

#### *6.2.1.2. Variables*

##### **Valor de comparación:**

en el momento de la cosecha (agosto - octubre), en cada árbol seleccionado se determinó la cantidad de frutos cosechados o carga real (Carga a cosecha).

En una primera instancia se trabajó con las variables que se detallan a continuación.

##### **Características de los huertos:**

Se trabajó con información general de los huertos, referida al geoposicionamiento global, Latitud (LatD) y Longitud (LongD) de cada huerto establecidas mediante un GPS Garmin 66i, y a condiciones productivas informadas por el productor: Tipo de suelo (Suelo), Riego (Irr) y Edad de los árboles (Edad).

##### **Información morfológica de los árboles:**

De las diferentes mediciones efectuadas de las copas de los árboles, descritas en el Capítulo V, se seleccionó el diámetro menor de copa ( $d_m$ ).

##### **Número de frutos por árbol:**

Entre las diferentes estimaciones de la carga explicadas en el Capítulo V, se utilizó el recuento total (RT) a modo de comparación por haber sido el de mejor comportamiento y con el fin de probar opciones de bajo costo, se trabajó con las mediciones en secciones predefinidas de copa, empleando el número de frutos dentro del marco de muestreo en los cuatro puntos cardinales y su promedio (NumN - NumS - NumE - NumO y NumP) y ese número promedio extrapolado a volumen de copa mediante las diferentes fórmulas ya presentadas (VolM1, VolM2, VolM3, VolM4, VolM5 y VolM6). Si bien con el uso de estas mediciones en secciones predefinidas de copa, en el Capítulo V se habían obtenido errores mayores a las estimaciones mediante recuento total, se decidió volver a probarlas en combinación

con otras variables debido al menor costo operativo que suponen respecto del recuento total.

Para la aplicación de las técnicas de aprendizaje automático, debido al elevado costo que supone la obtención de datos sobre el campo, en la primera instancia se probaron diferentes combinaciones de variables de entrada a fin de encontrar un método de estimación que equilibrara el compromiso entre una precisión adecuada y el costo asociado a la obtención de los datos. Las variables estudiadas se agruparon en diferentes conjuntos para predecir la carga y entrenar los algoritmos, en primer lugar, se emplearon únicamente características de los huertos, donde solamente el diámetro *d* es una variable que requiere mediciones en los árboles.

Tabla 6.3. Descripción de las variables que integran cada conjunto de datos

	Conjunto	Variables
<i>c1</i>	<i>Básicas</i>	LatD, LongD, Suelo, Irr, Edad, Diámetro ( <i>d</i> )
<i>c2</i>	<i>Básicas + Recuento a volumen de copa</i>	LatD, LongD, Suelo, Irr, Edad, Diámetro ( <i>d</i> )+ VolM1, VolM2, VolM3, VolM4, VolM5, VolM6
<i>c3</i>	<i>Básicas + Recuento por cuadrante</i>	LatD, LongD, Suelo, Irr, Edad, Diámetro ( <i>d</i> ) + <i>NumN, NumS, NumE, NumO.</i>
<i>c4</i>	<i>Completa</i>	LatD, LongD, Suelo, Irr, Edad, Diámetro ( <i>d</i> ) + VolM1, VolM2, VolM3, VolM4, VolM5, VolM6 +NumN, NumS, NumE, NumO.

En la segunda instancia, además de las variables anteriores se utilizó la información que se explica a continuación.

### **Variables climáticas**

En estaciones meteorológicas situadas próximas a los huertos (entre 5 y 45 km de distancia) se obtuvo en cada temporada, precipitación total, temperatura media y humedad, en plena floración (septiembre) (PrecipF, TempF, HumF), durante los periodos de crecimiento de los frutos (diciembre a marzo) (PrecipCF, TempCF, HumCF) y de maduración (abril a junio) (PrecipMAD, TempMAD, HumMAD).

## Información satelital

La obtención de la información mediante teledetección por satélite requiere una sólida adquisición de datos de alta resolución espacial y temporal para resolver la rápida transición fenológica a escala de campo. La creciente disponibilidad de datos satelitales de resolución moderada y de libre acceso ofrece una oportunidad sin precedentes para la obtención de información de tipos de cultivos en áreas extensas.

Los sensores disponibles en el momento del estudio eran Landsat y el espectrorradiómetro de imágenes de resolución moderada (MODIS), actualmente existen otros sensores como Sentinel 2, pero estuvo disponible recién a partir del año 2014.

Para el presente estudio se decidió utilizar el producto MOD09QG<sup>20</sup> del sensor MODIS, debido a que ofrece una cobertura temporal más consistente, con observaciones cada 15 días, en función de los objetivos de la investigación y el período de estudio que se extiende, para cada temporada, entre los meses de agosto y septiembre y de diciembre y junio. Este sensor, si bien presenta una resolución espacial menor que otros satélites (500 metros), ofrece mayor resolución temporal, con series de tiempo más largas, 2 observaciones por mes y una banda de máscara de nubes que permite bloquear los píxeles con altos porcentajes de nubosidad, lo cual disminuye el sesgo en la estimación del NDVI. Otros sensores como los Landsat 5, 7 y 8, ofrecen una resolución espacial mayor (30 metros de tamaño del píxel) pero su resolución temporal es menor, con un período de revisita de alrededor de 18 días, series de tiempo más cortas o con interrupciones y no poseen banda de máscara de nubes (Xiao, *et al.* 2021).

Por lo expuesto se decidió sacrificar resolución espacial en beneficio de la temporal, y de ponderar la consistencia de la serie de tiempo de los datos, ya que con MODIS se utiliza un sólo sensor para cubrir el período el análisis, mientras que con LANDSAT habría que combinar imágenes de los 3 sensores. A lo que se agrega el hecho que el sensor Landsat 7 sufrió una avería en el año 2003, por lo cual sería necesario realizar un proceso de corrección a los conjuntos o colecciones de imágenes.

---

<sup>20</sup> Disponible en: [https://modis-land.gsfc.nasa.gov/pdf/MOD09\\_UserGuide\\_v1.4.pdf](https://modis-land.gsfc.nasa.gov/pdf/MOD09_UserGuide_v1.4.pdf)



El sensor MODIS se encuentra a bordo de los satélites Terra (EOS-AM) y Aqua (EOS-PM), y fue utilizado para obtener la reflectancia 1 (infrarrojo cercano, nir) y la reflectancia 2 (rojo, red) en plena floración (septiembre: Ref10, Ref20, NDVI0), crecimiento y maduración de frutos (diciembre a junio: Ref11 a Ref17, Ref21 a Ref27 y NDVI1 a NDVI7). Se tomaron dos registros mensuales y se promediaron para definir un valor para cada mes. A partir de esos valores de reflectancia se calculó el NDVI mediante la siguiente expresión:

$$NDVI = \frac{(REF_{nir} - REF_{red})}{(REF_{nir} + REF_{red})}$$

El procedimiento de extracción de los valores reflectancia, asociados a los huertos y la conformación de la base de datos se automatizó utilizando el flujo de trabajo de cuatro fases detallado en la Figura 6.8.

Fase 1: localización del huerto y cálculo del centroide.

Fase 2: descarga de las series temporales de los sensores MODIS (producto MOD09GQ<sup>21</sup>) mediante una rutina del programa R Statistics, desarrollado por Arango (2016), que se describe más adelante.

Fase 3: estimación del NDVI según estación, momento y huerto.

Fase 4: creación de la base de datos.

---

<sup>21</sup> <https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/MOD09GQ/>

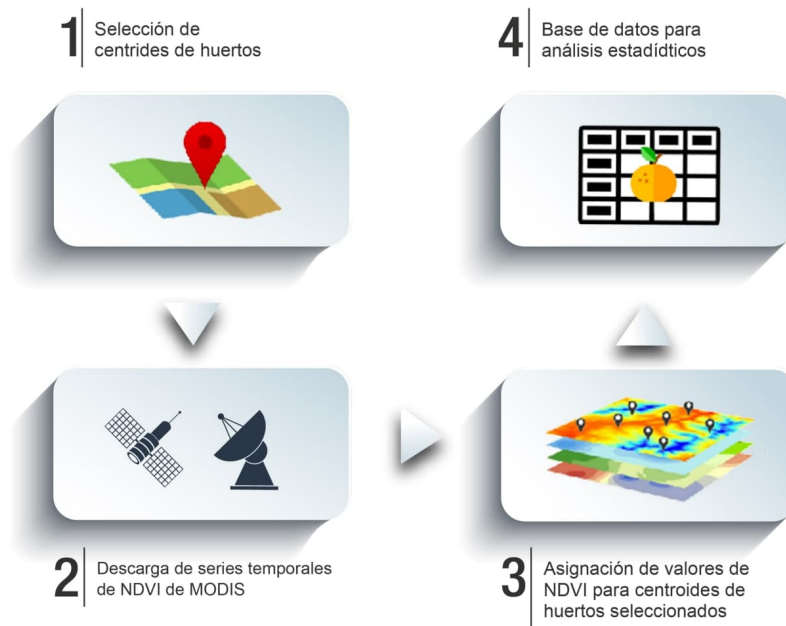


Figura 6.8. Flujo de trabajo utilizado en el proceso

La metodología utilizada para la obtención de los valores de reflectancia y posterior estimación del índice normalizado de vegetación sobre los huertos analizados está compuesta por dos rutinas en el lenguaje de programación R<sup>22</sup>, que incluyen la rutina principal *DescargaModis.r* y la función *getMOD09GQ.r*, cuya tarea es conectarse a los servidores de la plataforma MODIS y realizar la búsqueda y descarga de los datos, basado en los argumentos de entrada en la rutina principal. Para este análisis se utilizó la información del producto MOD09GQ<sup>23</sup> del sensor MODIS que está compuesto por escenas globales diarias con resolución espacial de 250 metros y resolución espectral de dos bandas, la banda 1 corresponde a la porción del rojo y la banda 2 al infrarrojo cercano (Mas, 2011).

El producto MOD09QG<sup>24</sup> captura la información devuelta desde las diferentes cubiertas de la tierra en términos de reflectancia de superficie. La reflectancia de la superficie es la cantidad de luz reflejada por la superficie de la tierra, se mide como una relación entre la radiancia de la superficie y la irradiancia de la superficie, no tiene unidades y toma valores entre 0,0 y 1,0. Los valores de reflectancia de la superficie de MOD09 se escalan por 10000 y luego se convierten en enteros de 16 bits. En términos

<sup>22</sup> Disponible en: <https://www.r-project.org/>

<sup>23</sup> Disponible en: <https://lpdaac.usgs.gov/products/mod09gqv006/>

<sup>24</sup> Disponible en: [https://modis-land.gsfc.nasa.gov/pdf/MOD09\\_UserGuide\\_v1.4.pdf](https://modis-land.gsfc.nasa.gov/pdf/MOD09_UserGuide_v1.4.pdf)

prácticos esto quiere decir que un píxel de este producto de MODIS puede tomar valores que van del 0 al 65.536.

Para la ejecución de la herramienta se requiere ingresar los siguientes argumentos de entrada:

- ✓ Caja del área de estudio:
- ✓ Valores de las coordenadas que definen la extensión de cada huerto.
- ✓ Fechas para la descarga de las imágenes: fechas de inicio y final para las que se requiere hacer la consulta y extracción de los valores.
- ✓ Definición del sistema de referencia de coordenadas (SRC): se especifica el sistema de referencia de coordenadas de los huertos, esto sirve para que la herramienta sea capaz de superponer la geometría de los huertos sobre las imágenes de MODIS.

Los datos de salida producto de la ejecución de la herramienta son:

- ✓ Coordenadas XY del centroide que corresponde al huerto.
- ✓ Fecha de la imagen utilizada.
- ✓ Valor de reflectancia de la banda 1 (refl\_b01).
- ✓ Valor de reflectancia de la banda 2 (refl\_b02).
- ✓ Valor de calidad de píxel (QC\_250m), que se expresa en porcentaje y mide la presencia de nubes y otros aerosoles presentes en la atmósfera en el momento de la toma de la imagen.

Estos resultados quedan almacenados dentro un marco de datos (*dataframe*) en el espacio de trabajo de R, el que luego es usado para la estimación del NDVI y posteriores análisis.

La Figura 6.9 representa el proceso, en el que las capas t son los mapas que contienen los valores resumen mensuales del NDVI para cada momento de medición (0 a 7).

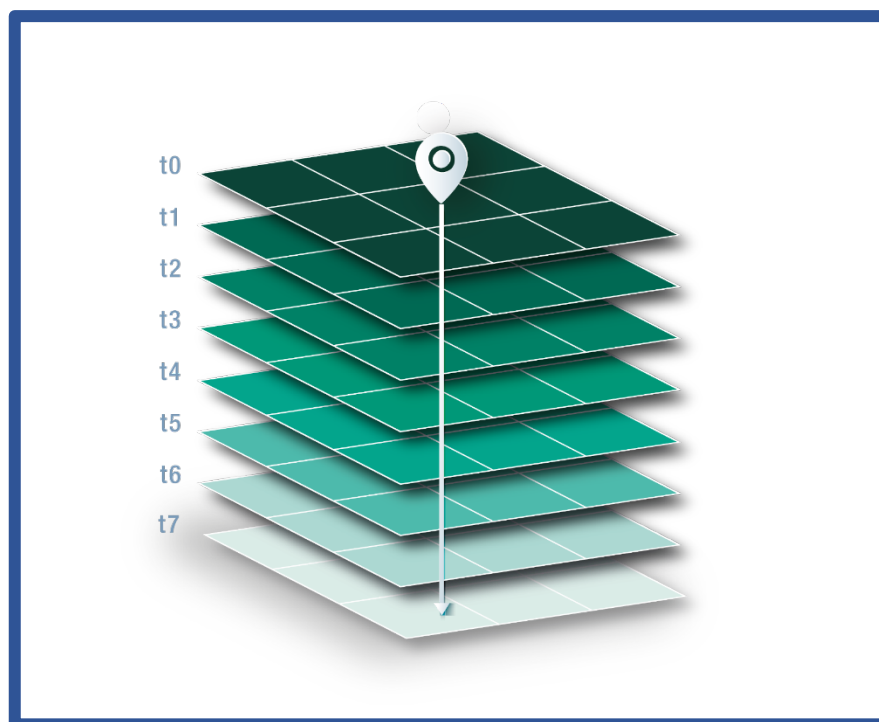


Figura 6.9. Extracción de los valores de NDVI en cada huerto

El resultado de esta extracción consistió en una matriz tabular de filas y columnas que contiene el identificador de cada huerto, campo y datos del satélite.

El costo de la obtención de datos depende de muchos factores, los más elevados corresponden a las mediciones *in situ* sobre los árboles porque requieren un arduo trabajo directamente dentro de los huertos. Las características de los huertos también son costosas si es necesario visitarlos, aunque a veces se dispone de una base de datos que evitan la visita. Las variables climáticas pueden ser obtenidas en las estaciones meteorológicas situadas en la zona. Los datos satelitales están disponibles gratuitamente desde diversos satélites. Teniendo en cuenta los costos y las dificultades en la medición de las variables, se crearon tres conjuntos de datos diferentes con el fin de estudiar el rendimiento de la predicción en función del costo de la recogida de información para producir dicha predicción. La composición de los diferentes conjuntos de datos considerados se detalla en la Tabla 6.4.

Tabla 6.4. Descripción de las variables que integran cada conjunto de datos

Conjunto		Variables
d1	Satélite	Ref11 a Ref17, Ref21 a Ref27, NDVI0 a NDVI7 y Carga a cosecha
d2	Huerto, climáticas y satélite	Sp, Temporada, Edad, Suelo, Irr, TempF, TempCF, TempMAD, PrecipF, PrecipCF, PrecipMAD, HumF, HumCF, HumMAD, Ref11 a Ref17, Ref21 a Ref27, NDVI0 a NDVI7 y Carga a cosecha.
d3	Características de árboles, huertos, climáticas y satélite	LD, Altura, NumN, NumS, NumE, NumO, NumP, Sp, Temporada, Edad, Suelo, Riego, TempF, TempCF, TempMAD, PrecipF, PrecipCF, PrecipMAD, HumF, HumCF, HumMAD, Ref11 to Ref17, Ref21 to Ref27, NDVI0 y Carga a cosecha.

En la Tabla 6.5 se presentan las frecuencias de los valores categóricos, considerados en el análisis.

Tabla 6.5. Frecuencias de las variables categóricas (Vd, variedad; Suelo, arenoso-arcilloso; Irr, presencia o ausencia de riego)

Variable	Valor	Frecuencias (%)
Vd	Murcott	44
	Salustiana	4
	Valencia late	52
Suelo	Arenoso	78
	Arcilloso	22
Irr	Si	40
	No	60

De los huertos incluido en el análisis el 44% correspondió a tangor Murcott y el 56% a naranjo dulce (52% Valencia late y 4% Salustiana), el 78% se encuentra implantado sobre suelos arenosos y el 22% en suelos arcillosos, el 40% posee riego y el 60% no lo posee.

### 6.2.2. Análisis de datos

Como ya se ha dicho, en este Capítulo se trata la estimación de la carga por árbol, mediante la aplicación de diferentes métodos de Aprendizaje Automático, combinados con variables de distinto tipo (datos de campo, información meteorológica y

proveniente de sensores remotos) con la finalidad de encontrar modelos que permitan disminuir los errores de estimación.

Para estimar el número de frutos por árbol, cada conjunto de datos se dividió en un subconjunto de entrenamiento y otro de prueba (75% y 25% de los datos, respectivamente), con el fin de evitar sesgos este proceso se repitió 50 veces. El entrenamiento se realizó siguiendo el modelo de validación cruzada con 10 cajas

La precisión de los métodos se estudió en términos del error cuadrático medio (RMSE) y el error medio absoluto (MAE) y se calculó el porcentaje de error en ambas instancias.

Error cuadrático medio, (RMSE), mide el promedio de la diferencia cuadrática entre los valores observados y estimados, es el indicador más utilizado para la validación de modelos de sistemas físicos y se puede calcular mediante la siguiente expresión:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{u}_i)^2}$$

El error absoluto medio (MAE), mide la magnitud media de la diferencia entre los datos medidos y los estimados por el modelo. Fue presentado por Hyndman & Koehler, (2006) y utilizado por (Gaiser *et al.*,2010) y (Karamirad *et al.*,2013). Su valor mínimo es cero y se produce cuando los datos medidos y estimados son iguales en todo el rango de la muestra, se puede calcular con la siguiente expresión

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{u}_i|$$

Para evaluar la precisión de cada método de estimación, se calcularon los errores porcentuales (% Error) en valor absoluto entre las cargas estimadas y las reales de la siguiente manera:

$$\% Error = \frac{\sum_i |\hat{\mu}_i - y_i|}{\sum_i y_i} \cdot 100$$

Donde  $\hat{\mu}_i$  es la predicción de carga para el *i-ésimo* árbol e  $y_i$  es el valor de comparación, esto es la carga real del *i-ésimo* árbol.

### 6.2.2.1. Estimación de carga basada en datos de campo (características de los huertos e información de los árboles)

La predicción de la producción se realizó mediante árboles de regresión, que han demostrado ser métodos adecuados para la predicción del rendimiento de los cultivos. Los algoritmos más utilizados para construir árboles de regresión son CART, M5 y M5-Prime.

Estos algoritmos utilizan una estrategia de "divide y vencerás" para construir el modelo. El primer nodo del árbol se llama raíz, un nodo con aristas salientes se llama nodo de prueba y un nodo sin aristas salientes es una hoja. Cada camino desde la raíz del árbol hasta una hoja determina una región, es decir, un subconjunto más homogéneo de los datos de entrada. En el caso de los datos analizados en este Capítulo, los huertos se agrupan según una producción estimada similar. Inicialmente, todo el conjunto de entrenamiento se asocia a una hoja. Aplicando un procedimiento recursivo, se utiliza una prueba para decidir si el conjunto asociado a una hoja se divide en subconjuntos más pequeños asociados a nuevas hojas.

Las principales diferencias entre los métodos descritos son los criterios de división, las reglas de poda y el mecanismo para estimar el valor de la hoja, CART utiliza la varianza como criterio de división, M5 utiliza la reducción de la desviación estándar (SDR). Por lo tanto, el valor estimado de una hoja es constante en CART, mientras que M5 aproxima los valores de las hojas mediante modelos de regresión lineal y es capaz de mejorar las predicciones introduciendo un procedimiento de suavizado.

Los experimentos se realizaron con el paquete RWeka (Hornik *et al.* 2009), utilizando la función *M5-Prime* con la configuración estándar, es decir, con poda, suavizado y con un número mínimo de 4 ejemplos por nodo.

### 6.2.2.2. Estimación de carga basada en características de los huertos, información de los árboles, variables climáticas e información satelital

Los algoritmos probados fueron RT (M5, Ctree, CART con dos implementaciones diferentes, rpart y rpart2), SVM con núcleos polinómico (svm1) o radial (svm2), ANN con perceptrón de una capa (mlp) o dos capas (mlpMP), KNN y RF. Los parámetros optimizados durante el proceso de validación cruzada, así como los valores de ajuste, se detallan en la Tabla 6.6.

Tabla 6.6. Parámetros introducidos en cada método de Aprendizaje Automático empleado (RT, árboles de regresión; ANN, redes neuronales artificiales; KNN, k-vecinos más cercanos; SVM, máquinas de soporte vectoriales; RF, bosques aleatorios)

Métodos	Parámetros	
RT	M5	pruned = {sí, no}, smoothed = {sí, no}, rules = {sí, no}
	ctree	seq (5, 20, 5), mincriterion = 0.75, 0.85, 0.95
	rpart	cp = seq (0, 5, 0.25)
	rpart2	maxdepth = seq (2, 15, 1)
ANN	mlp	size = seq (2, 20, 2)
	mlp7	layer1 = seq (2, 10, 2), layer2 = seq (2, 6, 2), layer3= 0
KNN	k	k = (2, 8, 2)
SVM	svm1	degree = 2, 3, scale = 0.1, 0.5, 1
	svm2	sigma = c 1, 3
RF	rf	mtry=-3, 3

El significado de los parámetros introducidos es el siguiente:

*Pruned* es *true* si el árbol M5 se poda, *false* en caso contrario. *Smoothed* es *true* si el árbol M5 está suavizado, *false* en caso contrario. *Rules* significa que el modelo producido por M5 puede ser un árbol o un sistema de reglas. *Maxdepth* representa la profundidad máxima del árbol producido para ctree y rpart2. *Cp* es el parámetro de complejidad en Rpart. Tanto rpart como rpart2 se refieren al mismo método, CART.

*Size*, *layer1* y *layer2* se refieren a la configuración de las diferentes capas en las redes neuronales consideradas.

El parámetro por optimizar en caso de KNN es el número de vecinos *k*.



Los parámetros *degree*, *scale* y *sigma* se refieren a las configuraciones de los kernels en SVM.

Finalmente, *mtry* se refiere al número de características consideradas para construir los árboles en el bosque aleatorio.

Con el fin de contrastar las medias de RMSE y MAE entre los diferentes conjuntos de datos y métodos estudiados, se aplicaron pruebas t pareadas con  $\alpha = 0,05$ , que permiten determinar si, para un mismo huerto, existe una variación significativa entre el número real de frutos cosechados y el estimado, a través del estadístico muestral t.

Se construyeron gráficos de caja con los valores de RMSE y MAE para los diferentes conjuntos de datos y métodos estudiados, los que permitieron obtener una visión general del comportamiento de estos valores y definir la combinación más adecuada.

### **Análisis de las características relevantes**

Cuando se quiere detectar la mejor combinación del conjunto de datos para predecir la carga mediante SVM se enfrenta al problema su escasa o nula interpretabilidad (diversos autores hacen referencia a que son “una caja negra”), por lo que no es posible analizar la importancia relativa de las características que intervienen en la construcción del modelo. Aunque existen algunos enfoques para seleccionar o identificar las características más prometedoras asociadas al modelo producido, no hay una forma estándar de evaluar la importancia de las variables asociadas a un modelo de clasificación obtenido por SVM. Sin embargo, es interesante en este marco estudiar de alguna manera las variables más relevantes.

El enfoque seguido en esta tesis es el explicado en la sección 6.4.2, utilizando como clasificador el método SVM (porque es el que mejor clasifica), utilizando el método de Borda como método de agregación, y dividiendo aleatoriamente en conjunto de entrenamiento y de prueba 50 veces.

Los análisis se realizaron utilizando los paquetes de *R versión 4.1.0*, *caret 6.0.86* y *ggplot2 3.3.2*.

## 6.3. RESULTADOS Y DISCUSIÓN

A continuación, se presentan y discuten los resultados obtenidos en las dos instancias realizadas.

### 6.3.1. Estimación de carga basada en datos de campo (características de los huertos e información de los árboles)

En primer lugar, se estudió el comportamiento de los métodos de recuento previamente definidos y posteriormente se aplicó el enfoque basado en árboles de regresión y selección de características para realizar dicha predicción.

Se calcularon los coeficientes de correlación entre la carga estimada por los diferentes métodos de recuento y la carga real de los huertos. Los coeficientes  $r$  resultaron estadísticamente significativos, sin embargo, sus valores indican asociaciones débiles entre la carga estimada y la real (los  $r$  varían entre 0,10 y 0,61), como se puede ver en la Tabla 6.7.

Tabla 6.7. Coeficientes de correlación de Pearson ( $r$ ) entre diferentes variables y la carga a cosecha, y valores  $p$  asociados, en naranjo dulce y tangor Murcott

Variables	Naranja dulce		Tangor Murcott	
	$r$	$p$ -valor	$r$	$p$ -valor
VolM1	0,33	< 0,001	0,25	< 0,001
VolM2	0,51	< 0,001	0,41	< 0,001
VolM3	0,61	< 0,001	0,36	< 0,001
VolM4	0,07	< 0,01	0,17	< 0,001
VolM5	0,10	< 0,001	0,17	< 0,001
VolM6	0,19	< 0,001	0,19	< 0,001

A continuación, se presentan los resultados obtenidos con diferentes combinaciones de variables, pero al analizarlos es necesario tener en cuenta las dificultades para obtener los distintos tipos de información, la información referida a los huertos se

encuentra disponible en los establecimientos, los datos descriptivos de las copas de los árboles se determinan una vez al año, la obtención de los recuentos totales de frutos o en marcos de muestreo incrementa el trabajo de campo necesario y consecuentemente el costo, y es importante evaluar las ventajas de obtener una mayor precisión a costa de un aumento de los gastos.

La Figura 6.9 muestra los errores de estimación en naranjo dulce, cuando se predijo la carga mediante M5-Prime, utilizando diferentes conjuntos de variables. Las líneas representan los errores obtenidos cuando la carga se estima a partir del recuento total de frutos (RT) a modo de comparación y utilizando los conjuntos de datos c1, c2, c3 y c4. Por otro lado, cada barra representa la predicción cuando se considera el conjunto c1 combinado con la carga estimada por el marco promedio combinado con cada una de las fórmulas de volumen de copa (VolM1 a VolM6).

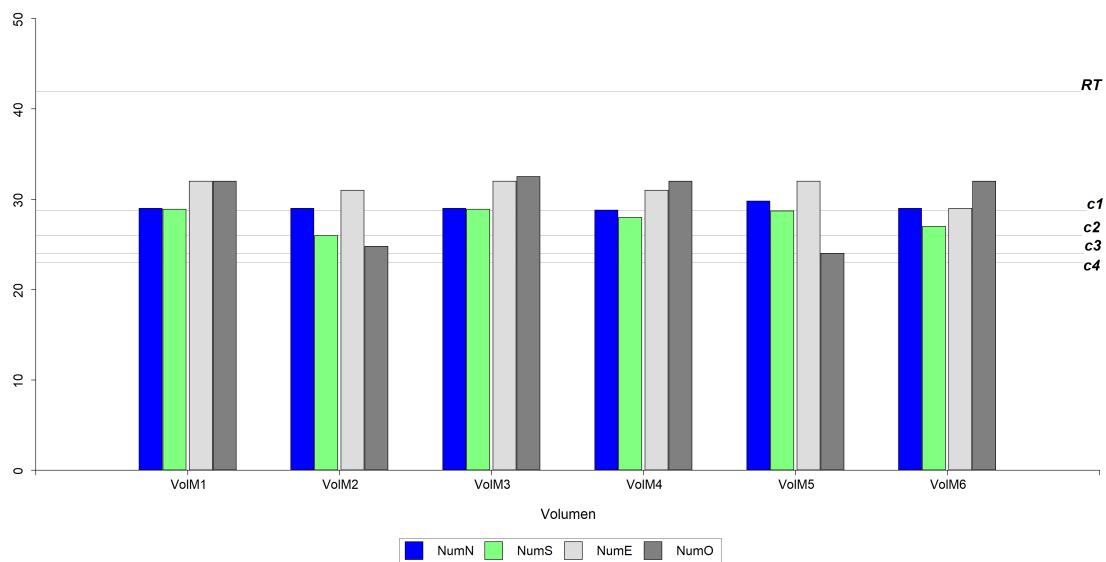


Figura 6.10. Errores obtenidos utilizando los diferentes conjuntos de variables para naranjas

Las mejores estimaciones se lograron a partir del conjunto *Completo*, luego las del conjunto *Básicas + recuento por cuadrante*, el conjunto *Básicas + recuento a volumen* y finalmente el conjunto *Básicas*. Como se puede observar, los errores obtenidos cuando se considera una sola combinación de volumen y marco contador fueron similares a los obtenidos al considerar todo el conjunto de datos c4, no obstante, la información del último conjunto es mucho más costosa. La información proporcionada por la aplicación

del marco contador ubicado en el cuadrante norte (NumO) combinada con la estimación del VolM2 (cono truncado) produjo los más bajos errores.

Los errores para tangor Murcott (Figura 6.11), tuvieron un comportamiento diferente. Cuando sólo se consideraron los conjuntos c1 y c2 como entrada del procedimiento de aprendizaje, el rendimiento fue más bajo que con las demás combinaciones o el recuento total. Se observaron varias combinaciones de posiciones del marco contador y fórmulas de volumen de copa con errores bajos, aunque la combinación de marco contador en el cuadrante oeste (NumO) y estimaciones de VolM6 (cilindro) fue la de mejor comportamiento.

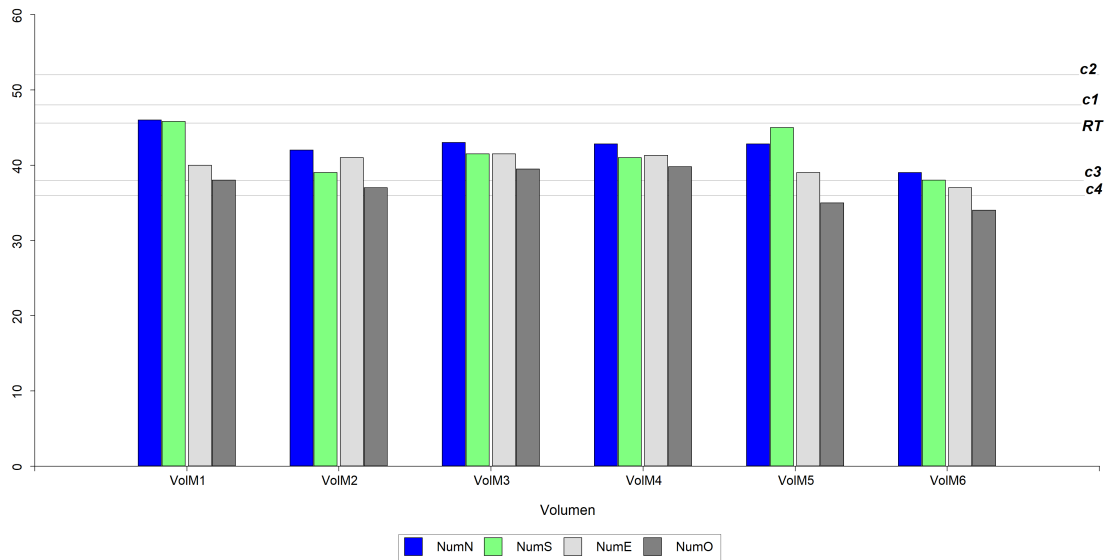


Figura 6.11. Errores obtenidos utilizando los diferentes conjuntos de variables para tangor Murcott

Las Tablas 6.8 y 6.9, presentan los resultados obtenidos con M5-Prime para las mejores combinaciones de variables en naranja dulce y tangor Murcott respectivamente.

En ambas tablas se presenta el conjunto de datos empleado, la complejidad del modelo a través del número de hojas y de la cantidad de variables utilizadas, el error de predicción del y el coeficiente de correlación asociado a cada combinación de variables.

Tabla. 6.8. Características de cada método empleado para predicción de carga mediante M5-PRIME en naranjo dulce (conjuntos de datos utilizado, complejidad de los modelos, errores en las predicciones y correlación entre carga estimada y real)

Conjunto de datos	Complejidad del modelo			Correlación
	Número de hojas	Número de variables	Error de predicción	r
C1	10	8	29,90	0,83
C2	14	11	57,80	0,84
C3	1	11	27,80	0,84
C5	10	17	27,80	0,91
VolM2-NumO	25	11	24,30	0,86

Tabla. 6.9. Características de cada método empleado para predicción de carga mediante M5-PRIME en tangor Murcott (conjuntos de datos utilizado, complejidad de los modelos, errores en las predicciones y correlación entre carga estimada y real)

Conjunto de datos	Complejidad del modelo			Correlación
	Número de hojas	Número de variables	Error de predicción	r
C1	7	4	29,90	0,64
C2	12	10	27,90	0,73
C3	5	10	27,80	0,81
C5	1	12	27,80	0,97
VolM6-NumO	7	10	24,30	0,84

La correlación entre los distintos métodos de estimación del número de frutos y la carga real en todos los casos fue significativa. Los coeficientes de correlación con el recuento total fueron similares a los valores encontrados anteriormente por Avanza (2011).

Estos resultados para la estimación utilizando marcos de muestreo en naranjo dulce y tangor Murcott son inferiores a los valores encontrados en naranjo dulce por Rezende & Barbosa (2004).

Los resultados de las predicciones de carga de naranjo dulce y tangor Murcott fueron ligeramente diferentes. En el caso de naranjo dulce la precisión de la predicción aumentó con el incremento de la información añadida, alcanzando el error más bajo (21,3%) cuando se consideraron las variables *Básica*, *Recuentos por cuadrante* (llevados o no a volumen de copa), sin embargo, ésta es la opción más costosa, por lo que se estudiaron las diferentes combinaciones  $VolM_i + Num_j$  y se encontró que la mejor combinación en términos de precisión fue la de las variables *Básicas* más VolM2 (cono

truncado) y NumO para naranjo dulce. Esto significa que solamente con la introducción de alguna información básica relacionada con los árboles, como la edad, la estimación mejoró.

El comportamiento de la predicción de la carga en tangor Murcott fue similar al de naranjo dulce, esto es, cuando se añadió más información al método el error disminuyó, aunque eso implica mayores costos. No obstante, se han podido obtener errores aceptables cuando se combinó el recuento en marco contador ubicado al oeste (NumO) combinado con el volumen de copa calculado mediante VolM6 (cilindro).

Se sabe que no existe un consenso en las ecuaciones utilizadas para el cálculo del volumen de copa en diferentes especies y variedades. No obstante, en este trabajo las fórmulas que mostraron mejor comportamiento no se condicen con las encontradas por otros autores para cítricos, como es el caso Orduz Rodríguez *et al.* (2007) en naranja Valencia y Monteverde *et al.* (1996) en mandarina Arrayana, quienes consideraron la forma de la copa como una esfera, otros autores como Zekri (2000) y Georgiou (2000) que utilizaron una semiesfera alargada en mandarinas Nova y Clementina. En nuestro caso, las formas de copa con mejor comportamiento corresponden al volumen de un cono truncado en el caso de las naranjas y a un cilindro en el caso de tangor Murcott, por Avilán & Mazzi (1973) y Avilán (1979), quienes consideraron como copa útil del árbol la contenida en un cono truncado, al determinar que la producción que está ubicada en el ápice de la planta es despreciable y que cerca del 98% de la producción se ubica en la parte media de la copa del árbol.

### 6.3.2. Estimación de carga basada en características de los huertos, información de los árboles, variables climáticas e información satelital

En esta sección se analiza el rendimiento de los diferentes enfoques de aprendizaje automático en función del conjunto de datos utilizado como entrada, detallados previamente y se incorpora información de otras fuentes como variables climáticas y datos satelitales.

La Figura 6.12, muestra los MAE y RMSE por conjunto de datos y método de aprendizaje automático. Cada figura muestra un gráfico de caja con la medida de evaluación correspondiente a través de las 50 repeticiones.

Los diferentes métodos tienen un comportamiento similar entre los diferentes conjuntos de datos. Los valores más altos de MAE y RMSE y las variaciones más altas se encontraron con el algoritmo MLP (mlp), que a diferencia de Han *et al.* (2019) que estiman la biomasa de maíz, en este trabajo no tuvo un buen comportamiento. KNN (knn), RT (M5, ctree, rpart, rpart2), ANN (mlp-ML) y SVM con kernel radial (svm2) tampoco mostraron buen comportamiento. Se obtuvieron menores MAE y RMSE con RF y SVM con kernel polinómico, los que tuvieron el mejor comportamiento en todo el conjunto de datos, lo que coincide con los resultados obtenidos por Kurtulmus *et al.* (2013), Feng *et al.* (2020), Han *et al.* (2019), Leroux *et al.* (2019), Maya Gopal & Bhargavi (2019) y Nyalala *et al.* (2019).

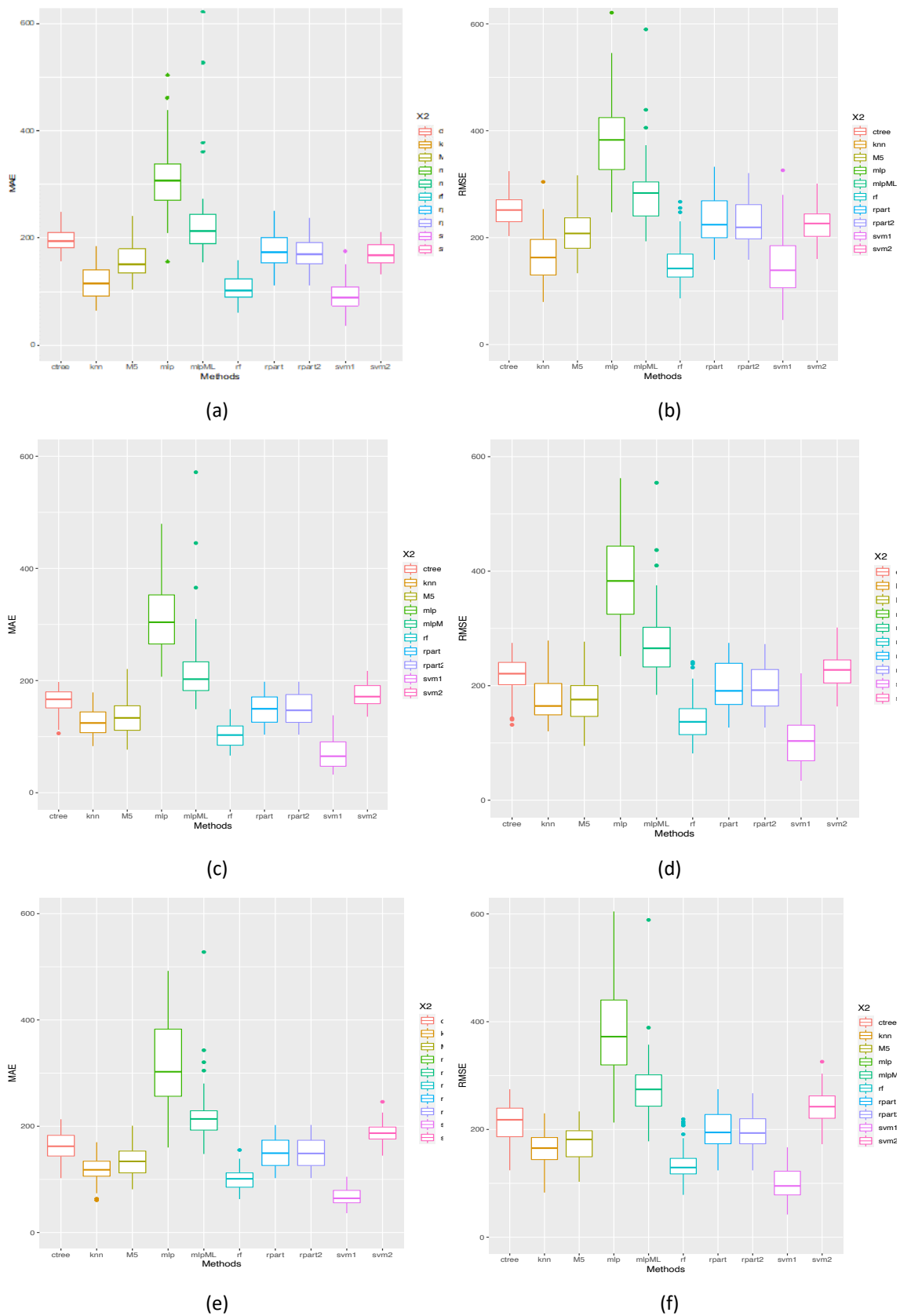


Figura 6.12. Gráficos de caja que representan: **a**, MAE conjunto d1; **b**, RMSE conjunto d1; **c**, MAE conjunto d2; **d**, RMSE conjunto d2; **e**, MAE conjunto d3; **f**, RMSE conjunto d3



Como se puede ver, los modelos que presentaron menores valores de MAE y RMSE fueron los bosques aleatorios (rf) y las máquinas de vectores de soporte con kernel polinómico (svm1), siendo este último el de mejor comportamiento independientemente del conjunto de datos (aunque en el caso de utilizar d1, los RMSEs obtenidos por rf y svm1 son similares).

Se realizaron pruebas t pareadas de una cola, para comparar las diferencias medias de MAE y RMSE entre svm1 y los demás métodos. En la Tabla 6.10 se presentan los valores p de cada prueba, que confirman que los errores medidos en términos de MAE y RMSE son significativamente menores cuando se estima la carga de cítricos utilizando svm1 (sólo hay una excepción cuando se compara el RMSE obtenido por svm1 y rf utilizando el conjunto de datos d1, y en ese caso no hay diferencias).

Así, las máquinas de vectores soporte con kernel polinómico (svm1) obtienen el mejor rendimiento en términos de MAE y RMSE independientemente del conjunto de datos de entrada.

Tabla 6.10. Valores p de la prueba t pareada para las diferencias medias de MAE y RMSE entre svm1 y otros algoritmos para cada conjunto de datos.

Métodos			d1	d2	d3
Árboles de regresión (RT)	Ctree	MAE	2,92e <sup>-24</sup>	1,37e <sup>-25</sup>	4,94e <sup>-23</sup>
		RMSE	1,50e <sup>-12</sup>	7,28e <sup>-20</sup>	5,04e <sup>-20</sup>
	rpart	MAE	4,91e <sup>-18</sup>	5,77e <sup>-22</sup>	1,11e <sup>-20</sup>
		RMSE	1,65e <sup>-08</sup>	3,69e <sup>-15</sup>	6,97e <sup>-19</sup>
	rpart2	MAE	7,94e <sup>-18</sup>	2,95e <sup>-20</sup>	6,09e <sup>-22</sup>
		RMSE	4,53e <sup>-08</sup>	4,31e <sup>-14</sup>	1,03e <sup>-20</sup>
	M5	MAE	4,95e <sup>-12</sup>	1,46e <sup>-13</sup>	1,10e <sup>-15</sup>
		RMSE	1,05e <sup>-03</sup>	3,38e <sup>-08</sup>	2,36e <sup>-14</sup>
Redes Neuronales	MP	MAE	4,12e <sup>-19</sup>	7,96e <sup>-15</sup>	4,87e <sup>-03</sup>
		RMSE	1,80e <sup>-16</sup>	9,46e <sup>-15</sup>	4,19e <sup>-02</sup>
	mlpML	MAE	1,54e <sup>-10</sup>	3,47e <sup>-16</sup>	8,51e <sup>-18</sup>
		RMSE	1,51e <sup>-09</sup>	3,70e <sup>-16</sup>	6,92e <sup>-19</sup>
Métodos Vagos (knn)	Knn	MAE	9,51e <sup>-02</sup>	1,38e <sup>-13</sup>	1,13e <sup>-17</sup>
		RMSE	1,23e <sup>04</sup>	1,27e <sup>07</sup>	9,16e <sup>-18</sup>
Máquinas de vectores soporte (SVM)	svm2	MAE	1,29e <sup>19</sup>	7,54e <sup>-30</sup>	7,36e <sup>-35</sup>
		RMSE	4,30e <sup>-07</sup>	2,42e <sup>-20</sup>	4,64e <sup>-28</sup>
Bosques Aleatorios (RF)	rf	MAE	1,17e <sup>02</sup>	1,48e <sup>-17</sup>	4,57e <sup>-18</sup>
		RMSE	4,86e <sup>05</sup>	2,15e <sup>-10</sup>	4,25e <sup>-13</sup>

Al analizar los resultados obtenidos con los diferentes conjuntos de datos se observa que el peor comportamiento se obtiene cuando se considera el conjunto de datos d1. Por otra parte, hay pequeñas diferencias entre los resultados obtenidos con los conjuntos de datos d2 y d3. Las pruebas t pareadas para comprobar si las diferencias entre ellos son significativas indican que, cuando se comparan los resultados entre d1 y d2 los valores p obtenidos son  $1,757206^{-07}$  (MAE) y  $1,007665^{-06}$  (RMSE), por lo que es mejor considerar el conjunto de datos d2. Por otro lado, cuando se comparan d2 y d3, los valores p son  $8,356207^{-01}$  (MAE) y  $1,339823^{-01}$  (RMSE), demostrando que no hay diferencias significativas entre ellos.

Teniendo en cuenta las variables incluidas en cada conjunto de datos, y la dificultad y los costos de medir diferentes tipos de variables, especialmente las que corresponden a los árboles, se enfocó el comportamiento de smv1 en los diferentes conjuntos de datos. En la Figura 6.13, se presentan los gráficos de caja de MAE y RMSE obtenidos con el algoritmo svm1 en los 3 conjuntos de datos (d1, d2, d3).

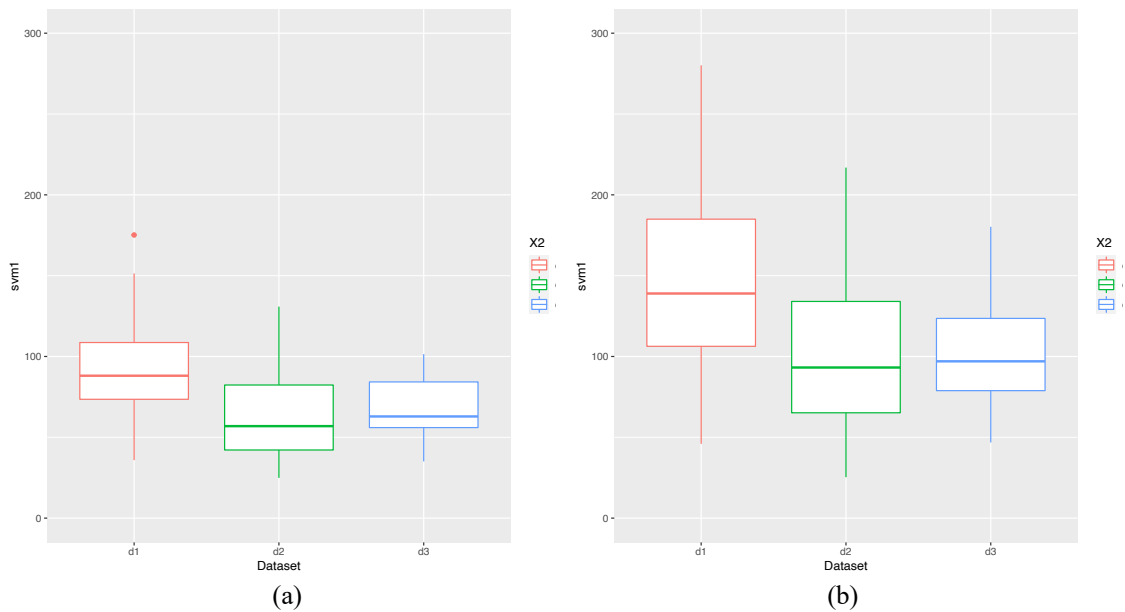


Figura 6.13. Gráficos de caja que representan: **a**, valores MAE y **b**, RMSE obtenidos con svm1 para diferentes conjuntos de datos (d1, d2, d3).

Cuando se analizan los valores MAE el comportamiento de smv1 es mejor en el conjunto de datos d2 y prácticamente igual si se considera el RMSE. Haciendo un contraste entre los valores de MAE y RMSE comparando d2 con el otro conjunto de datos, no hubo diferencias ni entre d2 y d3 ni entre d1 y d2, los valores p para d1 vs d2

fueron  $1757206e^{-07}$  (MAE) y  $1,007665e^{-06}$  (RMSE), y entre d2 vs. d3 fueron  $8,356207e^{-01}$  (MAE) y  $1,339823e^{-01}$  (RMSE), lo que indica el mejor comportamiento con d2, el conjunto de datos que incluye datos de huertos, climáticos y de satélite.

Estos resultados concuerdan con (Begué *et al.*, 2018) en una revisión de la literatura sobre teledetección para el mapeo de prácticas de cultivo, concluyendo que cuando se realizan pruebas a escala local hay una alta dependencia de los datos terrestres y se depende en gran medida de los conocimientos locales como las prácticas de gestión, el medio ambiente y el material biológico. También, con (Yadav *et al.*, 2002) que han demostrado que los datos satelitales de la temporada de verano proporcionan información sobre la superficie de los huertos de mango, pero para desarrollar modelos de rendimiento se requiere recoger datos fiables de rendimiento y otros datos agrometeorológicos. Con Robson *et al.* (2017), quienes encontraron una relación positiva consistente del índice de vegetación usando la banda infrarroja cercana y la banda del borde rojo con el peso total de la fruta y el tamaño promedio de la fruta y concluyeron que la ubicación del huerto y la temporada de crecimiento están influyendo en esta relación. Y con Rahman *et al.* (2017) que evaluaron el potencial de las imágenes satelitales de alta resolución para estimar el rendimiento del mango integrando el área de la copa del árbol y los índices espectrales de vegetación utilizando un modelo ANN, considerando que la combinación de estos tipos de datos es posible estimar el rendimiento total de la fruta y el número de frutos con alta precisión.

Los diferentes tipos de información varían en grado de dificultad y costos para su determinación. Se puede establecer un rango de dificultad y costo crecientes a partir de los datos satelitales, seguidos por los del clima, los huertos y, por último, la información de los árboles dentro de los huertos es la más difícil de obtener y costosa. Teniendo en cuenta las variables incluidas en cada conjunto de datos y basándonos en la falta de diferencias en los resultados de svm1 entre d2 y d3, se prefiere el conjunto de datos d2 porque no incluye variables de árboles.

Para esta combinación método-conjunto de datos, el error medio de predicción fue del 3,99% con una desviación estándar igual a 3,70%. Este error en torno al 4% al estimar el número de frutos por árbol coincide con Leroux *et al.* (2019) mejorado entre el 2% y el 10% al utilizar el índice de vegetación y por Kurtulmus *et al.* (2013), que detectaron

melocotones maduros e inmaduros utilizando imágenes en color, clasificadores estadísticos y RNA, con errores en el recuento de frutos entre el 15% y el 29%. Estos resultados son también inferiores a los obtenidos en una primera instancia en este Capítulo, al utilizar información de campo y el árbol de regresión para estimar el número de frutos en naranja dulce y tangor Murcott con un 29% de error.

Definido el conjunto de datos (d2) y el método analítico (svm1), y considerando que el funcionamiento interno del SVM es difícil de entender (se tratan como modelos de caja de bloques), el entrenamiento sucesivo del método incorporando y eliminando variables y analizando el RMSE y el MAE, mediante el recuento de Borda, permite definir el orden de importancia en la predicción obteniendo el siguiente rango **Sp > Edad > Ref23 > PrecipMAD > Ref11 > Irr > Ref13 > Ref21 > NDVI1 > HumCF > TempCF > NDVI3 > Ref24 > Ref22 > Suelo > NDVI2 > NDVI4 > NDVI0 > Ref14 > TempMAD > Ref17 > Ref26 > NDVI5 > HumMAD > PrecipF > Ref16 > PrecioCF > HumF > Temporad > TempF > NDVI7 > Ref15 ~ Ref12 > NDVI6 > Ref25 > Ref27**. De esta lista, se seleccionaron las 10 primeras variables y svm1 fue llevado con un nuevo conjunto de datos llamado d2 filtrado (Tabla 6.11).

Tabla 6.11. Variables seleccionadas del nuevo conjunto de datos d2 filtrado.

Variables seleccionadas	
1	Sp
2	Edad
3	Ref23
4	PrecipMAD
5	Ref11
6	Irr
7	Ref13
8	Ref21
9	NDV1
10	HumCF

La importancia de estas variables en la cosecha ha sido definida por diferentes autores. El efecto de la especie y la edad de los árboles en la producción de cítricos es una característica bien conocida (Palacios, 2013), por lo que tiene sentido que la especie y la edad estén entre las variables más significativas en la predicción. La importancia de la humedad, la lluvia y el riego está relacionada con el aporte de agua, necesario para el desarrollo fenológico de los vegetales y el crecimiento de los órganos, Kern *et al.* (2018) encontraron una asociación de la lluvia y el rendimiento en los cultivos de invierno. El comportamiento de los valores de NDVI y reflectancia en la estimación del rendimiento se determinó previamente y se esperaba que fueran buenos y tempranos predictores del rendimiento, según las conclusiones de Lopresti *et al.* (2015) y Kern *et al.* (2018).

La Figura 6.14 presenta los gráficos de caja de MAE y RMSE comparando smv1 con d2 y d2 filtrado. La selección de variables reduce el MAE y el RMSE (incluso las diferencias no son significativas (valor p 0,05455 para el MAE y 0,2808 para el RMSE), lo que permite reducir los costos del estudio.

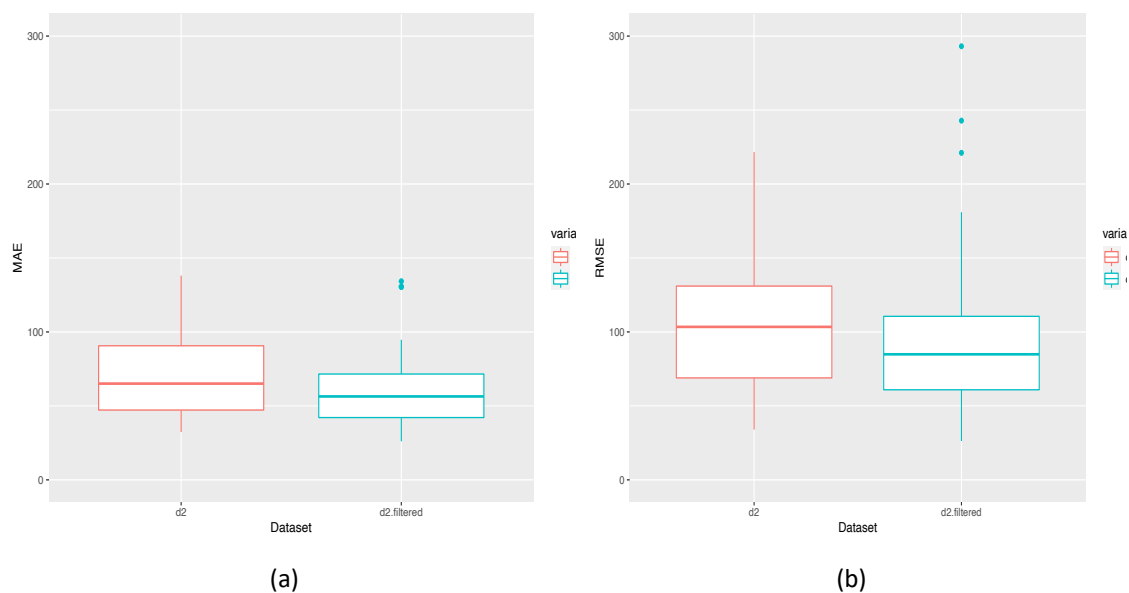


Figura 6.14. Gráficos de caja que representan: **a**, MAE y **b**, RMSE, comparando smv1 con d2 y d2 filtrado

Estos resultados permiten definir una metodología para estimar, con alta precisión, el número de frutos por árbol en tanger Murcott, y naranjas dulces tardías Salustiana y Valencia, utilizando una máquina de vectores soporte con un kernel polinómico basado en la especie, la edad de los árboles, las condiciones de riego, la

precipitación en el periodo de maduración de los frutos (abril a junio), la humedad en el periodo de crecimiento de los frutos (diciembre a marzo), la reflectancia roja e infrarroja cercana en febrero, y el NDVI, la reflectancia infrarroja cercana y roja en diciembre. La especie, la edad de los árboles y las condiciones de riego constituyen información de los huertos previamente registrada en cada uno de ellos, por lo que son variables terrestres fáciles de obtener, las precipitaciones y la humedad son accesibles en las estaciones meteorológicas, y las variables satelitales están disponibles en una amplia variedad de sensores a bordo de un número creciente de satélites civiles. La combinación seleccionada incluye variables de bajo coste y momentos tempranos de determinación, lo que asegura un método de estimación de cosecha con suficiente anticipación, alta precisión y bajo coste, lo que lo hace útil y aplicable.

## 6.4. RESUMEN DE RESULTADOS

Al analizar los resultados obtenidos con diferentes combinaciones de variables, es necesario tener en cuenta las dificultades para obtener los distintos tipos de información. La información referida a los huertos se encuentra disponible en los establecimientos, los datos descriptivos de las copas de los árboles se determinan una vez al año, la obtención de los recuentos en marcos de muestreo incrementa el trabajo de campo necesario y consecuentemente el costo, y es importante evaluar las ventajas de obtener una mayor precisión a costa de un aumento de los gastos.

Los métodos de recuento en secciones predefinidas de copa estudiados por sí solos no fueron capaces de estimar la carga real con adecuada precisión.

Los resultados de las predicciones de la carga mediante M5-Prime, de naranjo dulce y tangor Murcott fueron ligeramente diferentes. En el caso de naranjo dulce la precisión de la predicción aumentó con el incremento de la información añadida, alcanzando el error más bajo (21,3%) cuando se consideraron las variables *Básica*, *Recuentos por cuadrante* (llevados o no a volumen de copa), sin embargo, ésta es la opción más costosa, por lo que se estudiaron las diferentes combinaciones  $VolM_i + Num_j$  y se encontró que la mejor combinación en términos de precisión fue la de las variables *Básicas* más  $VolM_2$  (cono truncado). Esto significa que solamente con la introducción de alguna información básica relacionada con los árboles, como la edad, la estimación mejoró.

El comportamiento de la predicción de la carga en tangor Murcott fue similar al de naranjo dulce, esto es, cuando se añadió más información al método el error disminuyó, aunque eso implica mayores costos. No obstante, se han podido obtener errores aceptables cuando se combinó el recuento en marco contador ubicado al oeste (NumO) combinado con el volumen de copa calculado mediante  $VolM_6$  (cilindro).

Al analizar el rendimiento de los diferentes enfoques de aprendizaje automático en función del conjunto de datos utilizado como entrada y la incorporación de información de otras fuentes como variables climáticas y datos satelitales, se observa que los

diferentes métodos tienen un comportamiento similar entre los distintos conjuntos de datos.

Los modelos que presentaron menores valores de MAE y RMSE fueron los bosques aleatorios (rf) y las máquinas de vectores de soporte con kernel polinómico (svm1), siendo este último el de mejor comportamiento, independientemente del conjunto de datos (aunque en el caso de utilizar d1, los RMSEs obtenidos por rf y svm1 son similares). Se puede afirmar que las máquinas de vectores soporte con kernel polinómico (svm1) obtienen el mejor rendimiento en términos de MAE y RMSE independientemente del conjunto de datos de entrada.

Al analizar los resultados obtenidos con los diferentes conjuntos de datos, teniendo en cuenta las variables incluidas en cada uno, y la dificultad y los costos de medir diferentes tipos de variables, especialmente las que corresponden a los árboles, se enfocó el comportamiento de smv1 en los diferentes conjuntos de datos.

Se puede establecer un rango de dificultad y costo crecientes a partir de los datos satelitales, seguidos por los del clima, los huertos y, por último, la información de los árboles dentro de los huertos es la más difícil de obtener y costosa. Teniendo en cuenta las variables incluidas en cada conjunto de datos y basándonos en la falta de diferencias en los resultados de svm1 entre d2 y d3, se prefiere el conjunto de datos d2 porque no incluye variables de árboles.

Definido el conjunto de datos (d2) y el método analítico (svm1), y considerando que el funcionamiento interno del SVM es difícil de entender, el entrenamiento sucesivo del método incorporando y eliminando variables y analizando el RMSE y el MAE, permitió definir el orden de importancia en la predicción obteniendo el siguiente rango **Sp > Edad > Ref23 > PrecipMAD > Ref11 > Irr > Ref13 > Ref21 > NDVI1 > HumCF > TempCF > NDVI3 > Ref24 > Ref22 > Suelo > NDVI2 > NDVI4 > NDVI0 > Ref14 > TempMAD > Ref17 > Ref26 > NDVI5 > HumMAD > PrecipF > Ref16 > PrecipCF > HumF > Temporada > TempF > NDVI7 > Ref15 ~ Ref12 > NDVI6 > Ref25 > Ref27**. De esta lista, se seleccionaron las 10 primeras variables y svm1 fue aplicado con un nuevo conjunto de datos llamado d2 filtrado.

Los resultados obtenidos permitieron definir una metodología para estimar, con alta precisión, el número de frutos por árbol en tangor Murcott, y naranjas dulces tardías



Salustiana y Valencia, utilizando una máquina de vectores soporte con un kernel polinómico basado en la especie, la edad de los árboles, las condiciones de riego, la precipitación en el periodo de maduración de los frutos (abril a junio), la humedad en el periodo de crecimiento de los frutos (diciembre a marzo), la reflectancia roja e infrarroja cercana en febrero, y el NDVI, la reflectancia infrarroja cercana y roja en diciembre. La especie, la edad de los árboles y las condiciones de riego constituyen información de los huertos previamente registrada en cada uno de ellos, por lo que son variables terrestres fáciles de obtener, las precipitaciones y la humedad son accesibles en las estaciones meteorológicas, y las variables satelitales están disponibles en una amplia variedad de sensores a bordo de un número creciente de satélites civiles. La combinación seleccionada incluye variables de bajo coste y momentos tempranos de determinación, lo que asegura un método de estimación de cosecha con suficiente anticipación, alta precisión y bajo coste, lo que lo hace útil y aplicable.

## 6.5. BIBLIOGRAFÍA

1. Abbas, F., Afzaal, H., Farooque, A. A., & Tang, S. 2020. Crop yield prediction through proximal sensing and machine learning algorithms. *Agronomy*, 10 (7): 1046.
2. Agustí, M.; Mesejo, C.; Mesero, C., Martínez-Fuentes, A.; Zaragoza, S.; Primo Millo, E. 2020. El tamaño del fruto en los cítricos. *Levante Agrícola. Revista Internacional de Cítricos*. (451): 57-67.
3. Apolo-Apolo, O; Martínez-Guanter, J; Egea, G; Raja, P; Pérez-Ruiz, M. 2020. Deep learning techniques for estimation of the yield and size of citrus fruits using a uav. *European Journal of Agronomy*. (115).
4. Arango, R.B.; Campos, A; Combarro, EF; Canas, ER; Díaz, I. 2016 a. Mapping cultivable land from satellite imagery with clustering algorithms. *International Journal of Applied Earth Observation and Geoinformation*, 49: 99–106.
5. Arango, R.B.; Díaz, I; Campos, A; Canas, E.R.; Combarro, E.F. 2016 b. Automatic arable land detection with supervised machine learning. *Earth Science Informatics*. 9 (4): 535– 545.
6. Avanza, M.M. 2011. Desarrollo de una propuesta metodológica para la predicción de cosecha de naranjo dulce (*Citrus sinensis* L. Osbeck) var. Valencia late en la provincia de Corrientes, Argentina. Tesis Doctoral de la Universidad Nacional del Nordeste, Área Recursos Naturales
7. Avilán, L.R., Mazzi, L. 1973. Estimación de cosecha en frutales. 15 pp. (Monografía).
8. Avilán L. 1979. Efecto de las propiedades físicas del suelo sobre la distribución de las raíces del mango (*Mangifera indica* L.). *Turrialba* , 29 (2) 117-122.
9. Balzarini M.G.; Gonzalez L.; Tablada M., Casanoves F.; Di Rienzo J.A.; Robledo C.W. 2008. Infostat. Manual del Usuario, Editorial Brujas, Córdoba, Argentina.
10. Begué, A; Arvor, D; Bellon, B; Betbeder, J; De Abelleira, D; Ferraz, R; Lebourgeois, V; Lelong, C; Simões, M; Verón, S. 2018. Remote sensing and cropping practices: A review. *Remote Sensing*, 10 (1).
11. Cook, TR. 2020. *Neural Networks*, Springer International Publishing, Cham.161–189 p.

12. De Rezende Triboni, H., Barbosa, J. C. 2004. Estimativa do número de frutos por amostragem de parte da copa em laranjeiras. *Revista Brasileira de Fruticultura*, 26 (3): 454-458.
13. Díaz, I.; Mazza, S. Combarro, E.; Giménez, L.; Gaiad, José. 2017. Machine learning applied to the prediction of citrus production. *Spanish Journal of Agricultural Research*. (15).
14. Díaz, E; Fernández-Plaza, CI; Alonso, A; González, C; Díaz, I. 2021. Machine learning as a tool to study the influence of chrono-disruption in preterm births. *Journal of Ambient Intelligence and Humanized Computing*.
15. Dorj, U. O., Lee, M., and Yun, S. S. 2017. An yield estimation in citrus orchards via fruit detection and counting using image processing. *Computer Electronic Agriculture*. (140), 103–112.
16. Falivene, S.; Hardy D.S. 2008., *Assessing Citrus Crop Load Primefact 787*. ISSN 1832-6668. 4p.
17. Feng, P.; Wang, B.; Yu, Q. 2020. Machine learning-based integration of remotely sensed drought factors can improve the estimation of agricultural drought in south-eastern Australia. *Agricultural Systems*, (173): 303–316.
18. Gaiser, T; de Barros, I; Sereke, F; Lange, FM. 2010. Validation and reliability of the epic model to simulate maize production in small-holder farming systems in tropical sub-humid west Africa and semi-arid Brazil. *Agriculture, Ecosystems & Environment*, 135 (4): 318-327.
19. Georgiou, Androula. 2000. Performance of Nova mandarin on eleven rootstocks in Cyprus. *Scientia Horticulturae*. (84):115-126.
20. Gongal, A., Amatya, S., Karkee, M., Zhang, Q., and Lewis, K. 2015. Sensors and systems for fruit detection and localization: a review. *Computer Electronic Agriculture*. (116): 8–19.
21. Han, L; Yang, G; Dai, H; Xu, B; Yang, H; Feng, H; Li, Z; Yang, X. 2019. Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods*. 15 (1).

22. Hastie, T; Tibshirani, R; Friedman, J. 2017. The Elements of Statistical Learning Data Mining, Inference, and Prediction. 2nd Edition. Springer Series. 446 pp.
23. Hewson, J.; Steininger, M.K.; Pesmajoglou, S. 2014. Manual de Medición, Reporte y Verificación (MRV) de REDD+ Versión 2.0. Programa de Carbono Forestal, Mercados y Comunidades financiado por USAID. Washington, DC, Estados Unidos.
24. Hornik, K., Buchta, C. & Zeileis, A. 2009. Open-source machine learning: R meets Weka. *Comput Stat.* (24) 225-232.
25. Hothorn, T.; Hornik, K.; Zeileis, A. 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics*, 15 (3), 651-674.
26. Hyndman, R. J.; Koehler; A.B. 2006. "Another look at measures of forecast accuracy". *International Journal of Forecasting*, (22): 679-688 p.
27. Jain, AK; Duin, RPW; Mao, J. 2000. Statistical pattern recognition: a review. *IEEE. Trans. Pattern Anal. Mach. Intell.* 22 (1): 4-37.
28. James, G; Witten, D; Hastie, T; Tibshirani, R. 2017. An Introduction to Statistical Learning with Applications in R. Springer Science+Business Media New York 8th printing. 425 pp.
29. Jurecka, F; Lukas, V; Hlavinka, P; Semerádová, D; Zalud, Z; Trnka, M. 2018. Estimating crop yields at the field level using Landsat and MODIS products. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 66 (5): 1141–1150.
30. Kamilaris, A. and Prenafeta-Boldu, F. X. 2018. Deep learning in agriculture: a survey. *Computer Electronic. Agriculture.* ( 147): 70–90.
31. Karamirad, M; Omid, M; Alimardani, R; Mousazadeh, H; Heidari, S.N. 2013. Ann based simulation and experimental verification of analytical four- and five-parameters models of pv modules. *Simulation Modelling Practice and Theory*, (34): 86- 98.
32. Kern, A; Barcza, Z; Marjanovic, H; Arendas, T; Fodor, N; Bonis, P; Bogнар, P; Lichtenberger, J. 2018. Statistical modelling of crop yield in central Europe using climate data and remote sensing vegetation indices. *Agricultural and Forest*

- Meteorology. (260): 300-320.
33. Kurtulmuş, F; Lee, WS; Vardar, A. 2013. Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. *Precision Agriculture*.
  34. Lado, J; Rodrigo, M; Zacarías, L. 2014. Maturity indicators and citrus fruit quality. *Stewart Postharvest Review*, (10).
  35. Lantz; B. 2015. *Machine Learning with R: Packt Publishing Ltd*.
  36. Leroux, L; Castets, M; Baron, C; Escorihuela, MJ; Bégué, A; Lo Seen, D. 2019. Maize yield estimation in west Africa from crop process-induced combinations of multi-domain remote sensing indices. *European Journal of Agronomy*, (108): 11-26.
  37. Liakos, G.; Busato, P.; Simón Pearson D.; Bochtis, D. 2018. Review Machine Learning in Agriculture: A Review Konstantino Sensors. (18): 2674.
  38. Liu, H; Motoda, H. 1998. *Feature Selection for Knowledge Discovery and Data Mining*. Springer.
  39. Liu, H. 2005. Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on knowledge and data engineering*.
  40. Loh, WY. 2011. Classification and regression trees. *WIREs Data Mining and Knowledge Discovery*, 1 (1): 14-23.
  41. Lopresti, M.F.; Di Bella, C.M.; Degioanni, A.J. 2015. Relationship between MODIS-NDVI data and wheat yield: A case study in northern Buenos Aires province, Argentina. *Information Processing in Agriculture*, 2 (2): 73–84.
  42. Mas, J.F. *Aplicaciones del Sensor MODIS Para el Monitoreo del Territorio*, Primera edición. Secretaría de Medio Ambiente y Recursos Naturales: Mexico City, Mexico, 2011.
  43. Maya Gopal, P; Bhargavi, R. 2019. A novel approach for efficient crop yield prediction. *Computers and Electronics in Agriculture*, (165).
  44. Monteverde. E. 1996. Evaluación del Naranja 'Valencia' sobre siete patrones en Valles Altos de Carabobo-Yaracuy, Venezuela. *Agronomía Tropical*, 46 (4): 391-393.

45. Morín, C.H. 1980. Cultivos de cítricos. Lima. Stadium, S.A. 597 pp.
46. Norasyikin, F; Junita, MS; Haidi, I; Zaini, AH. 2012. Oil palm fresh fruit bunch ripeness classification using artificial neural network. 2012 4th International Conferences on Intelligence and Advances Systems (ICIAS2012). 18-21.
47. Nyalala, I; Okinda, C; Nyalala, L; Makange, N; Chao, Q; Chao, L; Yousaf, K; Chen, K. 2019. Tomato volume and mass estimation using computer vision and machine learning algorithms: Cherry tomato model. Journal of Food Engineering. (263): 288-298.
48. Orduz-Rodríguez, J.O.; Arango-Wiesner, L.; Monroy, H.; Fischer, G. 2006. Comportamiento de la mandarina Arrayana en seis patrones en suelos ácidos del piedemonte Llanero de Colombia. *Agronomía Colombiana*, 24 (2): 266-273.
49. Palacios, J. 2013. Citricultura. Editor Talleres Gráficos ALFA BETA. S. A. Argentina. ISBN: 9789-8743-266. 518 pp.
50. Parihar, J. S.; Panigrahi, S.; Dadhwal, V. K.; Bhatt, H. P.; Dass, N. K.; Ghosh, B. K.; Behra, D. 1987, Rice acreage estimation in South Orissa using stratified random sampling approach. Scientia c note, IRS-UP/SAC/CPR/SN/09/87.
51. Quilan, J. R. 1992. Learning with continuous classes. Proceeding's 5<sup>th</sup>. Australian Joint Conference on Artificial Intelligence. World Scientific, Singapore. 343-348 p.
52. Qureshi, W. S., Payne, A., Walsh, K. B., Linker, R., Cohen, O., and Dailey, M. N. (2017). Machine vision for counting fruit on mango tree canopies. *Precision Agriculture*. (18): 224-244.
53. Rahman, A. A.; Zhang, X. 2017. Prediction of cooling load for a standing wave thermoacoustic refrigerator through artificial neural network technique. Energy Procedia Proceedings of the 9th International Conference on Applied Energy. (142): 3780-3786.
54. Reeb, P. & Tassile, V. 2007. Estudios Estadísticos. Pronósticos de Producción. (<https://docplayer.es/42642541-Pronosticos-de-produccion.html> )
55. Refonaa, J; Lakshmi, M. 2021. Remote sensing-based rain fall prediction using big data assisted integrated routing framework. Journal of Ambient Intelligence and

- Humanized Computing. 1-11 p.
56. Rembold, F; Atzberger, C; Savin, I; Rojas, O; Dokuchaev, V. 2013. Using low resolution satellite imagery for yield prediction and yield anomaly detection. *Remote sensing* (5).
  57. Robson, A; Rahman, M; Muir, J. 2017. Using worldview satellite imagery to map yield in avocado (*Persea americana*): A case study in Bundaberg, Australia. *Remote Sensing*, (9).
  58. Steinberg, D. "The top ten algorithms in data mining," Ch 10, Taylor and Francis Group, LLC, 2009.
  59. Taghizadeh-Mehrjardi, R; Nabiollahi, K; Rasoli, L; Kerry, R; Scholten, T. 2020. Land suitability assessment and agricultural production sustainability using machine learning models. *Agronomy*, 10 (4).
  60. Teixidó, P.; Gómez-Galón, J.A.; Gómez-Bravo, F.; Sánchez-Rodríguez, T.; Alcina, J; Aponte, J. 2018. Low-power low-cost wireless flood sensor for smart home systems. *Sensors*, 18 (11).
  61. Torregrosa, A; Albert, F; Aleixos, N; Ortiz, C; Blasco, J. 2014. Analysis of the detachment of citrus fruits by vibration using artificial vision. *Biosystems Engineering* (119): 1-12.
  62. Wang, Y. and Witten, I. H. 1997. Induction of model trees for predicting continuous classes. 9th Eur Conf on Machine Learning, Prague (Czech Republic).
  63. Wang, W.; Ma, L.; Chen, M.; Du, Q. 2021. Joint correlation alignment-based graph neural network for domain adaptation of multitemporal hyperspectral remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, (14), 3170-3184.
  64. Williams, G 2011. *Data Mining with Rattle and R*. Springer New York. <https://doi.org/10.1007/978-1-4419-9890-3>.
  65. Xiao-Peng Song, X. Song, Wenli Huang, W. Huang, Matthew C. Hansen, M. C. Hansen, & Peter Potapov, P. Potapov. (2021). An evaluation of Landsat, Sentinel-2, Sentinel-1 and MODIS data for crop type mapping. *Science of remote sensing*. (3).

66. Yadav, I.; Rao, N. K.; Reddy, B.M.C.; Rawal, R.D.; Srinivasan, V.R.; Sujatha, N.T.; Bhattacharya, C.; Rao, P. P.; Ramesh, K.S.; Elango, S. 2002 . Acreage and production estimation of mango orchards using Indian Remote Sensing (IRS) satellite data. *Scientia Horticulturae*. (93): 105-123.
67. Yamamoto, K., Guo, W., Yoshioka, Y., and Ninomiya, S. (2014). On plant detection of intact tomato fruits using image analysis and machine learning methods. *Sensors*. (14): 12191-12206.
68. Zekri, M. 2000. Citrus rootstocks affect scion nutrition, fruit quality, growth, yield and economical return. *Fruits*, (55): 231-239.



## CAPITULO VII

### ESTIMACION ANTICIPADA DE VOLUMENES DE PRODUCCION Y DISTRIBUCION DE TAMAÑOS COMERCIALES A NIVEL PARCELARIO

#### 7.1. INTRODUCCIÓN

El modelo propuesto en el desarrollo de esta tesis (Figura 7.1) se basa en la combinación de tres componentes, la superficie (expresada por el número de árboles en producción o el número de hectáreas en producción multiplicada por la densidad de plantación), el peso del fruto medio esperado al momento de la cosecha y la carga estimada para el momento de la cosecha (cantidad de frutos por árbol).

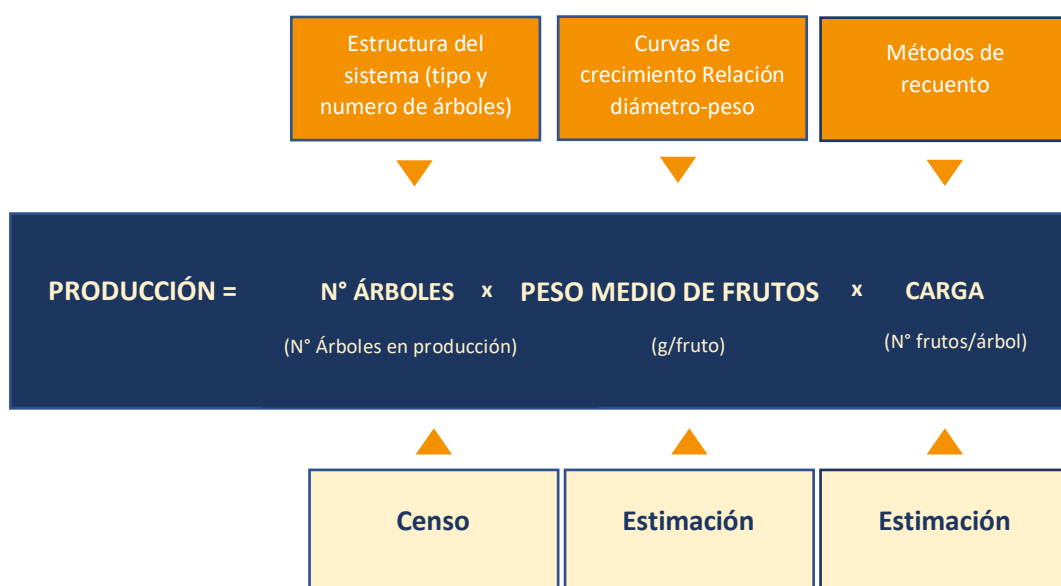


Figura 7.1. Modelo de pronóstico de producción

La determinación de la cantidad de plantas (número de árboles) se puede realizar mediante un censo de los árboles presentes, a través de croquis de los huertos (cuando estuvieran disponibles) o a partir imágenes obtenidas de sensores remotos, empleando los procedimientos que ya fueron comentados en el Capítulo I.

La estimación del peso medio que tendrán los frutos al momento de la cosecha puede ser efectuada mediante la definición de curvas de crecimiento de los frutos, para lo cual han sido estudiadas diferentes funciones matemáticas, y para el caso particular de las variedades de interés en esta tesis, han sido definidas en el Capítulo IV.

La tercera componente es la cantidad de frutos por árbol (carga) que se espera al momento de la cosecha, para cuya determinación, en el Capítulo V se han estudiado diferentes estrategias basadas en información parcelaria y en el Capítulo VI se probaron y seleccionaron diferentes combinaciones de variables provenientes de relevamientos a campo, de estaciones meteorológicas y de sensores remotos.

En el presente Capítulo se incluye la estimación de volúmenes de producción mediante 1) un método subjetivo, a partir de la opinión de informantes calificados, y 2) un método objetivo a partir de la aplicación del modelo propuesto mediante la combinación de las tres componentes definidas en capítulos previos y, complementariamente a la estimación de los volúmenes a producir, se estudia y define la distribución de tamaños de los frutos al momento de la cosecha según calibres comerciales, a partir de información de diámetros de frutos en el momento del pronóstico, transformados en peso de los frutos a la cosecha, utilizando las tablas de crecimientos desarrolladas en el Capítulo IV.

### 7.1.1. Estimación anticipada de los volúmenes de producción

Se describe metodología para el pronóstico de volúmenes de producción en frutales mediante métodos subjetivos y métodos objetivos.

#### *7.1.1.1. Pronóstico mediante métodos subjetivos*

La determinación anticipada de la producción esperada puede realizarse por métodos subjetivos a partir de la opinión de expertos, productores, técnicos o encargados de establecimientos productivos. Esta metodología implica, generalmente la realización de entrevistas en las que se solicita que calculen (estimación, antes de la cosecha) o recuerden (después de la cosecha), el rendimiento de un huerto o establecimiento individual. Antes de la cosecha se pide a los expertos que predigan qué cantidad esperan cosechar, basado en experiencias previas, comparando el rendimiento actual con el de años anteriores.

La encuesta cítrica del Estado de Florida proporciona un ejemplo específico de una encuesta de medición subjetiva. Las proyecciones para frutos cítricos se publican mensualmente en cada temporada, junto con un resumen anual de las estimaciones de producción y uso de las distintas variedades de frutos cítricos existentes en dicho país (Shields, 1992; Pollack & Perez, 2008).

Sapkota *et al.* (2016) indicaron que la estimación que realizan los agricultores es bastante precisa, pero necesita una supervisión adecuada. Consideran que si la realizan los agricultores, resulta un método económico y rápido, y cuando la realizan expertos es un método moderadamente rentable; pero las posibilidades de error aumentan si se utilizan diferentes equipos de expertos o personal de extensión para estimar el rendimiento en su propia área.

#### *7.1.1.2. Pronóstico mediante métodos objetivos*

El Instituto Nacional de Investigación Agropecuaria (INIA) de Salto, Uruguay, a través del Programa Nacional de Citricultura, ha desarrollado modelos de estimación de rendimiento, en función a las características de la floración y el desarrollo de frutos (Otero *et al.*, 2005).

En nuestro país se pueden citar como antecedentes el Programa de Pronóstico de Cosecha Frutícola de la provincia de Mendoza, con el objetivo de estimar con anterioridad, la cosecha de manzano, peral, membrillero, duraznero, ciruelo, damasco, cerezo, almendro, olivo y nogal<sup>25</sup>.

También el que se desarrolla entre la Facultad de Ciencias Agrarias de la Universidad Nacional del Comahue y los gobiernos provinciales de Río Negro y Neuquén, para la elaboración del pronóstico de las principales variedades de frutos de pepita existentes en la región<sup>26</sup>.

Y el desarrollado por Avanza (2011), para la predicción de cosecha de naranjo dulce variedad Valencia late para una zona de la provincia de Corrientes, Argentina.

### 7.1.2. Calibres comerciales y su distribución

Con el objeto de planificar las operaciones de cosecha y administrar las estrategias con antelación, los productores cítricos necesitan conocer anticipadamente, no sólo la cantidad de frutos a cosechar sino también el tamaño esperado de los mismos y la proporción de frutos aptos para destinar a los diferentes mercados nacionales y de exportación. Este conocimiento también les permitirá a los productores eliminar anticipadamente los frutos pequeños que no van a alcanzar el tamaño deseado para los mercados a los que aspiran enviar su producción.

Lötze & Bergh (2003) realizaron la predicción temprana de la distribución de frutos en la cosecha en manzano, a partir de la relación entre el tamaño de los frutos 42 días después de la plena floración, la carga y el tamaño de los frutos a la cosecha. La ecuación de regresión ajustada presentó una precisión superior al 80% para las predicciones.

Una predicción temprana de la distribución de tamaños ayudará a los productores a identificar aquellos huertos donde el porcentaje de frutos pequeños en la cosecha será

---

<sup>25</sup> Disponible en: <https://www.idr.org.ar/>

<sup>26</sup> Disponible en: <https://docplayer.es/42642541-Pronosticos-de-produccion.html>.

inaceptable, con lo que su eliminación temprana mejorará la distribución de los tamaños de frutos. La alta correlación entre el tamaño del fruto al momento de raleo y de la cosecha, estudiada por Lötze y Bergh (2002) y Williams *et al.* (1969), puede ser utilizada para identificar frutas pequeñas y ralear en consecuencia.

Giménez (2020) implementó la técnica de máquinas de vectores soporte en el pronóstico de producción en peral, logró predecir con suficiente antelación los tamaños comerciales a cosecha. Construyó una matriz de confusión contrastando los valores predichos por el método y los observados a campo, con una precisión global de 0,60 reduciendo la cantidad de categorías de tamaños de frutos logró mejorar la precisión, alcanzando un valor de 0,77.

## 7.2. MATERIALES Y MÉTODOS

A continuación, se describen los datos utilizados y las técnicas empleadas para la determinación de los volúmenes de producción a partir de informantes calificados y mediante la aplicación del modelo propuesto, como así también la metodología para la determinación de calibres comerciales y su distribución.

### 7.2.1. Estimación subjetiva de volúmenes de producción

Se seleccionaron algunos de los huertos en estudio (ver Tabla 7.1), en los que se disponía de informantes calificados en condiciones de realizar un pronóstico subjetivo de los volúmenes de cosecha, y se les solicitó que realizaran una estimación de la producción esperada para cada huerto (en tn) 15 días antes de la cosecha.

Los informantes calificados que actuaron en cada estimación subjetiva de los volúmenes de producción fueron: i) propietarios de los establecimientos, ii) encargados de los establecimientos, y iii) técnicos de la zona.

Esta estimación será comparada con los valores obtenidos mediante la aplicación del modelo basado en mediciones objetivas y con la cosecha real obtenida en esos huertos.

## 7.2.2. Estimación objetiva de volúmenes de producción

### 7.2.2.1. Datos

Para realizar la estimación de los volúmenes de producción se aplicó el modelo propuesto con las componentes establecidas en los capítulos precedentes. En la Tabla 7.1 se indican los huertos, que ya fueron descritos en el Capítulo II, de la provincia de Corrientes en los que se obtuvo la información para la determinación de las componentes del modelo objetivo, basado en mediciones parcelarias, en información obtenida de estaciones meteorológicas y de sensores remotos.

Tabla 7.1. Huertos utilizados para la estimación de volúmenes de producción y detalle de la información obtenida

Temporada	Establecimiento y huerto	N° Plantas	Superficie Efectiva (ha)	Carga estimada (frutos/árbol)	Peso fruto medio estimado (g)
2010-11	1.1	582	1,4	622,38	118,73
2015-16	1.2	1089	3,3	448,63	148,03
2015-16	1.5	1304	1,6	359,00	186,52
2013-14	2.1	733	1,8	274,49	112,61
2010-11	3.1	815	1,5	849,15	149,98
2013-14		815	1,5	434,66	151,66
2013-15		815	1,5	351,51	161,83
2015-16	3.2	949	3,3	320,61	136,89

### 7.2.2.2. Análisis de los datos

Para la realización de los pronósticos se combinaron los tres elementos del modelo propuesto:

(i) Plantas/Superficie: se determinó la cantidad de plantas en cada huerto, obtenida a partir de los croquis de los huertos. También se registró el distanciamiento en filas y entre plantas para obtener la superficie efectiva de la plantación.

(ii) Peso de los frutos: el tamaño a cosecha del fruto medio ( $y$ ) fue determinado en función de los días después de plena floración ( $x$ ) mediante las curvas de crecimiento definidas en el Capítulo III, especialmente desarrolladas para las condiciones de la región y aplicadas a la información de cada huerto, a partir de la siguiente función Logística en su quinta reparametrización:

$$y = \frac{1}{0,0140 + e^{(-2,5866) * 0,9831x}}$$

Posteriormente se realizó la estimación de la distribución de tamaños de frutos a cosecha ( $y$ ), descrita en el Capítulo IV y su transformación a peso, obteniendo el peso promedio ( $z$ ), a partir del modelo potencial que se presenta a continuación:

$$z = 0,00186 * y^{2,6643}$$

(iii) Carga: el número de frutos por árbol se determinó empleando los modelos desarrollados en los Capítulos V y VI.

- ✓ *Recuento total de frutos en cada árbol (RT)*, realizado en cada campaña en el mes de junio, de manera simultánea por dos operarios que utilizaban un contador manual y repetían la operación hasta lograr una diferencia entre ambos menor o igual al 10%. Se trabajó para cada huerto con la carga media por árbol corregida con el coeficiente de fruto oculto obtenido en el capítulo V.
- ✓ *Aprendizaje automático*, en este caso debido a que los algoritmos involucrados requieren de determinado volumen de datos se realizó una estimación de carga general considerando todos los huertos en conjunto. Posteriormente se obtuvo la carga estimada para con los huertos descritos en la tabla 7.1.

Para evaluar la precisión del modelo se calcularon los errores porcentuales (% Error) en valor absoluto, entre los volúmenes de producción estimados y los reales en cada huerto de la siguiente manera:

$$\% \text{ Error} = \frac{|\hat{\mu}_i - y_i|}{\sum_i y_i} \cdot 100$$

Donde  $\hat{\mu}_i$  es la predicción de los volúmenes de producción para el *i-ésimo* huerto e  $y_i$  es el valor de comparación, es decir volúmenes de producción real del *i-ésimo* huerto.

### 7.2.3. Distribución de calibres

Se trabajó con los calibres comerciales definidos para mandarina por la Unión Europea que comprende 10 clases, pero dadas las condiciones de los frutos de exportación de nuestro país, solamente se emplearon los 8 que se consignan en la Tabla 7.2.

Tabla 7.2. Rangos de diámetro ecuatorial (mm) mínimo, máximo y promedio, para cada tamaño comercial de frutos

Calibre comercial	Diámetro		
	Mínimo	Máximo	Promedio
1xxx	78	83	80,5
1xx	73	78	75,5
1x	68	73	70,5
1	63	68	65,5
2	58	63	60,5
3	54	58	56,0
4	50	54	52,0
5	46	50	48,0

La distribución de los tamaños comerciales se estimó a partir de información de diámetros de frutos al momento del pronóstico, transformados en peso de los frutos al momento de la cosecha, utilizando tablas de crecimientos desarrolladas en el Capítulo



IV. La predicción de los tamaños comerciales se realizó a partir de la información obtenida en el mes de junio a partir de los diámetros de frutos marcados, tomando como valor de comparación los diámetros que tenían esos mismos frutos al momento de ser recolectados (a los 325 días promedio después de la plena floración).

Con el objeto de comprobar si la distribución de tamaños comerciales estimada se correspondía con la real obtenida al momento de la cosecha, se utilizó una matriz de confusión multiclase, para lo que se asume que las categorías de los datos analizados se encuentran balanceados en similares proporciones (Karatzoglou *et al.*, 2006).

Se trata de una matriz cuadrada en la que se comparan las frecuencias reales obtenidas a cosecha para cada tamaño comercial definido, con las frecuencias predichas por la metodología propuesta. Permite evaluar la exactitud de la clasificación, situando en las filas las clases o categorías de los tamaños comerciales obtenidos a cosecha y en las columnas las mismas clases para la estimación de los tamaños comerciales. Presenta una visión general de las asignaciones, tanto de las clasificaciones correctas (elementos de la diagonal) como de las migraciones o fugas (elementos fuera de la diagonal) (Giménez, 2020).

La información que arroja la matriz de confusión no es de interpretación sencilla, por lo que existen distintos índices derivados para resumir su información (Liu *et al.*, 2007). En una clasificación estimada perfecta sólo quedarían elementos en la diagonal de la matriz.

Dos índices globales ampliamente aceptados para cuantificar la calidad de la distribución de tamaños de fruto comerciales estimada son:

- ✓ El Índice de Precisión Global, que se calcula como el número de unidades clasificadas correctamente, sobre el número total de unidades consideradas. Se obtiene sumando los elementos de la diagonal divididos por el total de observaciones. Sus valores se encuentran en el intervalo  $[0, 1]$ , y la clasificación es mejor cuanto más se acerque a la unidad, aunque este índice tiende a sobrestimar la bondad de la clasificación.

✓ El Índice Kappa: es una medida de la diferencia entre la exactitud lograda en la clasificación y la chance de lograr lo mismo con una clasificación correcta con un clasificador aleatorio. Se calcula mediante la expresión:

$$\hat{k} = \frac{n \sum_{i=1}^1 a_{ii} - \sum_{i=1}^1 (a_{i.} a_{.i})}{n^2 - \sum_{i=1}^1 (a_{i.} a_{.i})}$$

Donde  $i$  es la dimensión de la matriz (número de clases);  $a_{ij}$  el número de observaciones en la línea  $i$ , columna  $j$ ;  $a_{i.}$  y  $a_{.i}$  los totales marginales de línea  $i$  y columna  $i$ ; y  $n$  el número total de observaciones.

Tabla 7.3 Matriz de confusion

CLASES		A1	A2	...	AM	Total
	A1	$a_{11}$	$a_{12}^*$	...	$a_{1m}$	$a_{.1}$
	A2	$a_{21}$	$a_{22}$	...	$a_{2m}$	$a_{.2}$
	....			...		
	AM	$a_{m1}$		...	$a_{mm}$	$a_{.m}$
	Total c	$a_{1.}$	$a_{2.}$	...	$a_{.}$	$a_{..}$
	EXACTITUD	$a_{1.} / a_{..}$	$a_{2.} / a_{..}$	...	$a_{m.} / a_{..}$	

\*Esta notación indica en los subíndices la posición dentro de la matriz, por ejemplo en valor con  $a_{12}$  hace referencia al valor correspondiente a la fila 1, columna 2. El . (punto) indica que el valor corresponde a la sumatoria del mismo a lo largo de la fila o columna, por ejemplo:  $a_{1.}$  indica los valores de  $a$  sumados para toda la columna 1.

Existen diferentes interpretaciones de estadístico Kappa, que hacen referencia al grado de concordancia entre la clasificación lograda y la que corresponde a los valores reales. Powers (2012), propone la clasificación que se detalla en la Tabla 7.4.

Tabla 7.4 Valores Índice Kappa y su clasificación

Índice Kappa	Fuerza de la Concordancia
0,00-0,20	Pobre
0,21-0,40	Débil
0,41-0,60	Moderada
0,61-0,80	Buena
0,81-1,00	Muy buena

## 7.3. RESULTADOS Y DISCUSIÓN

### 7.3.1. Estimación anticipada de volúmenes de producción

En la Tabla 7.5 se presentan los volúmenes de producción estimados (expresados en tn por huerto) utilizando el modelo propuesto, con información de carga estimada mediante recuento total de frutos en una muestra integrada por el 2 al 3% de los árboles de cada huerto 90 días antes de la fecha esperada de cosecha, y la producción estimada por informantes calificados, lo que se comparó con el volumen de cosecha realmente producido.

Tabla 7.5. Estimación de volúmenes de producción (tn) en huertos de tangor Murcott

Temporada	Establecimiento y huerto	Producción Real (tn)	Producción Estimada (tn)	Informantes Calificados		
				Propietarios	Encargados	Técnicos
2010-11	1.1	40,00	43,01			
2015-16	1.2	79,12	72,32	85,00	80,00	75,00
2015-16	1.5	79,07	87,32	70,34	72,00	64,00
2013-14	2.1	21,91	22,66			
2010-11	3.1	106,95	103,8			
2013-14		51,52	53,72			
2013-15		47,17	46,36			
2015-16	3.2	38,3,	41,65	38,00	35,00	40,00

En la Tabla 7.6 se presentan los porcentajes de error de estimación mediante el modelo de pronóstico de producción y por informantes calificados.

Tabla 7.6. Porcentajes de error de estimación de volúmenes de producción (tn) de tangor Murcott.

Temporada	Establecimiento y huerto	Nivel Parcelario	Porcentaje de error		
			Informantes Calificados		
			Propietario	Encargado	Técnico
2010-11	1.1	7,53	7,43	1,11	-5,21
2015-16	1.2	-8,59	-11,05	-8,94	-19,06
2015-16	1.5	10,43			
2013-14	2.1	3,42			
2010-11	3.1	-2,95			
2013-14		4,27			
2013-15		-1,72	-0,78	-8,62	4,44
2015-16	3.2	8,75			
<b>Porcentaje de error medio</b>		<b>5,96</b>	<b>6,42</b>	<b>6,22</b>	<b>9,57</b>

Los errores en el pronóstico de los volúmenes de producción por huerto, a partir de mediciones objetivas parcelarias, oscilaron entre -1,72% y 10,43%, con un porcentaje de error absoluto medio de 5,96%, lo que comparado con los resultados obtenidos por otros autores se puede considerar una precisión aceptable, como los huertos no difieren mucho en superficie y lo usual es trabajar con la producción de todo el huerto, se promedia el error obtenido en cada uno de ellos.

Si bien no existen antecedentes de estimación de producción en tangor Murcott que permitan realizar una contrastación de los resultados obtenidos en esta tesis, se puede establecer que las estimaciones por métodos objetivos basadas en mediciones parcelarias son similares o incluso inferiores a los hallados por Tassile *et al.* (2013) para el pronóstico de producción de peras y manzanas, quienes a partir de un modelo semejante logran porcentajes de error de -6,88% en manzanas y de -12,26% en peras.

Estos errores también son inferiores a los obtenidos en el Centro de Información de Recursos Naturales (CIREN)<sup>27</sup> de Chile que, con un sistema basado en métodos de

<sup>27</sup> <https://www.ciren.cl>

muestreo y mediciones objetivas a nivel de huerto e integrando dos componentes (carga frutal y tamaño de fruto) obtuvieron errores de estimación en la producción no mayores al 10% para manzano, peral y uva de mesa, y del 20% para kiwi.

Los resultados obtenidos mediante estimaciones de informantes calificados indican que las estimaciones realizadas por los propietarios o los encargados de los establecimientos son mejores que las de los técnicos de la zona, probablemente por su mejor conocimiento de las historias de los huertos. No obstante, en todos los casos los errores son inferiores a los obtenidos por Apolo-Apolo *et al.* (2020), al comparar el rendimiento real y el estimado por un técnico experto con un error del 13,74%.

En el Capítulo VI, se estimó la carga seleccionando variables del huerto, climáticas y por sensores remotos, aplicando máquinas de vectores soporte, para un conjunto de huertos debido a las exigencias que tiene la técnica en cuanto al volumen de datos a procesar. Se obtuvo una carga estimada de 482,21 frutos promedio por árbol de los huertos descritos en la Tabla 7.1, en la superficie total efectiva fue de 15,8 ha en las que se involucran 6 huertos durante las temporadas 2010-15 hasta 2015-16, y el peso estimado se obtuvo de la relación peso-díametro, con esta información se aplicó en el modelo propuesto tal como se detalla en la Tabla 7.7.

Tabla 7.7. Predicción de los volúmenes de producción de tangor Murcott con información parcelaria, meteorológica y satelital

N° Plantas	Peso Fruto Medio Estimado (g)	Carga Estimada (frutos/árbol)	Producción total real (tn)	Producción total estimada (tn)	Porcentaje de error
7102	145,78	482,21	464,04	499,24	7,58

La producción estimada con información parcelaria, meteorológica y satelital fue de 498,21 tn para el conjunto de huertos de tangor Murcott descritos en la Tabla 7.1, con un error promedio de predicción de 7,58%.

Estos errores son similares a los obtenidos por Apolo-Apolo *et al.* (2020), quienes desarrollaron metodología de procesamiento de imágenes capturadas desde un vehículo aéreo no tripulado (UAV) para detectar, contar y estimar el tamaño de los frutos de cítricos en árboles utilizando técnicas de aprendizaje profundo, con un error de 7,22%.

### 7.3.2. Distribución de calibres comerciales

Del análisis realizado en el Capítulo IV, las frecuencias relativas con que se presentaron los diferentes calibres comerciales a cosecha estimados se representan en la Figura 7.1 junto a la distribución real a través de diagramas de tortas.

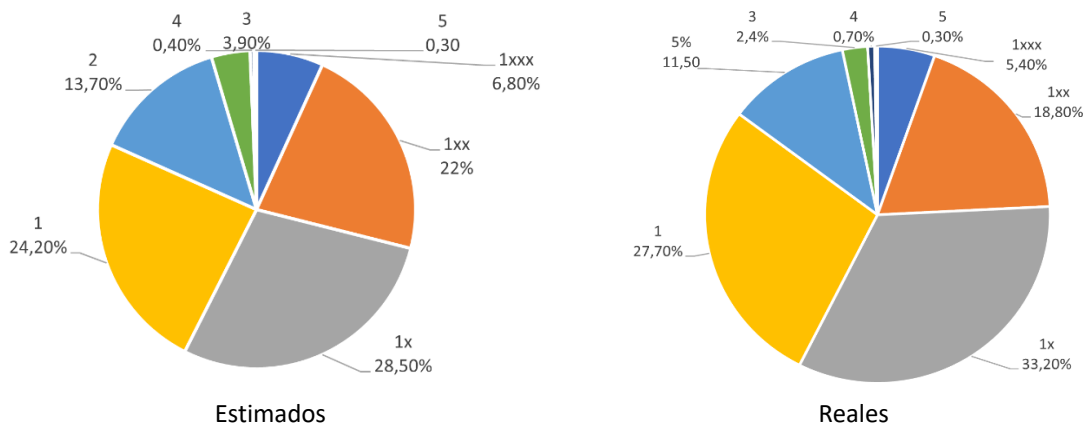


Figura 7.2. Distribución de calibres comerciales estimados y reales en tangor Murcott

Los porcentajes estimados para las categorías 1xxx, 1xx, 2 y 3 fueron superiores a los reales (con variaciones entre 1,4 y 3,3%) e inferiores para la categoría 1x, 1 y 4 (con variaciones entre -0,30 y -4,70%), las diferencias entre los porcentajes estimados y reales por categoría son pequeñas.

En la Tabla 7.8 se observa la matriz de confusión de los valores reales y predichos para los diferentes calibres comerciales.

Tabla 7.8. Matriz de confusión de las frecuencias predichas y observados a campo para los diferentes calibres comerciales

Predichos	Observados								Total
	1xxx	1xx	1x	1	2	3	4	5	
1xxx	<b>36</b>	5	0	0	0	0	0	0	41
1xx	14	<b>120</b>	10	0	0	0	0	0	144
1x	2	43	<b>189</b>	20	0	0	0	0	254
1	0	1	19	<b>157</b>	31	4	0	0	212
2	0	0	0	8	<b>74</b>	6	0	0	88
3	0	0	0	0	0	<b>18</b>	0	0	18
4	0	0	0	0	0	2	<b>3</b>	0	5
5	0	0	0	0	0	0	0	<b>2</b>	2
Total	52	169	218	185	105	30	3	2	764

Precisión Global 0,784

Kappa 0,721

Los índices obtenidos a partir de la matriz de confusión fueron: precisión global = 0,784 y  $\kappa = 0,721$ . De acuerdo con la clasificación propuesta por Powers (2012) este valor de  $\kappa$  puede considerarse bueno, lo que indica que las tablas de crecimiento desarrolladas en el Capítulo IV a partir del modelo logístico en su 5<sup>ta</sup> reparametrización presentan una adecuada capacidad predictiva de las frecuencias con que se presentarán los diferentes calibres comerciales en la cosecha.

El ajuste entre los valores predichos y los observados en esta tesis fue mejor que el reportado por Giménez (2020) quien obtuvo un  $\kappa = 0,50$  y una precisión global = 0,60 para la predicción de crecimiento de peras, aplicando máquinas de vectores soporte sin ponderar y al reducir el número de categorías de tamaños pudo lograr una precisión global = 0,77, similar al hallado en esta tesis con mayor cantidad de categorías.

## 7.4. RESUMEN DE RESULTADOS

En este Capítulo se realizó la estimación de volúmenes de producción mediante un método subjetivo (opinión de informantes calificados) y otro, objetivo, a partir de la aplicación del modelo propuesto con la combinación de número de árboles, peso medio de frutos y carga, definidos en capítulos previos. A su vez, la estimación de la carga se realizó empleando únicamente información parcelaria analizada con técnicas clásicas o información parcelaria más datos climáticos e imágenes satelitales, aplicando técnicas de aprendizaje automático.

Los errores porcentuales medios registrados se encontraron entre 6,22 y 9,57% cuando la estimación se realizó por informantes calificados, y fueron inferiores cuando lo realizaban los encargados o los dueños de los establecimientos respecto de los técnicos de la zona. En la estimación objetiva mediante el modelo de pronóstico de volúmenes a producir, cuando se efectuaron estimaciones por huerto basadas en información parcelaria, el error medio fue de 5,96% y cuando se realizaron pronósticos globales para todos los huertos empleando información parcelaria, meteorológica y satelital el error medio fue del 7,58%.

El pronóstico de volúmenes de producción realizado en esta tesis se considera que presenta una precisión aceptable comparado con las descriptas por otros autores. Tanto cuando se realizó la estimación subjetiva de los volúmenes de fruta a producir por huerto a través de las opiniones de encargados o propietarios de establecimientos como cuando se efectuó una estimación basada en mediciones objetivas, en las dos situaciones estudiadas, realizando el pronóstico por huerto basado en información parcelaria o estableciendo un pronóstico para el conjunto de huertos, empleando información de campo, meteorológica y de sensores remotos.

Complementariamente a la estimación de los volúmenes de producción, se definió la distribución de tamaños de los frutos a la cosecha según calibres comerciales, a partir de información de diámetros de frutos al momento del pronóstico utilizando tablas de crecimientos desarrolladas previamente.



Los porcentajes estimados para las categorías 1xxx, 1xx, 2 y 3 fueron superiores a los reales (con variaciones entre 1,4 y 3,3%) e inferiores para la categoría 1x, 1 y 4 (con variaciones entre -0,30 y -4,70%), las diferencias entre los porcentajes estimados y reales por categoría son pequeñas.

El índice  $\kappa$  obtenido de la matriz de confusión puede considerarse bueno, lo que indica que las tablas de crecimiento desarrolladas en el Capítulo IV a partir del modelo logístico en su 5<sup>ta</sup> reparametrización presentan una adecuada capacidad predictiva de las frecuencias con que se presentarán los diferentes calibres comerciales en la cosecha.

## 7.5. BIBLIOGRAFIA

1. Apolo-Apolo, O; Martínez-Guanter, J; Egea, G; Raja, P; Pérez-Ruiz, M. 2020. Deep learning techniques for estimation of the yield and size of citrus fruits using a uav. *European Journal of Agronomy*. (115).
2. Avanza, M.M. 2011. Desarrollo de una propuesta metodológica para la predicción de cosecha de naranjo dulce (*Citrus sinensis* L. Osbeck) var. Valencia late en la provincia de Corrientes, Argentina. Tesis Doctoral de la Universidad Nacional del Nordeste, Área Recursos Naturales.
3. Giménez, G. 2020. Análisis y Aprovechamiento De Bases De Datos Agronómicos Recurriendo Al Proceso “Knowledge Discovery In Databases”(Kdd) y Algoritmos de “DATA MINING”(DM). Una Aplicación al Pronóstico de Producción de Frutas de Pepita en los Valles de Río Negro y Neuquén. Tesis Doctoral de la Universidad Nacional de Rosario. Facultad de Ciencias Agrarias
4. Karatzoglou, A., Meyer, D., and Hornik, K. 2006. Support Vector Machines in R. *Journal of Statistical Software*, 15(9):1–28.
5. Lotze, E. and Bergh, O. 2004. Early Prediction of Harvest Fruit Size Distribution of an Apple and Pear Cultivar. *Scientea Horticulturae*, 101:281–290.
6. Lin, H.-T., Lin, C.-J.; Weng, R. C. 2007. A note on Platt’s probabilistic outputs for support vector machines. *Machine Learning*, 68:267–276.
7. Otero, A. Goñi, C., Carrau, F.; Zefferino, E.; Brennam, M. 2005. Previsión de productividad y calidad en citrus: herramientas para una gestión sustentable. *Revista INIA*, 4: 21-24.
8. Peña-Barragan, J.M.; Jurado-Exposito, M.; Lopez-Granados, F.; Atenciano, S.; Sanchez-de la Orden, M.; Garcia-Ferrer, A.; Garcia-Torres, L. Assessing land-use in olive groves from aerial photographs. *Agric. Ecosyst. Environ.* 2004, 103, 117–122
9. Pollack, S.; Perez, A. 2008. Fruit and Tree Nuts Situation and Outlook Yearbook. A report from de economic Research Service. United States Departament of Agriculture. October 2008. p 200. ([www.ers.usda.gov](http://www.ers.usda.gov)).
10. Powers, D. M. W. (2012). The Problem with Kappa. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, p. 345–355. Avignon, France. Association for Computational Linguistics

11. Sapkota, T.; MI, J.; Raj, J.; Kapoor, P.; Stirling, C. 2016. Yield Estimation of Food and Non-food Crops in Smallholder Production Systems. *Methods for Measuring Greenhouse Gas Balances and Evaluating Mitigation Options in Smallholder Agriculture*. 163-174.
12. Shields, D.A. 1992. USDA's Method of Estimating Fruit and Tree Nut Production. *Special Article. Fruit and Tree Nuts*, p. 58 - 60.
13. Tassile, V., Giménez, G., Bramardi, S., Sepúlveda, M., García, A. (2013). Resultados del Pronóstico de Producción 2013-2014. Technical report. Secretaría de Estado de Fruticultura de Rio Negro-Ministerio de Desarrollo Territorial de la Provincia de Neuquén-Facultad de Ciencias Agrarias de la Universidad Nacional del Comahue.
14. Torres, L.G.; Pena-Barragan, J.M.; Lopez-Granados, F.; Jurado-Exposito, M.; Fernandez-Escobar, R. Automatic assessment of agro-environmental indicators from remotely sensed images of tree orchards and its evaluation using olive plantations. *Comput. Electron. Agric.* 2008, 61, 179–191
15. Véga, C.; Hamrouni, A.; El Mokhtari, S.; Morel, J.; Bock, J.; Renaud, J.-P.; Bouvier, M.; Durrieu, S. Ptrees. 2014. A point-based approach to forest tree extraction from LiDAR data. *Int. J. Appl. Earth Obs. Geoinf.*, (33): 98–108.
16. Winter, F. 1976. **Ein simulationsmodell zur Untersuchung der Wirtschaftlichkeit von Apfel und Birnenanlage**. *Gartenbauwissenschaft*, 41, p: 26-34

## CAPITULO VIII

### CONCLUSIONES Y RECOMENDACIONES

#### 8.1. CONCLUSIONES

**En esta tesis se ha podido definir un modelo para la estimación anticipada de volúmenes de producción y distribución de tamaño de frutos a cosecha en tangor Murcott, para la provincia de Corrientes, Argentina.**

A partir del modelo inicialmente propuesto, basado en mediciones objetivas a campo referidas a la cantidad de árboles en producción, el peso medio de los frutos y la cantidad de frutos por árbol se ha logrado establecer metodología para la determinación o estimación de cada una de sus componentes.

Complementariamente se ha definido metodología para la estimación de la carga en tangor Murcott y naranjo Salustiana y Valencia late, que logra adecuada precisión utilizando información de bajo costo y es aplicable a pronósticos de producción para nivel de cuenca o región.

## 8.1.1. Determinación de componentes del modelo

### 8.1.1.1. Número de árboles

La determinación de la primera componente del modelo se puede realizar de manera censal, y se define a partir de croquis de los huertos o de imágenes de sensores remotos.

### 8.1.1.2. Peso medio de los frutos

**Con la finalidad de estimar anticipadamente el peso medio que tendrán los frutos se ha definido un modelo que describe las curvas de crecimiento para frutos de tangor Murcott en la región, basadas en la evolución del diámetro entre los 60 días después de plena flor y el momento de cosecha.**

Se estableció la quinta parametrización del modelo Logístico (L5) como la combinación datos-modelo-parametrización más adecuada para describir el crecimiento de frutos de tangor Murcott:

$$y = \frac{1}{\alpha + \exp(\beta)\gamma^x}$$

Donde  $y$  es diámetro ecuatorial (en mm),  $1/\alpha$  es la asíntota superior,  $1/(\alpha + \exp^\beta)$  es el intercepto y  $-\ln\gamma$  corresponde a la tasa de crecimiento de la curva desde un valor inicial ( $\beta$ ) a un valor final ( $\alpha$ ).

La aplicación de modelos no lineales mixtos permitió incluir flexibilidad en el modelo seleccionado y contemplar la variabilidad existente entre individuos (frutos). Las estimaciones de las componentes de varianza establecieron que  $(\alpha + \mu)$ , tamaño máximo de los frutos (asíntota superior), está influenciado por las condiciones del huerto, en cambio  $(\beta + \mu)$ , tamaño inicial de los frutos (intercepto), presenta mayor

variabilidad a nivel de temporada, y el parámetro  $(\gamma + \mu)$ , relacionado con la velocidad de crecimiento no presentó variabilidad a nivel temporada ni huerto.

Estos resultados evidencian que el crecimiento de los frutos está influenciado por las condiciones particulares correspondientes a la temporada y al huerto, razón por la cual se estimó una curva que modela el comportamiento medio de crecimiento.

$$y = \frac{1}{0,0140 + e^{(-2,5866) * 0,9831x}}$$

**A los efectos de facilitar la estimación del tamaño final de los frutos en cualquier momento previo a la cosecha para su aplicación en las predicciones de volúmenes de producción o en el raleo de frutos, se han construido tablas de crecimiento en diferentes diámetros de frutos destinados al mercado externo.**

Se han desarrollado tablas de crecimiento que expresan el diámetro ecuatorial medio que los frutos presentarán través del tiempo (expresado en días después de la plena floración), referenciados a 325 días después de la fecha de plena floración (momento de cosecha promedio para Corrientes, Argentina).

Se ha definido y explicado un procedimiento que permite aplicar las tablas de crecimiento desarrolladas para el raleo de frutos en tangor Murcott.

**Con la finalidad de estimar el peso de frutos a cosecha en función del diámetro, se ha definido un modelo matemático que relaciona ambos caracteres.**

Mediante el análisis de regresión no lineal, la relación entre los diámetros (y) de los frutos de tangor Murcott estudiados y sus pesos (z) se ajustó al modelo potencial:

$$z = 0,00186 * y^{2,66435}$$

A los efectos de expresar el tamaño en unidades de peso, mediante el modelo potencial descrito, se establecieron los rangos de pesos correspondiente a cada tamaño comercial y transformaron los diámetros en pesos, considerando el número de frutos para cajas de 10 y 15 kg, envases utilizados en los empaques de mandarinas en Argentina.

### *8.1.1.3. Número de frutos por árbol*

**Se han definido diferentes estrategias para la estimación de la carga en cítricos.**

#### ***A nivel parcelario***

Se ha seleccionado metodología para la estimación de carga, en tangor Murcott y naranjo dulce, mediante el recuento total de frutos corregido por un coeficiente de fruto oculto, con dos meses de anticipación a la cosecha, con un error de 9,85% luego de corregir la carga estimada mediante un coeficiente de fruto oculto igual a 0,88.

Se han analizado las diferentes fuentes de variabilidad que inciden sobre la carga y se ha podido establecer que ésta depende principalmente de las condiciones particulares de cada árbol, más que de las condiciones del huerto o la temporada.

#### ***A nivel de cuenca o región***

Mediante la aplicación de metodología para la estimación anticipada de carga empleando información de huertos, meteorológica y satelital, combinada con técnicas de aprendizaje automático y el estudio de diferentes combinaciones de variables considerando el costo de obtención de los datos, se ha logrado estimar la carga por árbol con un error del 3,99%.

El número de frutos por árbol en tangor Murcott y naranjas dulces tardías, Salustiana y Valencia, se pudo estimar utilizando una máquina de vectores soporte con kernel polinómico basado en la especie, la edad de los árboles, las condiciones de riego, las

precipitaciones en el periodo de maduración de los frutos (abril a junio), la humedad en el periodo de crecimiento de los frutos (diciembre a marzo), la reflectancia roja e infrarroja cercana en febrero, y el NDVI, la reflectancia infrarroja cercana y roja en diciembre. La combinación seleccionada incluye variables de bajo coste y momentos tempranos de determinación, lo que asegura un método de estimación de cosecha con suficiente anticipación, alta precisión y bajo coste, lo que lo hace útil y aplicable.

### 8.1.2. Precisión del modelo de pronóstico

Cuando la estimación se realizó por informantes calificados, los errores porcentuales medios registrados se encontraron entre 6,22 y 9,57%, y fueron inferiores cuando lo realizaban los encargados o los dueños de los establecimientos respecto de los técnicos de la zona. En la estimación objetiva mediante el modelo de pronóstico de volúmenes a producir, cuando se efectuaron estimaciones por huerto basadas en información parcelaria, el error medio fue de 5,96% y cuando se realizaron pronósticos globales para todos los huertos empleando información parcelaria, meteorológica y satelital el error medio fue del 7,36%.

La metodología para el pronóstico de volúmenes de producción definida en esta tesis se considera que logra una precisión aceptable, comparada con las descritas por otros autores. Tanto cuando se realizó la estimación subjetiva de los volúmenes de fruta a producir por huerto a través de las opiniones de encargados o propietarios de establecimientos como cuando se efectuó una estimación basada en mediciones objetivas, realizando el pronóstico por huerto basado en información parcelaria o estableciendo un pronóstico para el conjunto de huertos, empleando información de campo, meteorológica y de sensores remotos.



### 8.1.3. Precisión de las tablas de crecimiento

Complementariamente a la estimación de los volúmenes de producción, se definió la distribución de tamaños de los frutos a la cosecha según calibres comerciales, a partir de información de diámetros de frutos al momento del pronóstico utilizando tablas de crecimientos desarrolladas previamente.

Los porcentajes estimados para las categorías 1xxx, 1xx, 2 y 3 fueron superiores a los reales (con variaciones entre 1,4 y 3,3%) e inferiores para la categoría 1x, 1 y 4 (con variaciones entre -0,30 y -4,70%), las diferencias entre los porcentajes estimados y reales por categoría son pequeñas.

El índice  $\kappa$  obtenido de la matriz de confusión puede considerarse bueno, lo que indica que las tablas de crecimiento desarrolladas en el Capítulo IV a partir del modelo logístico en su 5<sup>ta</sup> reparametrización presentan una adecuada capacidad predictiva de las frecuencias con que se presentarán los diferentes calibres comerciales en la cosecha.

## 8.2. VERIFICACIÓN DE HIPÓTESIS

Confirmando la hipótesis de investigación, se ha desarrollado metodología para la estimación anticipada de los volúmenes de producción y la distribución de tamaño de frutos a cosecha en tangor Murcott, mediante diferentes estrategias que permiten alcanzar una precisión igual o superior a la reportada por otros autores. Esta metodología ha demostrado ser adecuada para el pronóstico de producción a nivel parcela o establecimiento.

Complementariamente, se ha desarrollado metodología para estimación de carga mediante una combinación de información de campo, meteorológica y proveniente de sensores remotos, aplicable a tangor Murcott y naranjo dulce Salustiana y Valencia late para las provincias de Corrientes y Entre Ríos, con adecuada precisión. Debido al gran volumen de datos que requiere, esta metodología es aplicable para realizar pronósticos de producción en cuencas o grandes regiones.

### 8.3. RECOMENDACIONES

**Los resultados obtenidos en esta tesis permiten abrir líneas complementarias de investigación, por lo que se recomienda continuar investigaciones en relación con:**

- ✓ La determinación de metodología para la estimación anticipada de tamaño de fruto en otros cítricos de interés.
- ✓ La estimación anticipada de la carga en otros frutales cítricos para la región nordeste de la Argentina.
- ✓ El pronóstico de los volúmenes de producción de cítricos a nivel cuenca o región en el nordeste argentino, empleando información de campo, meteorológica y satelital.