



**FACULTAD DE CIENCIAS AGRARIAS
UNIVERSIDAD NACIONAL DE ROSARIO**

**MODELIZACIÓN DE SISTEMAS BIOLÓGICOS PARA MANEJO DE FRUTALES
MEDIANTE SENSORES REMOTOS Y PROXIMALES**

DOLORES DEL BRIO

TESIS PARA OPTAR AL TÍTULO DE DOCTORA EN CIENCIAS AGRARIAS

**DIRECTOR: PABLO DANIEL REEB
CODIRECTORES: SERGIO BRAMARDI Y DARIO EDUARDO FERNANDEZ**

AÑO 2023

**MODELIZACIÓN DE SISTEMAS BIOLÓGICOS PARA MANEJO DE FRUTALES
MEDIANTE SENSORES REMOTOS Y PROXIMALES**

Dolores del Brio

Ingeniera Agrónoma – Universidad Nacional del Comahue

Esta Tesis es presentada como parte de los requisitos para optar al grado académico de Doctor en Ciencias Agrarias, de la Universidad Nacional de Rosario y no ha sido previamente presentada para la obtención de otro título en ésta u otra Universidad. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en La Estación Experimental 'Ing. Agr. Carlos H. Casamiquela – Alto Valle' del Instituto Nacional de Tecnología Agropecuaria, durante el período comprendido entre los años 2018 y 2023, bajo la dirección del Dr. Pablo Daniel Reeb.

Dolores del Brio

Pablo Daniel Reeb

Darío Eduardo Fernández

Sergio Jorge Bramardi

Agradecimientos

A mis co-directores de beca, por acompañarme en este proceso.

A los técnicos del grupo de Fruticultura y de Ingeniería en Biosistemas de INTA por su gran ayuda a la hora de realizar los trabajos de campo y cosechas.

A Mariela Curetti y Valentín Tassile por su gran ayuda y excelente predisposición durante estos años. También por compartir conmigo todos sus conocimientos y experiencias.

A Gustavo Giménez, por compartirme sus experiencias en la UNR y motivarme a elegir esta carrera. También por todos sus aportes a esta tesis.

A mis compañeras, las becarias, por los cursos compartidos, ensayos, charlas, mates y por hacer más lindos y entretenidos los días de trabajo.

Agradezco especialmente a Nacho y a mis hijos, por brindarme siempre su apoyo, su amor, y su paciencia. Porque les ha tocado ayudarme en el campo, acompañarme en viajes de estudio, o esperarme en casa, y siempre lo hicieron con mucho amor.

A mis hermanas por estar siempre que lo necesité, y por último, quisiera destacar a mi madre, por su ayuda incondicional y por inculcarme siempre la importancia y el valor del estudio y el conocimiento.

Dedicatoria

Esta tesis se la dedico a mis tres hijos. Porque creo que el ejemplo tiene más fuerza que las reglas, y también creo que no es la única, pero sí la mejor forma de enseñar. Deseo que ellos tengan la certeza de que siempre se puede aprender algo más, que sean curiosos, perseverantes y que sepan que a pesar de los obstáculos que se interpongan en el camino, todo se puede lograr con trabajo, esfuerzo y amor.

Publicaciones en revista

1. Del Brio, D., Tassile, V., Fernandez, D.E., Bramardi, S.J., Reeb, P.D. (2023). Apple (*Malus domestica*) and pear (*Pyrus communis*) yield prediction after tree image analysis. Revista de la Facultad de Ciencias Agrarias UNCuyo, Aprobado el 02/08/2023, en prensa.
2. del Brio, D., Tassile, V., Reeb, P.D., Bramardi, S.J. (2024). Incorporation of environmental covariates to nonlinear mixed models to describe fruit growth. Revista de Investigaciones Agropecuarias. 49:3. <https://doi.org/10.58149/14h1-sp68>.

Presentaciones a congresos

1. del Brio, D., Tassile, V., Fernández, D.E., Reeb, P.D. (2022). Estimación del tamaño de frutos a partir de imágenes digitales. Presentado en XXVI Reunión Científica del grupo argentino de Biometría, Córdoba, Argentina.
2. del Brio, D., Tassile, V., Fernández, D.E., Reeb, P.D. (2022). Detección y conteo de peras a partir de imágenes mediante uso de técnicas de inteligencia artificial. Presentado en XXVI Reunión Científica del grupo argentino de Biometría, Córdoba, Argentina.
3. del Brio, D., Tassile, V., Reeb, P.D. (2021). Detección y conteo de Manzanas Mediante Análisis de Imágenes para Pronósticos de Producción. Presentado en las 50ª Jornadas argentinas de Informática, Congreso argentino de agroinformática, Virtual, Argentina.
4. del Brio, D., Tassile, V., Reeb, P.D. (2020). Aplicación de modelos mixtos no lineales en el crecimiento de manzanas cv. 'Red chief'. Presentado en XXV Reunión Científica del grupo argentino de Biometría, Tandil, Argentina.
5. del Brio, D., Illescas, J., Fernández, D.E. y Reeb, P.D. (2019). Estimación del número de flores y frutos de manzana a partir de imágenes digitales con ImageJ. Presentado en XXIV Reunión Científica del grupo argentino de Biometría, Mendoza, Argentina.
6. del Brio, D., Illescas, J., Raverta-Capua, F., Fernández, D. y Reeb, P.D. (2019). Detección de manzanas en imágenes digitales a partir de rasgos de color utilizando

Python. Presentado en XXIV Reunión Científica del grupo argentino de Biometría, Mendoza, Argentina.

7. del Brio, D., Fernández, D., Mañuecos, L., Montenegro, A., Reeb, P. D. (2019). Estimación del Tamaño de Manzanas mediante Análisis de Imágenes Digitales para Pronósticos de Producción. Presentado en el 11° Congreso argentino de agroinformática, Salta, Argentina.
8. Del Brio, D., Ocampo, S., Reeb, P.D. (2018). Estimación del número de ramilletes florales en perales Williams mediante análisis de imagen para pronósticos de producción. Presentado en la XXIII Reunión Científica del grupo de Biometría - II Encuentro Argentino- Chileno de Biometría. Neuquén Capital, Argentina.

Data sets

1. del Brio, D., Reeb, P.D., Fernandez, D.E. (2020). IMAGE SET FOR FLOWER DETECTION IN APPLE TREES. Mendeley Data, V2, doi: 10.17632/bpbsfs26td.2
2. del Brio, D., Reeb, P.D., Fernández, D.E. (2020). IMAGE SET FOR FLOWER DETECTION IN PEAR TREES. Mendeley Data, V1, doi: 10.17632/w2t9w75cv6.1

Índice de Contenidos

1. CAPITULO I: MARCO TEÓRICO.....	22
1.1. INTRODUCCION GENERAL	23
1.1.1. Ubicación y caracterización de la zona en estudio	23
1.1.2. Producción regional de frutales de pepita.....	25
1.1.3. Características de los cultivos.....	26
1.1.4. Aspectos agronómicos de los cultivos.....	27
1.1.5. Crecimiento de frutos de pepita.....	29
1.2. PLANTEO DEL PROBLEMA.....	30
1.3. PRONOSTICO DE PRODUCCION	31
1.4. OBJETIVOS	33
1.4.1. Objetivo general	33
1.4.2. Objetivos específicos	33
1.5. HIPÓTESIS DE TRABAJO	33
2. CAPITULO II: DETECCIÓN Y CONTEO DE FRUTOS	34
2.1. INTRODUCCION	35
2.1.1. Estimación de la carga frutal	35
2.1.2. Incorporación de Análisis de imágenes	35
2.1.3. Procesamiento y análisis de imágenes.....	37
2.1.4. Inteligencia artificial, aprendizaje automático y aprendizaje profundo	40
2.1.5. Modelos pre-entrenados de ML para la detección de objetos:	42
2.1.6. Métricas de evaluación de modelos de visión artificial.....	45
2.1.7. Antecedentes del uso y evaluación de los modelos de visión artificial para la detección de frutos	46
2.2. MATERIALES Y METODOS	49

2.2.1.	Características de los montes frutales.....	49
2.2.2.	Obtención de datos para la incorporación de información proveniente de sensores a los modelos predictivos.....	51
2.2.3.	Análisis de imágenes	54
2.3.	RESULTADOS Y DISCUSION	59
2.3.1.	Detección y conteo de frutos	59
2.3.2.	Correlación entre frutos contados vs estimados	66
2.3.3.	Estimación del tamaño de frutos	70
2.4.	CONCLUSIONES	74
3.	CAPITULO III: CURVAS DE CRECIMIENTO	75
3.1.	INTRODUCCION	76
3.1.1.	Estimación del tamaño final de los frutos	76
3.1.2.	Curvas de crecimiento	77
3.1.3.	Incorporación de variables climáticas e índices verdes a las curvas de crecimiento	78
3.1.4.	Modelos mixtos no lineales	81
3.1.5.	Aspectos estadísticos del ajuste y evaluación de los modelos mixtos no lineales	82
3.2.	MATERIALES Y METODOS	86
3.2.1.	Obtención de los datos.....	86
3.2.1.1.	Temporada 2017/2018	86
3.2.2.	Obtención de datos meteorológicos.....	87
3.2.3.	Obtención de índices verdes a partir de imágenes satelitales.....	87
3.2.4.	Ajuste de curvas de crecimiento	88
3.3.	RESULTADOS Y DISCUSIÓN	90
3.3.1.	Análisis exploratorio de los datos.....	90
3.3.2.	Resultados de los ajustes de los modelos de crecimiento para peras	92
3.3.3.	Resultados de los ajustes de los modelos de crecimiento para manzanas	100

3.3.4.	Discusión acerca de los modelos de crecimiento obtenidos.....	106
3.4.	CONCLUSIONES	109
4.	CAPITULO IV: PRONÓSTICOS DE PRODUCCIÓN	110
4.1.	INTRODUCCION	111
4.1.1.	Pronósticos de producción.....	111
4.1.2.	Comparación de modelos de predicción.....	113
4.1.3.	Estimación de la distribución de tamaños	114
4.2.	MATERIALES Y METODOS	116
4.2.1.	Pronostico de producción	116
4.2.2.	Cálculo de presupuestos parciales y comparación	118
4.2.3.	Estimación de la distribución de tamaños comerciales	119
4.3.	RESULTADOS	120
4.3.1.	Pronóstico de producción	120
4.3.2.	Comparación de costos y precisión de ambos pronósticos.....	124
4.3.3.	Estimacion de la distribución de tamaños	125
4.4.	CONCLUSIONES	129
5.	CONCLUSIONES GENERALES	130
6.	BIBLIOGRAFIA.....	133
7.	ANEXO.....	147
7.1.	Gráficos de residuales para el modelo seleccionado en peras.....	148
7.2.	Gráficos de residuales para el modelo seleccionado de manzanas:	150

Índice de figuras

Figura 1: Alto Valle del Rio Negro y Neuquén. Elaboración propia.	24
Figura 2: Curva doble sigmoidea. Elaboración propia.	29
Figura 3: Diagrama de flujo del análisis de imágenes. Elaboración propia.	39
Figura 4: ML: nuevo paradigma de programación. (Extraído de Chollet, 2018).....	40
Figura 5: Esquema de capas en una RNA. (Extraído de Bagnato, 2020).....	41
Figura 6: Izquierda: Parcela de perales cv ‘Williams’. Derecha: Parcela de manzana. Ambas de la Estación experimental INTA Alto Valle.....	49
Figura 7: a) Fotos de reposo invernal. Izquierda: árbol de pera cv William’s. Derecha: árbol de manzana cv Red chief; b) Fotos en floración de un árbol de pera cv ‘William’s’. Izquierda: foto tomada de día. Derecha: foto tomada de noche; c) Fotos con frutos de un árbol de manzana cv ‘Red chief’. Izquierda: foto tomada de día. Derecha: foto tomada de noche. 53	
Figura 8: Mapas de anotaciones en set de imágenes e histograma de objetos contados por imagen. A) Manzanas; B) Peras. Tomados de https://app.roboflow.com/doloresdb-fbaqe/williams-lfylv/health	55
Figura 9: Imágenes de manzanas y peras numeradas en plantas y bandeja; y de la escala utilizada	58
Figura 10: a) Gráfico de precisión vs pasos del entrenamiento; b) Gráfico de sensibilidad vs pasos del entrenamiento. c) Gráfico de precisión promedio 0.5 versus pasos del entrenamiento. Gráficos de los MPDO utilizados para imágenes de manzana. Obtenido de https://wandb.ai/delbrioyolov5s-W?workspace=user-delbrioyolov5s-W	60
Figura 11: a) Gráfico de precisión vs pasos del entrenamiento; b) Gráfico de sensibilidad vs pasos del entrenamiento. c) Gráfico de precisión promedio 0.5 versus pasos del entrenamiento. Gráficos de los MPDO utilizados para imágenes de pera. Obtenido de https://wandb.ai/delbrioyolov5s-W?workspace=user-delbrioyolov5s-W	61
Figura 12: Izquierda: Gráfico de F1 score para el MPDO seleccionado para manzanas. Derecha: Gráfico de F1 score para el MPDO seleccionado para peras. Obtenido de https://wandb.ai/delbrioyolov5s-W?workspace=user-delbrioyolov5s-W	62
Figura 13 : a) Izquierda: foto de árbol de manzana tomada de día con los frutos etiquetados manualmente; Derecha: misma foto con frutos detectados por MPDO; b) Izquierda: foto de	

árbol de manzana tomada de noche con frutos etiquetados manualmente; Derecha: misma foto con frutos detectados por MPDO	65
Figura 14: Diagramas de dispersión entre frutos contados manualmente del lado que se tomó la imagen, los contados manualmente en el total de la planta, los etiquetados en la imagen y los cosechados. Izquierda: Manzanas. Derecha: Peras	69
Figura 15: Diagramas de dispersión de las mediciones manuales de los diámetros de los frutos versus sus estimaciones a partir de análisis de imágenes.	71
Figura 16: Gráfico de frecuencias absolutas para las diferencias entre los diámetros obtenidos por imágenes (DI) y los medidos manualmente con calibre digital (DM).	72
Figura 17: Izquierda: Imagen Sentinel 2 del cuadro de manzanas cv `Red chief` con grilla que indica el tamaño de resolución de pixel para el cálculo del NDVI. Derecha: índices NDVI para cada pixel de la grilla.	88
Figura 18: Diagramas de dispersión de los diámetros de los frutos a lo largo de los DDPF para cada especie durante las temporadas 2018-2018; 2018-2019 y 2019-2020 en INTA.. Los colores representan el tamaño de frutos: azul corresponde a grandes, verde a medianos y naranja a pequeños.....	90
Figura 19: Diagramas de dispersión de los diámetros de los frutos a lo largo de los DDPF para cada sitio y especie de la temporada 2020-2021. Los colores representan el tamaño de frutos: azul corresponde a grandes, verde a medianos y naranja a pequeños.	91
Figura 20: Tabla de crecimiento de frutos, tomada de Reeb et al., 2018.	115
Figura 21: esquema del modelo general del pronóstico de producción regional	117
Figura 22: Histogramas de frecuencias relativas observadas y pronosticadas para cada clase de tamaño de frutos a cosecha en peras.	126
Figura 23: Histogramas de frecuencias relativas observadas y pronosticadas para cada clase de tamaño de frutos a cosecha en manzanas.....	127

Índice de tablas

Tabla 1: Características de los montes frutales utilizados para los muestreos	50
Tabla 2: Tamañado de peras y manzanas. Clases y rangos de pesos y diámetros asociados.	54
Tabla 3: Pueba t para muestras apareadas sobre la igualdad a cero de la media de las diferencias entre los números de frutos detectado por los modelos entrenados (visión artificial) y los números de frutos contados manualmente sobre el mismo lado del árbol para cada árbol.	64
Tabla 4: Correlaciones entre frutos contados manualmente al momento del pronóstico del lado de la imagen (Manual_lado_imagen); los frutos contados manualmente al momento del pronóstico de todo el árbol (Manual_Total); los detectados por el ojo humano en las imágenes (Etiquetados), los detectados por los MPDO (Detectados) y los cosechados (Cosechados).	67
Tabla 5: Prueba t de comparación de medias de los diámetros de frutos medidos manualmente (grupo 1) y los estimados a partir de imágenes (grupo 2) para manzanas en campo.	73
Tabla 6: Pruebas Wilcoxon de comparación de medias de los diámetros de frutos medidos por imágenes (grupo 1) y los medidos manualmente (grupo 2), para manzanas en laboratorio y peras a campo.	73
Tabla 7: Comparación de modelos de crecimiento de pera que contemplan los diferentes efectos aleatorios en estudio	93
Tabla 8: Comparación de los modelos con diferentes covariables basados en Pera_3	95
Tabla 9: Comparación de los modelos candidatos al momento y aquellos modelos que contemplan ambos tipos de covariables en peras.	96
Tabla 10: Estimaciones de efectos aleatorios a nivel de sitio para el modelo Pera_3_MIX_3	97
Tabla 11: Estimaciones de efectos aleatorios a nivel de fruto anidado a sitio para el modelo Pera_3_MIX_3	97
Tabla 12: Estimaciones de efectos fijos para el modelo Pera_3_MIX_3	98
Tabla 13: Incidencia de los efectos aleatorios en los diferentes niveles considerados.....	98
Tabla 14: indicadores de bondad de ajuste y capacidad predictiva obtenidos para los modelo candidatos	99
Tabla 15: Comparación de modelos de crecimiento de manzana que contemplan los diferentes efectos aleatorios en estudio	100

Tabla 16: Comparación de los modelos con diferentes covariables basados en Manzana_3	102
Tabla 17: Comparación de los modelos candidatos al momento y aquellos modelos que contemplan ambos tipos de covariables en manzanas.....	103
Tabla 18: Estimaciones de efectos aleatorios a nivel de sitio para el modelo Manzana_3_MIX_1	104
Tabla 19: Estimaciones de efectos aleatorios a nivel de fruto anidado a sitio para el modelo Manzana_3_MIX_1	104
Tabla 20: Estimaciones de efectos fijos para el modelo Manzana_3_MIX_1	104
Tabla 21: coeficientes de variación para los efectos aleatorios en los diferentes niveles considerados.	105
Tabla 22: indicadores obtenidos en los modelos candidatos para el análisis de la capacidad predictiva	105
Tabla 23: Clases creadas para la estimación de la distribución de tamaños a cosecha en peras	119
Tabla 24: Clases creadas para la estimación de la distribución de tamaños a cosecha en manzanas	119
Tabla 25: Rendimientos obtenidos y estimados mediante pronóstico tradicional en tn/ha para manzanas	120
Tabla 26: Rendimientos obtenidos y estimados mediante pronóstico por imágenes en tn/ha para manzanas.....	120
Tabla 27: Rendimientos obtenidos y estimados mediante pronóstico tradicional en tn/ha para peras.....	121
Tabla 28: Rendimientos obtenidos y estimados mediante pronóstico por imágenes en tn/ha para peras.....	121
Tabla 29: Comparación de la producción estimada por el pronóstico de producción regional de las provincias de Rio Negro y Meuquen, respecto a la producción total registrada por la secretaria de fruticultura de Rio Negro (en miles de toneladas).....	123
Tabla 30: tareas y horas de trabajo asociados a cada metodología de pronóstico.....	125

Índice de abreviaturas

AIC: criterio de información de Akaike
BIC: criterio de información Bayesiano
CCC: coeficiente de concordancia
DDPF: días después de plena floración
DL: deep learning
EBEs: estimadores bayesianos empíricos
Faster R-CNN: Faster Regional Convolutional Neural Network
FN: falso negativo
FP: falso positivo
GPU: unidad de procesamiento de gráficos
HSL: hue, saturation, luminance
IA: inteligencia artificial
INTA: Instituto Nacional de Tecnología Agropecuaria
IoU: intersección sobre unión
LRT: cociente de verosimilitud
mAP: media de la precisión
ML: machine learning
MP: megapíxeles
MPDO: modelos pre-entrenados de detección de objetos
MSI: índice de estrés por humedad
NDVI: índice de vegetación normalizado
NLMM: modelos mixtos no lineales
NIR: infrarrojo cercano
NN: neural network o redes neuronales
 R^2 : coeficiente de determinación
 R^2_{aj} : coeficiente de determinación ajustado
 R^2_{cond} : coeficiente de determinación condicionado
RED: banda del rojo

REML: máxima verosimilitud restringida
RGB: red, green and blue
RI: región de interés
RNA: redes neuronales artificiales
RNC: redes neuronales convolucionales
RSME: raíz cuadrada del error medio
SM: sesgo medio
SMA: sesgo medio absoluto
SMR: sesgo medio relativo
SMRA: sesgo medio absoluto relativo
SRL: sociedad de responsabilidad limitada
SSD: Single Shot MultiBox Detector
SVM: support vector machine
SWIR: infrarrojo medio
TBVI: índice de vegetación de dos bandas
UE: unidades experimentales
VN: verdaderos negativos
VP: verdaderos positivos
YOLO: You Only Look Once

Resumen

Estimar el rendimiento de un monte frutal de manera anticipada es muy importante para mejorar la organización y la logística de las actividades tanto en la etapa de producción, como en la de cosecha y posterior comercialización. Al momento de realizar pronósticos de producción en frutales de pepita, existen dos aspectos fundamentales que se deben considerar. Por un lado, conocer el número de frutos presentes en los árboles, y, por otro lado, conocer el tamaño de los mismos para poder estimar su peso a cosecha. Hasta el momento, la metodología utilizada a nivel regional para recolectar esta información es el conteo manual del número de frutos, y la medición del diámetro ecuatorial de los frutos utilizando un calibre digital. Sin embargo, estas metodologías no siempre son precisas y requieren de mucho tiempo de medición a campo y son costosas en mano de obra. Por ello resulta de gran interés la búsqueda de alternativas como el análisis de imágenes para realizar estas tareas.

El objetivo de la presente tesis fue implementar un modelo predictivo del rendimiento parcelario a cosecha en frutales de pepita, adaptado a las condiciones del Alto Valle de Río Negro y Neuquén. Los objetivos específicos fueron: analizar y seleccionar variables relevantes para el modelo de predicción de cosecha; adaptar y aplicar metodologías de escaneo de frutos en base a sensores proximales; comparar calidad de predicción y costos entre metodologías que impliquen conteo de fruto manual y técnicas de conteo con sensores y desarrollar un modelo de predicción de tamaño, peso y número de frutos (rendimiento) a cosecha en manzanos y perales.

Para ello se trabajó en dos montes frutales de la Estación Experimental Agropecuaria INTA Alto Valle, uno de manzanas cv 'Red chief' y uno de peras cv 'Williams' y en cuatro montes frutales comerciales de empresas con los mismos cultivares.

Como se mencionó previamente, el primero de los elementos a conocer para realizar pronósticos de cosecha es el número de frutos presentes en los árboles. Para esto se realizaron conteos manuales de frutos por árbol (metodología tradicional) al momento de realizar los pronósticos (principios de diciembre) y además el mismo día del conteo se capturaron imágenes provenientes de cámaras digitales RGB. Se tomaron fotos de los árboles de día y de noche durante 4 temporadas. También se obtuvo el dato de número total de frutos por planta a cosecha. Con las imágenes se entrenaron algoritmos de visión artificial para la detección de objetos basados en redes neuronales que puedan detectar y contar manzanas y peras a partir de imágenes, como alternativa al conteo manual. Los modelos se evaluaron según

sensibilidad, precisión y F1score; y se calculó la correlación entre frutos detectados por imágenes y los efectivamente cosechados. La precisión, sensibilidad y F1score alcanzados por los modelos fueron 0,86; 0,83 y 0,84; respectivamente. El coeficiente de correlación entre el número de frutos por árbol obtenido manualmente y el número de frutos detectados por el modelo fue 0,88 para manzanas y 0,69 para peras. Estos resultados pueden mejorarse si se aumenta la cantidad de datos de entrada a los modelos utilizados con más imágenes y etiquetas de frutas tomadas en la mayor diversidad de condiciones posibles. La posibilidad de estimar los frutos a partir de imágenes, reduce el tiempo de mano de obra destinada a dicha tarea, y esto permitiría aumentar el número de árboles muestreados para realizar los pronósticos.

También se trabajó en la estimación de diámetros a partir de imágenes. Para esto se tomaron imágenes de frutos en arboles colocando elementos de dimensiones conocidas como referencia. Luego se trazaron líneas rectas sobre cada fruto para estimar su tamaño y se calculó el coeficiente de correlación entre las mediciones con calibre digital y las estimaciones a partir de imágenes. Las correlaciones entre diámetros estimados y medidos manualmente fueron de 0,73 en manzanas y 0,80 en peras. Las estimaciones no resultaron del todo precisas y los resultados variaron en función de la cercanía de los frutos al elemento de dimensiones conocidas, y según la posición que tome el fruto en el árbol y cuan oculto quede por otras partes del árbol.

El segundo elemento a conocer para realizar pronósticos es el tamaño de los frutos a cosecha. Para esta estimación se trabajó en el ajuste de modelos mixtos no lineales que describan el crecimiento de los frutos de pera y manzana para la región en estudio. Para ello se seleccionaron frutos cada temporada y se midió semanalmente sus diámetros con calibre digital desde los 30 días después de plena floración hasta cosecha. Con los datos obtenidos se ajustaron diferentes modelos que contemplaron tanto efectos fijos como aleatorios a diferentes niveles de agrupamiento. Como resultado se obtuvo que el crecimiento de manzanas cv 'Red Chief' y peras cv 'Williams' en el Alto Valle de Rio Negro quedó descrito por la curva logística en su tercera parametrización. El uso de modelos mixtos no lineales multiniveles presentó la ventaja de poder discernir y cuantificar fuentes de variabilidad a diferentes niveles. El efecto aleatorio que mayor influencia tuvo en los parámetros estimados de los modelos fue el fruto. La incorporación de variables ambientales e índices verdes en los modelos mejoró el comportamiento de los mismos, permitiendo contemplar la variabilidad asociada a cada temporada y sitio analizado. Si bien los modelos que incorporan índices verdes lograron describir bien el crecimiento de los frutos, a los fines del pronóstico no resultan prácticos si no están acompañados de un paquete tecnológico apropiado, por lo cual para hacer pronósticos

se utilizarán los modelos sin dichos índices. La capacidad predictiva de los modelos cuando se calculó con datos no contemplados en el proceso modelado fue razonable, con valores de sesgo medio cercanos a un milímetro en peras y menores para manzanas; y de sesgo medio relativo cercanos al dos y al uno por ciento para peras y manzanas respectivamente.

El número de árboles o tamaño del monte frutal se obtuvo por registros del productor y con todos estos elementos se realizaron dos pronósticos diferentes para luego compararlos con los rendimientos reales: i) basado en información recolectada por metodologías de conteo manual o tradicional; ii) basado en información obtenida por imágenes.

Los pronósticos realizados de forma tradicional presentaron errores relativos aceptables tanto para peras (hasta 10%) como para manzanas (hasta 5%), por lo cual representan una herramienta de gran utilidad para estimar la producción a nivel de parcela. Por su parte, los pronósticos realizados a partir de imágenes arrojaron errores relativos desde tres a diez por ciento para manzanas y un poco mayores en peras (hasta 25% en uno de los montes frutales). Esto se explicó por la estructura de los montes de pera y la mayor dificultad de detección de frutos de peras en las imágenes. En casi todos los montes frutales el pronóstico tradicional fue más preciso que el obtenido a partir de imágenes, sin embargo, el costo de realización es muy superior respecto a este último. Considerando que los errores son aceptables, y que es posible mejorar el pronóstico por imágenes, representa una gran alternativa.

Las estimaciones de distribución de tamaños a cosecha no representaron adecuadamente las distribuciones obtenidas realmente, lo cual se asoció a un tamaño de muestra pequeño y a errores en el diseño del experimento.

A partir de los resultados surge también la necesidad de seguir trabajando a futuro en la mejora de la precisión de los algoritmos de detección de frutas a partir de imágenes y en el desarrollo de metodologías precisas de estimación de tamaños de frutos ya sea de forma manual o también a partir de imágenes. Además, sería interesante evaluar la posibilidad de construir un paquete tecnológico que permita medir (ya sea tamaños o número de frutos) y a la vez registrar las coordenadas geográficas de las respectivas mediciones para contar con datos de índices verdes de cada registro y poder incorporarlos a los pronósticos.

Palabras clave: peras, manzanas, análisis de imágenes, pronóstico de producción, visión artificial, modelos mixtos no lineales.

Abstract

Anticipating the yield forecast in fruit orchards is crucial to improve the organization and logistics within the activities in the production stage, fruit picking labors and ultimately commercialization. There are two key aspects to be consider in apple and pear yield forecast: having the number of fruit present in the trees and getting to know the size in order to estimate their weight in harvest. Up to know manual fruit counting and manual diameter measurement have been the only methodology used regionally to collect this information. However, these techniques have not proved to be accurate enough and eventually require many hours of work. Thus image analysis becomes crucial as an alternative method. The aim of this research is to implement an innovative model for yield forecast in orchards in the Upper valley of Rio Negro and Neuquén. The specific goals are: analyzing and selecting relevant variables to the yield forecast model; adjusting and applying fruit scanning methodologies with proximal sensors; comparing precision and costs in the yield forecast between manual fruit counting and image counting; ultimately, developing model to estimate the size, weight and number of harvest fruit in apple and pear trees.

The trial was carried out in two plots of the National Institute of Agricultural Technology (INTA) Experimental Station and in four commercial orchards. Apple cultivar was ‘Red Chief’ and pear cultivar was ‘William’s Bon Chretien’.

In order to get the number of fruits, images were taken during four seasons from many trees when the pears reached an average diameter of 40 mm and apples 35 mm approximately (first week of December) in two different conditions: i) natural daylight between 10 am and 1 pm, ii) at night with the artificial flash light of the camera. The same day images were captured, the fruits on each photographed tree were manually counted by a single operator. Object detection pre-trained models based on neural networks (YOLO) were used to train artificial vision models to detect and count pear and apple numbers in trees from images. Trained models were evaluated according to recall, precision and F1score; and the correlation between detected and counted fruits was calculated. The precision, recall and F1score achieved by the models were up to 0.86, 0.83 and 0.84, respectively. Also, they could be improved by increasing the amount of input data to the deep learning models. This includes more images taken in many different conditions (angles, distances to the focus, and growth stages of the fruits, among others). Also in different tree size, spacing, age and management conditions. The possibility of estimating the fruit numbers from images could reduce the time spent on

this task, and above all, the costs. This allows growers to increase the number of trees sampled to make yield forecasts. Size estimation from images was made by drawing straight lines on each fruit and using reference elements. Pearson correlation coefficient between the manually measured diameter and the image measured diameter were 0.73 in apples and 0.80 in pears. The estimation of the fruit size from images was not accurate in this trial, varying according to the proximity of the fruits to the known dimension object. It also depends on the position that the fruits take on the tree and how hidden they are by other parts of the plant.

In order to get the fruit size no lineal mixed models were fitted to describe pear and apple growth in the studied area. Many fruits were randomly selected in each orchard and each season and its equatorial diameter was measured every week with an electronic digital caliper. Measurement started 30 days after full bloom (when fruits have reached an average diameter of 10-15 mm) until harvest time. Different models that included fixed and random effects in different levels were fitted with the obtained data. As a result, Red chief apples and William's pears growth in the Upper Valley of Rio negro and Neuquén was described by a logistic curve in its third parameterization. Multilevel no lineal mixed models implementation showed the advantage of discerning and quantifying the sources of stochastic variability at different levels. Fruit was the random effect that showed the biggest incidence. The incorporation of environmental covariates and green indexes improved model performances allowing the variability associated to each site and season to be considered. Although these models were able to describe the fruit growth, they were not regarded as practical if they are not used alongside an appropriate technological package. Therefore, models without green indexes will be used for yield forecast. Predictive capacity indexes were calculated with data not used at the fitted stage of the models. Results were valid, showing mean bias up to one mm in pear and lower in apples; and relative mean bias up to 2%.

The number of trees and sizes of the fruit orchards were obtained from growers. Regarding these data two different forecast were made to be compared with real yield: i) based on information collected from fruit manual or traditional counting; ii) based on information obtained from images. Traditional yield forecast showed relative bias up to 5% in apples and 10% in pear becoming a useful tool to estimate the orchard harvest. Whereas image yield forecast showed relative bias from 3 to 10% in apples and higher in pears (up to 25% in one of the orchards). These can be related to the fact that there is a high difficulty in fruit detection in pear images. In most orchards the traditional forecast was more accurate than the one

obtained from images, however the first one has a higher cost. As bias can be acceptable and improved the image forecast represent a feasible alternative.

Harvest size distribution estimation did not represent well the real ones. These results could be due to the fact that the sample size was not representative or big enough.

These research results suggest the need to continue working to improve the vision artificial algorithm precision to detect and count fruits from images and in the development of more precise methodologies for size estimation (manual or from images). In addition, it could be interesting to develop a technological package that could measure size and number of fruits and also record the geographical coordinates of each measurement to have green indexes information for the yield forecast models.

Key words: pear, apple, image analysis, yield forecast, artificial vision, no lineal mixed models.

1. CAPITULO I: MARCO TEÓRICO

1.1. INTRODUCCION GENERAL

1.1.1. Ubicación y caracterización de la zona en estudio

El alto Valle de Río Negro y Neuquén (Figura 1) representa un área muy extensa que abarca los valles aluviales de los ríos Neuquén inferior, Limay inferior y Negro superior, formando una faja angosta, larga y llana de 130 Km de largo y ancho variable entre 2 y 12 km. Está comprendido entre 38°40' y 39°20' de latitud Sur y entre los 66°50' a 68°20' de longitud Oeste. Los valores de altitud van de 400 a 200 metros sobre el nivel del mar (Rodríguez y Muñoz, 2005).

Los suelos de la región son de origen aluvial, de diferentes clases texturales, escasos en materia orgánica, con velocidad de infiltración media a alta y con un perfil profundo, sin gran desarrollo edáfico (Sánchez et al., 2011). En las cercanías de los ríos Neuquén y Negro, se encuentran los mejores suelos, por lo cual la actividad agrícola se hace más intensiva y continua. Los suelos de esta zona tienen más presencia de limo, mayor capacidad de retención de humedad y fertilidad, y es por lo tanto donde se practica la mayor parte de la actividad productiva (Storti, 2008).

El clima de la región se caracteriza por presentar escasas precipitaciones, con una media anual de 235,5 mm. Los meses más lluviosos son marzo, abril, septiembre y octubre. La temperatura media anual es de 15,5 °C, siendo enero el mes que presenta el mayor valor de temperatura con 24,2 °C y julio el mes con menor valor con 6,7 °C. El periodo más cálido es de noviembre a febrero con una temperatura media mayor a 20 °C, y el periodo más frío es de junio a agosto con valores de temperatura medios entre 6,7 °C y 9,1 °C. El período medio con heladas es de 190 días, en tanto la fecha media de la primera helada es el 14 de abril y la fecha media de la última el 3 de octubre. La región se caracteriza por tener altos valores de radiación global y de insolación. Los valores más altos de heliofanía efectiva corresponden al período estival con una media de 9,8 horas por día para enero y febrero, con una disminución gradual desde fines de verano hasta el mes de junio, para el cual se registran 3,6 horas. (Rodríguez y Muñoz, 2005). Una característica de la zona es la presencia de vientos intensos dominantes del sector oeste sudoeste, con mayor frecuencia en primavera. La velocidad media del viento es entre 18,8 km/h y 15,9 km/h.

La región presenta, por un lado, vegetación natural xerófita, con predominio de plantas bajas, resistentes a sequía, en su mayoría arbustos, arbustos espinosos y pastos duros; y por otra parte, la vegetación implantada por el hombre, ya que se trata de valles irrigados destinados a la producción agrícola, que está compuesta principalmente por frutales de pepita y en menor medida frutales de carozo, forrajes y hortalizas.

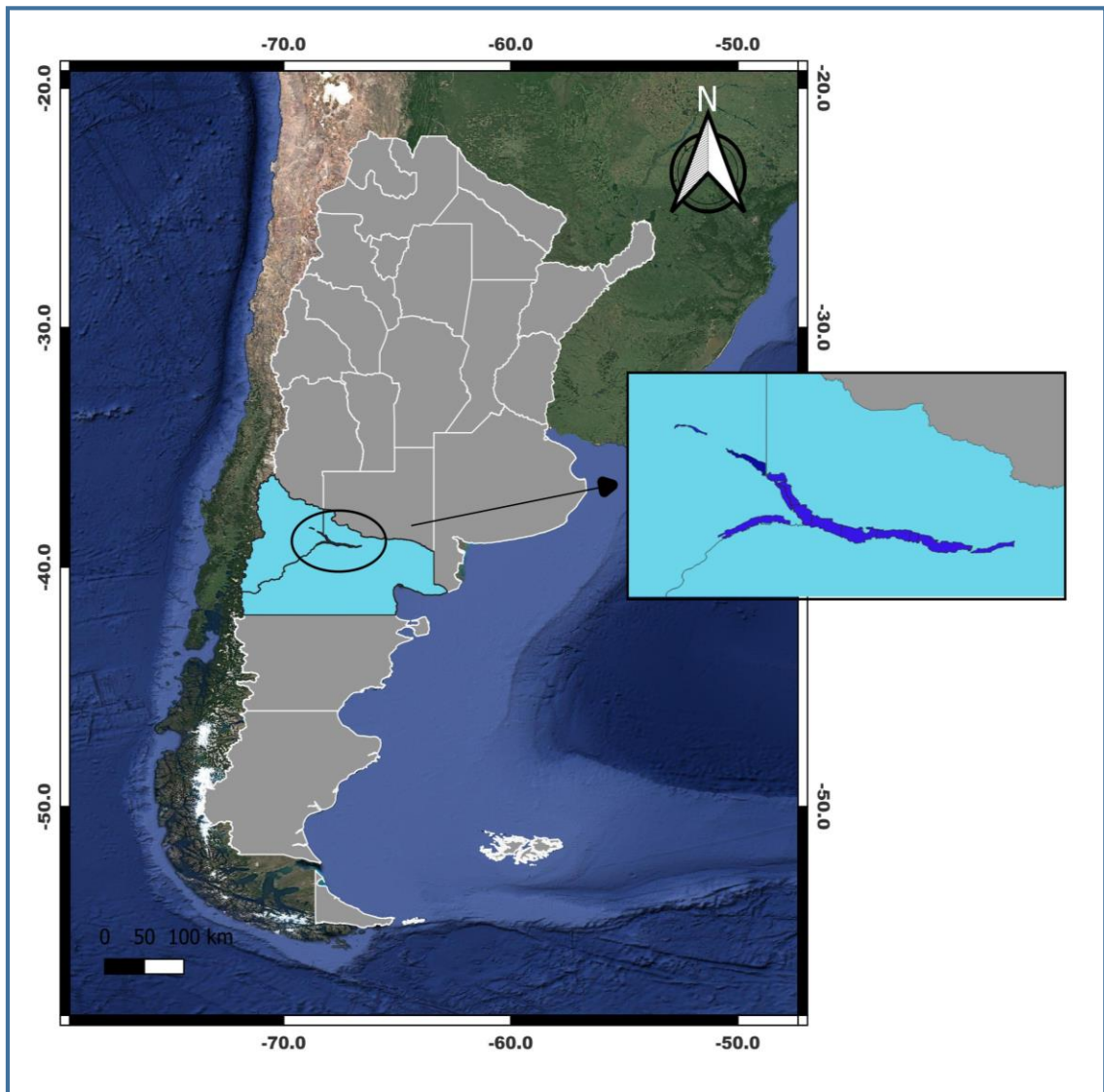


Figura 1: Alto Valle del Río Negro y Neuquén. Elaboración propia.

1.1.2. Producción regional de frutales de pepita

El área productiva de la región del Alto Valle de Río Negro y Neuquén abarca unas 100.000 ha aproximadamente, de las cuales alrededor del 60% se encuentra bajo riego y unas 40.000 ha se cultivan. El área es destinada principalmente a la actividad frutícola, en particular al cultivo de peras y manzanas para la exportación en fresco y la industria de jugos concentrados. También se destaca el cultivo de la vid para la elaboración de vinos y en menor medida frutales de carozo como ciruelas, duraznos y pelones (CPIARN, 2015).

El Alto Valle constituye la principal zona de producción de frutales de pepita de Argentina, representando el 90% del total producido, mientras que Mendoza participa con el diez por ciento de la producción nacional. Existen 18.266 ha implantadas de peras (*Pyrus communis*) distribuidas en aproximadamente en 1.540 establecimientos productivos en todo el Alto Valle, y 17.279 ha de manzana (*Malus domestica*) distribuidas en 1.515 establecimientos productivos (Senasa, 2021). El destino principal de la fruta es el consumo en fresco, tanto para mercado externo como para mercado interno. Disponer de grandes cantidades de fruta de calidad en contraestación de los principales mercados del Hemisferio Norte, permitieron la inserción y permanencia de nuestro país dentro del mercado mundial de peras y manzanas. Argentina es uno de los principales exportadores mundiales de peras frescas alcanzando en el año 2019 unas 304.000 tn exportadas (Senasa, 2021). En manzanas, en cambio, el mercado interno absorbe aproximadamente el 40% de la producción, un alto porcentaje se destina a la industria (45%) especialmente para para la elaboración de jugos y un menor porcentaje va a exportación (Avellá et al. 2018).

De las variedades cultivadas de pera, el 41% aproximadamente corresponde a William's y el 29% a Packham's Triumph. Le siguen Beurre D'Anjou con el 15%, Abate Fetel y Red Bartlett con un 5% cada una. El porcentaje restante incluye las variedades Beurre Bosc, Beurre Giffard, Clapps Favourite y Red Beurre D'Anjou, entre otras (Senasa, 2020).

De las variedades cultivadas de manzana, el 65% aproximadamente corresponde a Red Delicious y sus clones. Luego le siguen Granny Smith, Gala y Cripps Pink con el 13, 12 y 5% del área cultivada, respectivamente (Senasa, 2020).

1.1.3. Características de los cultivos

Los manzanos y perales son especies leñosas de hojas caducas originarias de climas templados. Estos se caracterizan por tener una fase inicial de vida en vivero de dos a tres años para luego ser implantadas en forma definitiva en la parcela comercial, alcanzando generalmente la plena producción recién al quinto o sexto año de vida, dependiendo de la especie, el cultivar y el portainjerto. Las arboles frutales de las parcelas productivas son individuos bimembres conformadas por dos secciones: i) el portainjerto o pie, correspondiente a la parte radicular, a partir del cual se busca control vegetativo, acelerar el comienzo de la producción, y la resistencia a plagas y/o enfermedades del suelo; ii) y la parte aérea conocida como cultivar, que es la parte productiva.

Botánicamente los perales y manzanos corresponden a los géneros y especies *Pyrus communis* y *Malus doméstica*, respectivamente. Ambos pertenecen la familia Rosácea, orden Rosales, subclase Rosidae, clase Dicotiledonea, subdivisión Angiosperma.

1.1.3.1. Pera cv 'William's'

Esta variedad, conocida también como 'William's Bon Chretien' o 'Bartlett' es originaria de Inglaterra y en el Alto Valle es la más cultivada ya que las condiciones climáticas regionales son excepcionales para su producción (Calvo, 2008). Es una variedad estratégica a nivel comercial debido a su maduración temprana y a sus sobresalientes características organolépticas (Calvo et al., 2013).

El árbol es piramidal, de tronco alto, grueso, de corteza agrietada y gris. Florece a fines de septiembre y es una variedad autocompatible. Se cosecha aproximadamente en la segunda quincena del mes de enero cuando el fruto tiene aproximadamente 106 días de edad. Posee un periodo de conservación de tres a seis meses en frío convencional o atmósfera controlada.

El fruto es piriforme con cintura definida, de tamaño mediano a grande (230 a 235 g, calibre ecuatorial de 65 mm y altura 105 mm). El pedúnculo es de tamaño y grosor mediano. El cáliz es semicerrado y la epidermis fina y delicada de color verde claro, tornando al amarillo en su madurez. La pulpa es blancocremosa, de textura fina, aromática, muy jugosa, dulce suavemente acidulada, con muy buena calidad organoléptica (Benítez, 2001).

1.1.3.2. Manzana cv 'Red chief'

Este cultivar es una mutación gemaria de 'Starkrimson', originaria de USA. También se la conoce como 'Chañar 90' y en muchos lugares se la toma como referente de las manzanas del grupo Delicious. Es un árbol dardífero, de poco vigor, compacto y muy productivo. Florece a fines de septiembre, es autoincompatible, se poliniza con 'Granny Smith', 'Golden Delicious', 'Rome Beauty' o 'Gala'. Se cosecha aproximadamente en la segunda quincena del mes de febrero cuando el fruto tiene entre 140 a 145 días de edad.

El fruto es de forma tronco-cónica, con cinco lóbulos o protuberancias bien marcados. El tamaño es medio y el color de la epidermis precoz, rojo intenso brillante y ligeramente estriado. La pulpa es blancoamarillenta, de textura finamente granulada, jugosa, algo perfumada, dulce, de buen sabor y con tendencia a la pérdida de firmeza y a la harinosidad bajo ciertas condiciones (Benítez, 2001).

1.1.4. Aspectos agronómicos de los cultivos

Los montes frutales deben recibir una serie de cuidados o manejos agronómicos específicos durante cada temporada productiva para lograr una óptima cosecha. Entre las principales tareas se pueden mencionar:

- ✓ Riego: la producción de frutales de pepita en el Alto Valle es posible gracias a los sistemas de riego. Se utilizan diferentes tipos como el gravitacional, aspersion o goteo; pero en todos los casos se pretende suplir las necesidades del cultivo. Esta última se conoce como evapotranspiración y es la combinación de dos procesos: evaporación y transpiración (Adaro et al., 2010).
- ✓ Conducción y poda: estas labores se realizan desde el año de plantación para lograr una producción lo más temprana posible, y a su vez continua del monte frutal. El tipo de conducción más utilizado en la actualidad a nivel regional es la espaldera. La poda, por su parte, puede seguir diferentes objetivos: de plantación (rebaje, eliminación de competencia para definición de ejes principales, etc.); de fructificación (eliminación de ramas fuertes o chupones, favorecer entrada de luz y aire, etc.).
- ✓ Raleo de frutos: es una práctica de vital importancia para obtener fruta de buena calidad, y consiste en reducir la carga frutal en un período temprano de ciclo productivo para lograr un mayor tamaño final de los frutos y un óptimo retorno a floración del año siguiente,

controlando el añerismo característico de estas especies. Esta tarea se puede realizar en forma química o manual.

✓ Fertilización: el contenido de materia orgánica de los suelos del Alto Valle generalmente es menor al uno por ciento en los primeros centímetros del suelo y la fruticultura consume un alto contenido de la misma, por lo cual es necesario el aporte anual de biomasa al suelo (Adaro et al., 2010). También son muy comunes las fertilizaciones foliares con macro elementos (nitrógeno, fosforo y potasio) y micro elementos (boro y calcio). Un adecuado equilibrio nutricional de la planta asegura una buena calidad de las yemas de flor (Rodríguez, 2011).

✓ Manejo de plagas: el control de insectos y enfermedades es esencial para llevar adelante la producción. La plaga principal de los frutales de pepita es la carpocapsa (*Cydia pomonella*), y las plagas secundarias incluyen pulgones o áfidos, cochinillas y ácaros, entre otras.

✓ Control de adversidades climáticas: se realiza monitoreo de heladas primaverales y control pasivo o activo de las mismas, se controla el efecto de los fuertes vientos mediante el uso de cortinas rompivientos, y también se intenta controlar el efecto adverso de las altas temperaturas con mallas o protectores solares.

Con todo esto se pretende obtener al final de cada temporada de crecimiento una cosecha adecuada en términos de cantidad de frutos y de tamaño final de los mismos. Sin embargo, el tamaño final de los frutos no solo depende de la correcta labor de dichas tareas, sino que también puede verse afectado por factores intrínsecos, tales como:

✓ Relación hoja/fruto: la disponibilidad de hidratos de carbono depende de la superficie foliar y la intensidad de fotosíntesis. Una relación normal para obtener una buena calidad y cantidad de cosecha es de 50 hojas por fruto (Adaro et al., 2010).

✓ Estado nutricional de las yemas: la energía inicial para el crecimiento tanto vegetativo como de floración proviene de las reservas de la planta; si estas son escasas, la división celular de los frutos también lo será, limitando el crecimiento potencial de los frutos.

✓ Ubicación de las estructuras fructíferas en el árbol: de esto dependerá la disponibilidad de luz y de foto asimilados; en general, frutos de la parte baja del árbol (más sombreados), son más chicos.

✓ Número de semillas: el tamaño del fruto es proporcional al número de semillas viables. Estas son fuente de diferentes fitohormonas que favorecen el crecimiento.

✓ Carga frutal: debe regularse con la práctica de raleo

✓ Tamaño inicial del fruto: el número de células de un fruto queda definido al inicio del crecimiento y determinará el tamaño potencial que podrá alcanzar. El número de células a su vez queda definido por las condiciones térmicas durante la fase de división celular (Stanley et al., 2000).

1.1.5. Crecimiento de frutos de pepita

El crecimiento de los frutos sigue un modelo representado por una curva sigmoidea (Figura 2). Este tipo de modelo se caracteriza por presentar una asíntota superior que indica el tamaño máximo que puede alcanzar el fruto, un punto de inflexión donde la velocidad del proceso comienza a decrecer y una asíntota inferior (Álvarez y Boché, 1999).

En términos de crecimiento se pueden distinguir tres fases:

- *fase I* o de crecimiento exponencial en la cual se produce la multiplicación celular;
- *fase II* o de crecimiento lineal en la cual se produce el agrandamiento y expansión celular;
- *fase III* o etapa final donde se produce el crecimiento del fruto y su maduración.

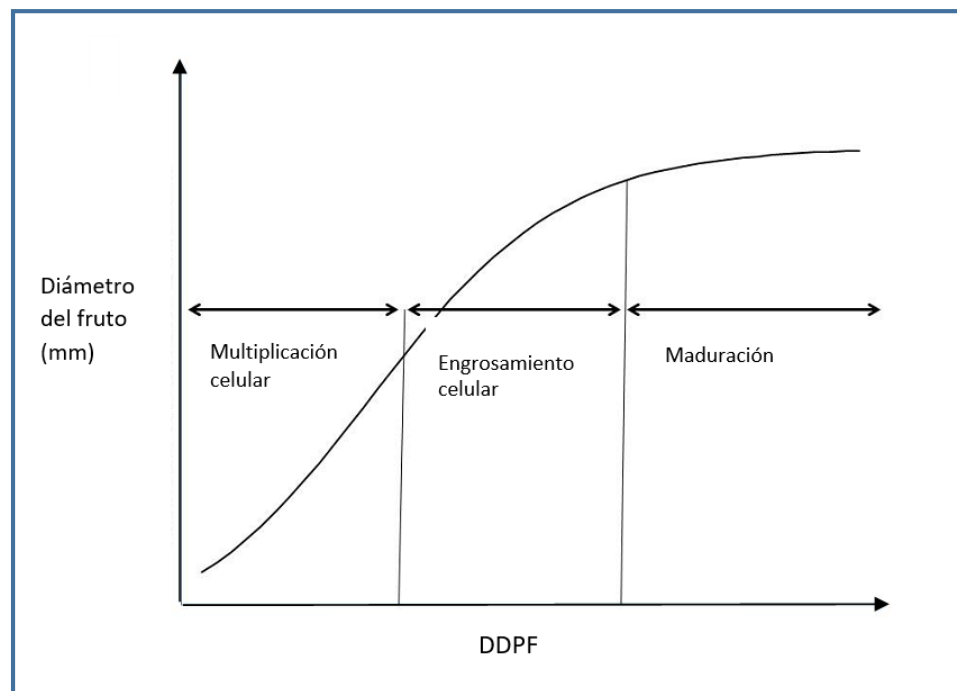


Figura 2: Curva doble sigmoidea. Elaboración propia.

La fase I tiene una duración relativamente constante en cada especie, en manzanas particularmente el periodo de división celular dura entre 21 y 30 días después de plena

floración (DDPF) y en peras se puede extender hasta dos meses (Sozzi, 2007). La fase de elongación celular consiste en la acumulación de agua y sustancias hidrocarbonadas en las células, originando el aumento de volumen y peso del fruto y se puede extender de 30 a 90 días (Rodríguez, 2011), variando en función de la especie y cultivar del que se trate. En pera cv William's dicha fase tiene una duración de 65 a 70 días (Adaro et al., 2010). La última fase, que corresponde al proceso de maduración, dura de diez a 30 días aproximadamente (Adaro et al., 2010).

1.2. PLANTEO DEL PROBLEMA

Los cambios sustanciales que están ocurriendo actualmente en nuestra cultura y forma de vida al introducirnos en la era del post-industrialismo, exigen estar preparados para un mundo donde la tecnología, el conocimiento y la sustentabilidad marcarán la agenda en materia de desarrollo. La fruticultura en particular no es ajena a estas tendencias, ya que es una actividad intensiva, con un gran aporte de capital, tecnología, conocimiento y mano de obra. Estos factores de desarrollo, marcarán las pequeñas diferencias que permitirán mantener e incrementar su competitividad. El desafío de la fruticultura moderna es la incorporación de tecnología con el fin de dar un salto cualitativo y agregar valor a la producción a través del conocimiento, la innovación y el desarrollo tecnológico. En este sentido, un aspecto muy importante a considerar en la búsqueda de nuevas herramientas es el desarrollo de modelos de predicción de la producción que permitan mejorar la competitividad de las empresas, tanto locales como a nivel nacional. Esta mejora se producirá fundamentalmente por el conocimiento "a priori" del comportamiento productivo de los cultivos de peras y manzanas, lo que permitirá la organización logística de la mano de obra, la compra anticipada de insumos, la planificación de los empaques, disposición de los sistemas de transporte, organización del almacenamiento en frío y planificación de la venta de la producción.

Se ha realizado durante varios años en la región un pronóstico de producción en base a conteos y mediciones manuales de frutos, considerando variables climáticas y otros aspectos relevantes de cada temporada, el cual provee información del potencial productivo regional pero no aporta información a nivel predial o de un productor en particular. Es decir, no existe

una metodología validada localmente para estimar el rendimiento a nivel de parcela. Además, el conteo manual de frutos en especies arbóreas es difícil de realizar, poco preciso y requiere de mucho tiempo, por lo cual resulta muy interesante encontrar alternativas más sencillas y que permitan obtener resultados confiables. Por este motivo se pretende utilizar la información actualmente disponible en la región (datos climáticos, datos de fenología, modelos de crecimiento de frutos, prácticas culturales, etc.) sumada al uso de tecnología de captura de datos mediante sensores para desarrollar un modelo de predicción del número, peso y tamaño de los frutos al momento de cosecha para peras, como así también adaptar los modelos ya existentes para manzanas a las condiciones locales.

1.3. PRONOSTICO DE PRODUCCION

La producción de frutas de peras y manzanas es una actividad de alto riesgo, lo cual, junto a la gran interdependencia de los mercados, genera la necesidad de contar con información anticipada y confiable sobre la producción y los precios que alcanzará la misma. Por ello, la planificación se convierte en una tarea indispensable. Sin embargo, la producción se ve afectada por numerosos factores internos y externos que hacen muy difícil cualquier intento de planificación (Avanza, 2010). En este sentido la predicción de volúmenes de fruta a obtener en cada temporada es una herramienta muy útil para definir recomendaciones de manejo del cultivo y estrategias de comercialización (Bergh, 1982).

La cosecha es una de las tareas más intensivas de la producción de peras y manzanas y se debe realizar en el momento en que la fruta posee la madurez apropiada para ser comercializada o almacenada. Es decir, debe cumplir con ciertos requerimientos óptimos de calidad, tamaño y madurez.

El periodo de tiempo en el cual la fruta presenta estos requerimientos es acotado, por lo cual las tareas de cosecha deben ser cuidadosamente planificadas y en este sentido la predicción de volúmenes es fundamental para una correcta organización de la misma. Contar con una predicción temprana de los rendimientos permite a los productores estimar el volumen de frutos a cosechar, calcular la cantidad de insumos necesarios para la recolección de la fruta,

planificar la mano de obra necesaria, organizar y administrar los espacios de almacenamiento, el empaque y el transporte (Zhou et al., 2012).

Sin embargo, la estimación de los rendimientos en cultivos frutales no es una tarea sencilla. Por un lado, los árboles frutales son especies perennes, en las cuales las prácticas culturales y los diferentes cultivares implican una gran variabilidad en la producción obtenida año tras año (Avanza, 2010). Por otra parte, los rendimientos dependen en gran medida de las condiciones meteorológicas del año que afectan tanto la floración como el desarrollo de los frutos y su permanencia en la planta hasta el momento de la cosecha (Aruani et al., 2014).

A nivel regional, hasta el momento, todo programa de predicción anticipada de volúmenes de producción en montes frutales se ha basado en mediciones objetivas a campo en las que es necesario conocer tres elementos esenciales:

- i) carga de los arboles al momento del pronóstico expresado en número de frutos por árbol;
- ii) tamaño a cosecha que se espera tendrán los frutos;
- iii) la cantidad de plantas o superficie plantada en el área en estudio.

Las diversas estrategias propuestas para realizar estas predicciones difieren esencialmente en la forma en que se estiman o predicen estos parámetros.

Como se mencionó previamente, se ha realizado un Pronóstico de Producción regional por más de 23 años, que, si bien provee información del potencial productivo regional y es un sistema muy útil para la toma de decisiones a nivel macroeconómico y de logística en general, no aporta suficiente información a nivel predial o de un productor en particular. Dicho pronóstico, a partir del año 2015, dejó de realizarse.

El hecho de conocer acabadamente la producción por lote o parcela, permitiría ajustar con mayor precisión la logística y la compra de insumos en la cosecha de cada explotación, haciendo más eficiente consecuentemente esta etapa de producción. El dato de rendimiento debería ser obtenido preferentemente con una antelación de 30 a 45 días previos a la cosecha, para que tenga la utilidad pretendida.

1.4. OBJETIVOS

1.4.1. Objetivo general

Implementar un modelo predictivo del rendimiento parcelario a cosecha en frutales de pepita, adaptado a las condiciones del Alto Valle de Río Negro y Neuquén.

1.4.2. Objetivos específicos

- i) Analizar y seleccionar variables relevantes para el modelo de predicción de cosecha.
- ii) Adaptar y aplicar metodologías de escaneo de frutos en base a sensores proximales.
- iii) Comparar calidad de predicción y costos entre metodologías que impliquen conteo de fruto manual y técnicas de conteo con sensores.
- iv) Desarrollar un modelo de predicción de tamaño, peso y número de frutos (rendimiento) a cosecha en manzanos y perales.

1.5. HIPÓTESIS DE TRABAJO

- i) Es posible contar y medir el tamaño de los frutos de pepita en tiempo real, a partir de sensores proximales.
- ii) Se puede predecir anticipadamente el tamaño, peso y cantidad de frutos producidos en una parcela de frutales de pepita a partir del desarrollo de un modelo de predicción que incluya variables climáticas, de manejo del cultivo y tecnología de sensores.
- iii) Las metodologías de conteo de frutos con sensores son precisas y menos costosas que el conteo manual.

2.CAPITULO II: DETECCIÓN Y CONTEO DE FRUTOS

2.1. INTRODUCCION

2.1.1. Estimación de la carga frutal

Como se dijo en el capítulo anterior, uno de los tres elementos que se tiene que conocer para poder realizar predicciones de la producción es la carga frutal. Para estimar el número de frutos por árbol o la carga de los arboles al momento del pronóstico existen diferentes metodologías de conteo manual de frutos, la mayoría de las cuales no han demostrado ser prácticas ni lo suficientemente precisas (Avanza, 2010; Wulfsohn, 2012). Esto se debe a que requieren de mucha capacitación de los operarios y son mano de obra intensiva, dos aspectos altamente limitantes en los cultivos frutícolas.

Por otra parte, cada empresa o productor frutícola utiliza su propio método de estimación, muchas veces sin un sustento metodológico apropiado, lo que conlleva grandes esfuerzos de muestreo y con resultados frecuentemente muy pobres. Además, el número de árboles inspeccionados generalmente es bajo, lo cual conduce a resultados poco precisos debido a la gran variabilidad que suele existir en los montes frutales (Reeb et al., 2003; Stajniko et al., 2009). En este sentido la incorporación del uso de imágenes y su análisis, surge como una alternativa muy interesante como reemplazo a los métodos tradicionales o manuales de conteo, con resultados muy prometedores (Maheswari et al., 2021).

2.1.2. Incorporación de Análisis de imágenes

Si bien el uso de imágenes en agricultura es extenso, en fruticultura, en particular, ha aumentado de manera considerable en la última década. Para la obtención de las imágenes existen diferentes tipos de sensores a utilizar según los objetivos perseguidos (Gongal et al., 2015). Un sensor de imagen puede definirse básicamente como un dispositivo eléctrico que convierte magnitudes físicas (luz) en señales eléctricas (Vargas Baeza, 2020). Es decir, el sensor de imagen de una cámara digital, convierte la luz que proviene del sujeto en una señal electrónica que finalmente se transforma en un archivo de imagen digital. En función de la distancia que existe entre el sensor y el objetivo, éstos se pueden clasificar en proximales o remotos (Velez et al., 2016). Mientras que los primeros requieren cierto contacto con el objetivo, los últimos captan información de los objetos a distancia. En esta tesis se utilizará

información proveniente de sensores proximales (cámaras digitales); y sensores remotos (satélites). Como ventajas de los primeros se puede mencionar que en general poseen mayor resolución espacial, son tecnologías más simples, brindan la posibilidad de usar varios sensores al mismo tiempo. Como desventajas se pueden mencionar la necesidad de ir al terreno en busca de la información, que las mediciones son a pequeña escala y que requieren mayor tiempo de mano de obra.

Por su parte los sensores remotos miden a gran escala y hay información disponible sin costo, pero normalmente poseen menor resolución espacial, y son más dependientes de las condiciones ambientales. La percepción remota o teledetección, denominada en inglés '*remote sensing*' hace referencia a la adquisición de datos de la superficie terrestre con un sensor remoto, y al procesamiento e interpretación de esos datos.

Durante las últimas décadas se ha producido un importante avance de la teledetección con el desarrollo de sensores con sistemas de barrido multiespectral, que permiten registrar información en formato digital de muchas regiones del espectro electromagnético. Con este avance también surge el incremento en el interés de utilizar esta información para diferentes aplicaciones, y entre ellas las agronómicas.

En fruticultura en particular, se ha utilizado información proveniente de sensores remotos en el desarrollo de índices no destructivos para la medición de degradación de clorofila en frutos de manzana (Zude, 2003); en el desarrollo de modelos de predicción de cosecha de cítricos basados en el tamaño de copa y en el índice de vegetación de dos bandas (Ye et al., 2008); en la determinación de factores causantes de variabilidad espacial del rendimiento de perales a partir del tamaño de los árboles, el vigor (tomado del NDVI) y las propiedades de los suelos (Perry et al., 2010); entre otros.

Por su parte, el uso de información proveniente de sensores proximales también ha resultado de gran interés en fruticultura. Los sensores laser, si bien no se utilizan tan frecuentemente como las cámaras digitales, se han empleado en el desarrollo de inventarios y estimación de rendimientos en función de índices verdes calculados a partir de imágenes multiespectrales y de información topográfica (Viau et al., 2005). También se los ha utilizado para hacer mapeos de la variabilidad del área foliar de viñedos (del Moral Martínez et al., 2016) y para obtener volumen de copa de árboles de olivos (Escola et al., 2009); entre otros.

Por otro lado, el uso de cámaras digitales, en general tiene como principal objetivo la detección automática de frutos a partir de imágenes con diferentes propósitos. Entre ellos se puede destacar la predicción de los rendimientos de montes frutales (Stajanko et al., 2009; Aggelopoulou et al., 2010; Cohen et al., 2011; Zhou et al., 2012; Dorj et al., 2017), el ajuste

en la toma de decisiones de raleo químico (Hočevár et al., 2014; Farjón et al., 2019), la cosecha mecanizada o robotizada (Mao et al., 2009; Wachs et al., 2010).

Se han utilizado diversos tipos de sensores proximales a los fines de lograr visión por computador y detectar frutos automáticamente, desde cámaras en blanco y negro, por ejemplo, para evaluar la viabilidad de un sistema automatizado de recolección de manzanas (Parrish y Goksel, 1977), hasta cámaras multiespectrales, para estimar producción acumulada, peso medio y contenido de sólidos solubles en manzanas (Best et al., 2008.). Si bien estas últimas son equipamientos de muy alto costo (Kurtulmus et al., 2014), ofrecen imágenes con numerosas longitudes de onda simultáneamente, es decir, abarcan un amplio espectro y de allí su nombre. Una cámara multiespectral tiene el potencial de detectar frutos aun cuando el ambiente que los rodea es muy similar a ellos (Henten et al., 2002; Kane, 2006; Bulanon y Kataoka, 2010; Wang et al., 2012; Gongal et al., 2015), ya que la información que ofrece es muy detallada. Sin embargo, gran parte de las metodologías desarrolladas hasta el momento, basan sus análisis en imágenes obtenidas de cámaras digitales a color, también llamadas imágenes RGB (por las siglas de los colores ‘Red’, ‘Green’, ‘Blue’). Estas imágenes proveen información del color, textura y forma de los objetos presentes en la fotografía y tienen como ventaja que son tecnologías de bajo costo y sencillas de utilizar para los usuarios.

2.1.3. Procesamiento y análisis de imágenes

No solo existen diferentes sensores para obtener imágenes, sino que también existen diversas estrategias que se pueden abordar para el procesamiento y análisis de las mismas, abarcando desde las más sencillas como el uso de rasgos de color, hasta métodos mucho más complejos y modernos como el aprendizaje automático o ‘*machine learning*’ (ML). Es importante destacar que los conceptos de procesamiento y análisis de imágenes, si bien resultan similares, tienen diferentes implicancias. El primero hace referencia a la elaboración, transformación, filtración, entre otros, que se realizan sobre las imágenes, mientras que el segundo implica las mediciones y cuantificaciones de los objetos presentes en las imágenes al finalizar el procesamiento, es decir, se obtiene información numérica que puede ser interpretada de manera matemática y estadística (Potriansky, 2013).

Cuando se trabaja con técnicas de procesamiento y análisis de imágenes para la detección de objetos existen cuatro rasgos básicos que los caracterizan en las imágenes: el color, la intensidad, la forma y la textura. Todos ellos pueden ser utilizados para reconocer objetos como flores o frutos a partir de una imagen. Uno de los más utilizados para detectar frutos en

imágenes de plantas frutales es el color. Sin embargo, cuando la detección está basada en este rasgo, su eficacia puede verse afectada por variaciones en el mismo, en la madurez o variedad de los frutos y por cambios en el ambiente que los rodea (Gongal et al., 2015). En este sentido, el análisis a partir de los rasgos de forma se presenta como más robusto mientras que el análisis a partir de color e intensidad puede ser más veloz (Mao et al., 2009).

Se han desarrollado algoritmos de detección de frutos solo a partir de rasgos de color logrando buenos resultados, por ejemplo, en arboles de manzanas cv ‘Gala’ (Zhou et al., 2012), manzanas cv ‘Fuji’ (Bulanon et al., 2002; Mao et al., 2009), mandarinas (Dorj et al., 2017), ‘Orlando’ tangelos (Annamalai, 2003).

Otros algoritmos de detección más complejos integran rasgos de color, textura y forma. En arboles de manzana cv ‘Golden Delicious’ y ‘Jonagold’, se desarrolló un algoritmo de múltiples pasos para detectar y contar frutos con propósitos de pronóstico, a partir de segmentaciones y transformaciones de las imágenes (Stajnko et al., 2009). Además, en ‘Golden Delicious’ se desarrolló un algoritmo para estimar el número de manzanas maduras en imágenes capturadas de noche utilizando operaciones de umbralización de las imágenes y detección de círculos concéntricos, logrando errores menores al 6% entre los frutos reales y los estimados (Linker, 2016). En cítricos se logró detectar frutos inmaduros a partir un algoritmo que utiliza componentes de intensidad y saturación de color y análisis de textura, detectando aproximadamente el 75% de los frutos presentes en el árbol (Kurtumulus et al., 2011). En mangos también se han propuesto algoritmos de detección de frutos en arboles a partir del análisis de rasgos de textura y de forma de las imágenes (Qureshi et al. 2016).

También se han desarrollado algoritmos de detección de flores en manzana cv ‘Golden Delicious’ utilizando umbrales para los parámetros de la transformación del espacio de color: HSL (siglas de ‘*hue*’, ‘*saturation*’, ‘*luminance*’, es decir matiz, luminosidad y saturación de color) y algunos parámetros de optimización. Las correlaciones obtenidas entre las flores estimadas y contadas a campo fueron menores a 0,6 (Hočevan et al., 2014).

Mayoritariamente todos estos algoritmos de detección de objetos a partir de imágenes implican los siguientes pasos: captura de la imagen, preprocesamiento, segmentación, remoción del ruido, operaciones morfológicas, clasificación y extracción de rasgos (Gongal et al., 2015), los cuales se pueden ver esquematizados en la Figura 3.

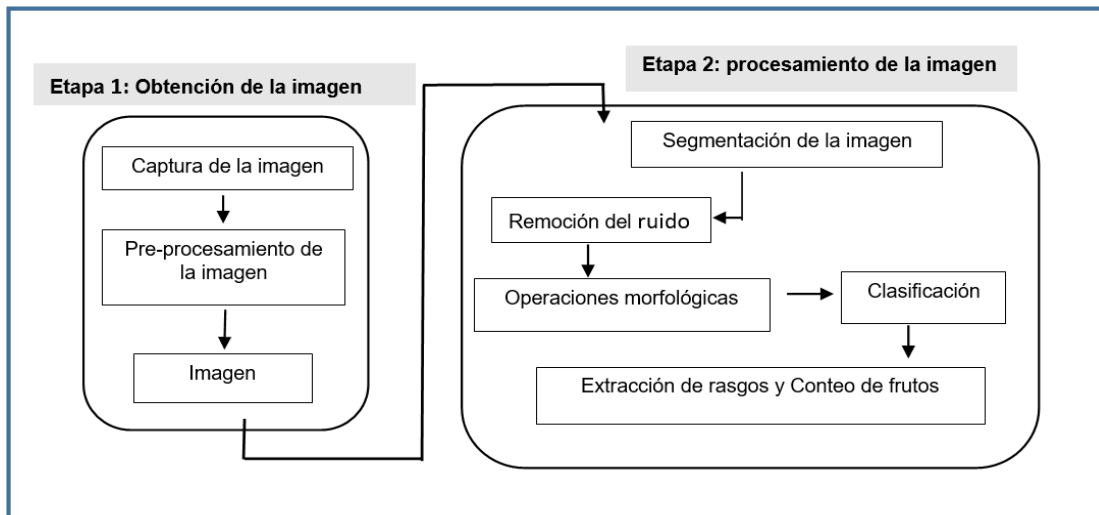


Figura 3: Diagrama de flujo del análisis de imágenes. Elaboración propia.

Estas técnicas utilizan métodos clásicos de programación, en los cuales la lógica es brindar al equipo los datos y las reglas (por ejemplo, imágenes y umbrales de color, respectivamente) y, en función de éstos, esperar una respuesta (por ejemplo, fruto/no fruto).

Sin embargo, las técnicas más prometedoras actualmente en análisis de imágenes y, particularmente, en detección de objetos, pertenecen a la Inteligencia Artificial (IA) o Machine Learning (ML), en las cuales surge otra lógica diferente de programación. En estos casos, el investigador brinda los datos de entrada y las respuestas que espera obtener de estos datos (por ejemplo, las imágenes y los frutos señalados dentro de las imágenes) y espera que la computadora genere sus propias reglas (Chollet, 2018). En la agricultura de precisión el ML es una de las técnicas más utilizadas para la toma de decisiones (Maheswari et al., 2021).

Algunos autores sugieren que los métodos de ML basados en IA logran mejores resultados que los métodos clásicos o simples de programación utilizados para analizar imágenes, aunque requieren mayores recursos computacionales (Gongal et al., 2015). Según Koirala et al., (2019b) esto último ha dejado de ser un inconveniente en los últimos años, en los cuales es común disponer de una unidad de procesamiento de gráficos (GPU), sumado a que existen herramientas de anotaciones gráficas de uso libre (como LabelImg <https://github.com/tzutalin/labelImg>) y técnicas de aprendizaje profundo de última generación como Faster R-CNN (Faster Regional Convolutional Neural Network) (Ren et al., 2015) o YOLO (You Only Look Once) (Redmon y Farhadi, 2017), entre otros.

2.1.4. Inteligencia artificial, aprendizaje automático y aprendizaje profundo

Como se dijo previamente las técnicas de ML o aprendizaje automático se engloban dentro de lo que se denomina IA, pero con un nuevo paradigma de programación (Figura 4). El ML puede definirse como la ciencia o el arte de programar computadoras para que éstas puedan aprender a partir de los datos (Gerón, 2005). Toda técnica de ML incluye tres elementos básicos: datos de entrada, ejemplos de respuestas esperadas a partir de dichos datos, y alguna forma de medir cuan bien funciona el algoritmo.

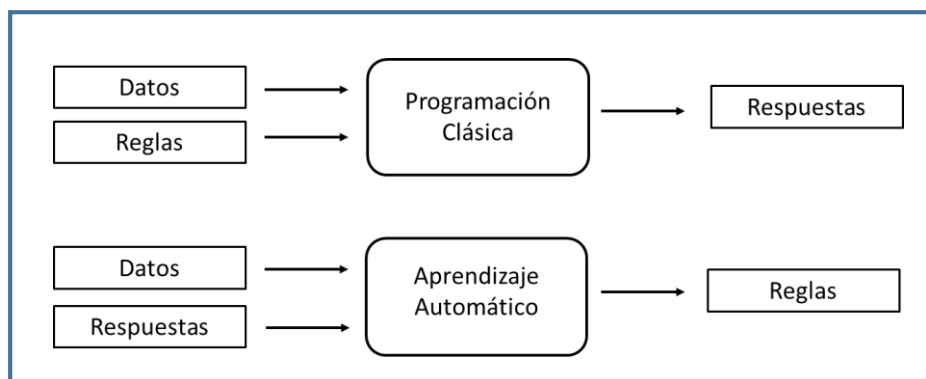


Figura 4: ML: nuevo paradigma de programación. (Extraído de Chollet, 2018)

Es decir, durante el procesamiento de los datos en una técnica de ML se generan reglas que luego pueden ser aplicadas a nuevos conjuntos de datos, y brindar nuevas respuestas. No se dan ordenes explícitas de programación, sino que los algoritmos se van entrenando solos. Entre las técnicas de aprendizaje automático más destacadas, están las redes neuronales artificiales (RNA), que si bien existen hace más de 70 años, en las últimas dos décadas han evolucionado notoriamente (junto a la disponibilidad de mayor capacidad tecnológica de procesamiento, memoria, disco, nube, etc.), logrando cada vez mejores resultados.

El denominado ‘deep learning’ o aprendizaje profundo (DL) para el procesamiento de la información, usa RNA pero concatenando un gran número de capas formadas por unidades de procesamiento sencillas (neuronas), que van transformando progresivamente la información (Rozada Raneros, 2021). Estas RNA imitan comportamiento biológico del cerebro, trabajan con tres tipos de capas de neuronas: de entrada, ocultas y de salida (Figura 5) y son muy utilizadas para el análisis de grandes bases de datos (Bagnato, 2020). Dichas técnicas trabajan con los datos en la búsqueda de diferentes representaciones de los mismos,

que resulten útiles dentro de un espacio predeterminado de posibilidades y utilizan una señal de retroalimentación como guía. Estas representaciones generalmente se denominan ‘Layers’ o capas y utilizan modelos de aprendizaje denominados ‘neural networks’ (NN). Cada capa o representación es parametrizada mediante una medida de peso y para saber cómo funciona la técnica, y que tan aproximados son los resultados a la realidad, se utiliza la llamada función de pérdida. Esta última calcula la distancia entre las predicciones y los datos reales, y es utilizada como señal de retroalimentación para ir ajustando de a poco los pesos de forma tal que la función perdida vaya decreciendo (Chollet, 2018).

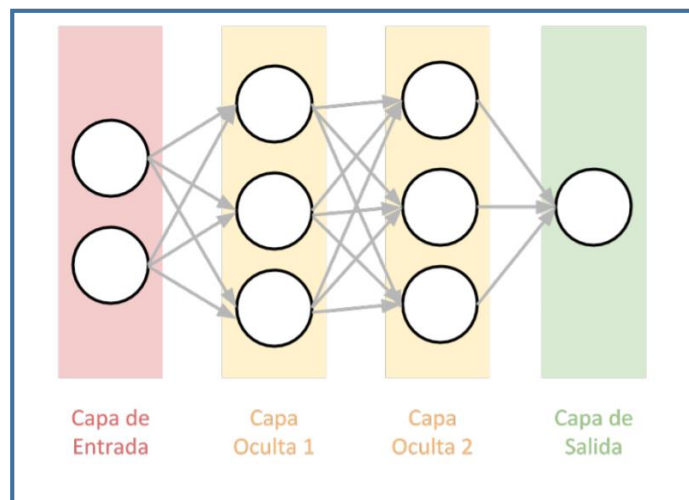


Figura 5: Esquema de capas en una RNA. (Extraído de Bagnato, 2020).

En la detección de objetos a partir de análisis y procesamiento de imágenes se utilizan principalmente redes neuronales convolucionales (RNC) (Rozada Raneros, 2021). Las RNA tradicionales en sus capas completamente conectadas presentan el problema de no escalar adecuadamente los parámetros sobre todo cuando las imágenes son grandes, lo cual provoca sobreajustes. Por el contrario, las RNC toman como premisa que trabajan con imágenes y restringen su arquitectura para resolver este inconveniente. Las neuronas en las capas de las RNC están dispuestas en tres dimensiones: ancho alto y profundidad; y en la salida las imágenes completas de la red se reducen a un solo vector de puntuaciones de clase. Las capas de entrada contienen toda la imagen, es decir, un volumen igual a las dimensiones de la imagen; las capas convolucionales calculan la salida de las regiones conectadas localmente en la entrada de la capa aplicando un producto escalar entre sus pesos y la región a la que están

conectados; y las capas de salida son completamente conectadas y es donde se calcula el puntaje de clase (Escola, 2017).

En manzanas cv Braeburn' y 'Golden Delicious', Črtomir et al., (2012) han utilizado RNA para realizar predicciones de rendimientos a partir de imágenes. Los autores desarrollan una metodología híbrida que incluye un algoritmo para calcular el número de frutos presentes en la imagen, y las RNA para modelar los rendimientos con resultados muy alentadores. Este trabajo, sin embargo, utiliza las técnicas de ML en la etapa de modelación de los rendimientos, no así en la detección y conteo de frutos. Contrariamente, en los últimos años se han publicado varios trabajos que utilizan modelos pre-entrenados de ML para detectar y contar frutos a partir de imágenes logrando altas performances (Chen et al., 2016; Sa et al., 2016; Bargoti y Underwood, 2017; Liang et al., 2018; Bresilla et al., 2019; Tian et al., 2019; Itakura et al., 2021; Wang y He, 2021).

2.1.5. Modelos pre-entrenados de ML para la detección de objetos:

El entrenamiento de una técnica de DL involucra la optimización de una función de pérdida para lograr un error mínimo en la localización y clasificación de los objetos. Como sus arquitecturas poseen miles de parámetros a optimizar, hay sets de datos de libre acceso disponibles para los entrenamientos de modelos y también hay modelos pre-entrenados disponibles para ser utilizados en nuevas aplicaciones, los cuales en esta tesis se mencionarán como modelos pre-entrenados de detección de objetos (MPDO). Estos son utilizados en diversos estudios y el énfasis de los mismos generalmente radica en poder realizar recomendaciones de aspectos prácticos para adoptar estas técnicas estándar, y también en hacer contribuciones originales como el uso de métricas de comparación de modelos (Koirala et al., 2019b).

Entre los MPDO desarrollados para la detección de objetos, podemos mencionar como los más relevantes a R-CNN (Girshick et al., 2014), el cual usa el concepto de '*Region Proposal Network*' o región de interés (RI). Este modelo lo primero que hace es determinar las RI dentro de la imagen, es decir, de una manera selectiva determinar dónde está el objeto. Luego realiza una clasificación sobre las RI mediante un clasificador '*support vector machine*' (SVM) para definir si la RI se clasifica como objeto de interés o no. Esto implica varios algoritmos y dos etapas, una en la que se generan las potenciales RI en la imagen y una segunda que usa un clasificador para que funcione en dichas RI. Después de la clasificación se usa un procesamiento posterior para afinar las ventanas de detección y puntuar de nuevo el objeto de

interés en base a otros que aparecen en la imagen y eliminar ventanas duplicadas. Cuando hay muchas RI el método se vuelve lento, siendo su principal desventaja.

En la búsqueda de mejorar este problema de velocidad surgen dos nuevos MPDO, Fast R-CNN (Girshick et al., 2015), que mejora un poco el algoritmo inicial en tiempo de entrenamiento y detección por el uso de una sola capa o *layer* para el paso de la clasificación en lugar de utilizar el SVM. Sin embargo, la búsqueda de las RI sigue siendo el cuello de botella de la técnica; y Faster R-CNN (Ren et al., 2015) que posee un método de búsqueda de RI que usa RN y cajas de anclaje para incrementar la velocidad. Desliza una venta de tres por tres a través del mapa final y en cada ventana considera k cajas de anclaje diferentes, centradas en la localización para generar las RI (Koirala et al. 2019b). Estas regiones propuestas como RI son luego filtradas según una puntuación (probabilidad de que dicha RI contenga el objeto) y pasa a la siguiente etapa (que esencialmente es el Fast R-CNN) para la detección del objeto. Faster R-CNN posee dos alternativas: ZFNet y VGGNet, cuya diferencia radica en la cantidad de capas que utiliza cada una (7 y 16 respectivamente). Sin embargo, a pesar de la mejora en la velocidad de detección lograda, sigue siendo una técnica demasiado lenta para aplicar, por ejemplo, en videos en tiempo real (Koirala et al., 2019b).

Las MPDO mencionadas anteriormente trabajan en dos etapas, una que localiza el objeto de interés y otra que lo clasifica. Sin embargo, hay otras técnicas trabajan en una sola etapa (es decir, clasifican y localizan los objetos en un mismo paso) como el Single Shot MultiBox Detector (SSD) (Liu et al., 2016) que predice simultáneamente la clase de objeto y el cuadro delimitador con la localización del mismo en la imagen haciéndolo más rápido que el Faster R-CNN. SSD alimenta la imagen de entrada a través de una serie de capas y genera diferentes mapas de características de la imagen a distintas escalas. En cada mapa se usa un filtro (de tres por tres) para evaluar las cajas de anclaje predeterminadas que luego son clasificadas en cada posición de la imagen y en cada escala. Se genera un amplio número de ejemplos negativos y algunos positivos lo cual se balancea con una técnica denominada *'hard negative mining'*. Debido a este enfoque de detección multiescala, SSD es muy eficaz en detectar objetos de diferentes tamaños en una imagen.

Otro MPDO que trabaja en una sola etapa es YOLO (Redmon et al., 2016), el cual implementa RNC. En las capas iniciales se extraen las características de la imagen y las capas completamente conectadas predicen las probabilidades de salida y las coordenadas. Es un sistema muy rápido, lo cual lo vuelve apto para la detección de objetos en tiempo real, y además entrena y detecta con la imagen completa logrando menos falsos positivos. Existen ya varias versiones de esta técnica, ya que está en constante mejoramiento. La primera versión

fue YOLOv1, que divide la imagen de entrada en celdas de cuadrícula, y cada celda es responsable de la detección de objetos. Predice una clase por celda de cuadrícula y posee errores de localización para objetos pequeños y objetos en grupos. Luego, YOLOv2 (Redmon y Farhadi, 2017) proporciona una mejora en la precisión y la velocidad respecto de YOLOv1, posee más cajas de predicción y también implementó una capa de transferencia que concatena las características de una capa de mayor resolución a una de menor (26 x 26 capas a 13 x 13 capas), proporcionando la ventaja de la detección de objetos pequeños (Redmon y Farhadi, 2017). La estructura principal de extracción de rasgos de YOLOv2 es Darknet-19 (el cual tiene 19 capas, la mayoría de las cuales utiliza filtros de 3 x 3) y posee 22 capas de convolución y una capa de detección (Koirala et al. 2019b).

Posteriormente surge YOLOv3 (Redmon y Farhadi, 2018) como una mejora sobre YOLOv2 en términos de detección y precisión, utiliza Darknet-53 (el cual tiene 53 capas), y la predicción de probabilidades se basa en una regresión logística. Posee 75 capas convolución y tres capas de detección. Más tarde surge YOLOv4 que tiene como objetivo principal optimizar el detector para poder realizar cálculos en paralelo. En su arquitectura aparece CSPDarknet53 que es una red que aumenta la capacidad de aprendizaje, un módulo de agrupación PANet y utiliza la cabeza de YOLOv3.

Una versión posterior fue YOLOv5 que se presenta en ambiente de trabajo Pytorch, es más rápido y preciso en la detección de objetos respecto de las versiones tres y cuatro. La versión YOLOv5s no solo mejora la velocidad y precisión de inferencia, sino que tiene una estructura más pequeña. Utiliza también CSPDarkNet para la extracción de rasgos de la imagen, un módulo de agrupación PaNet y utiliza la cabeza de YOLOv3 (Wang y He, 2021). La versión cinco y su posterior, YOLOv6, creada por la compañía China Meituan, no son consideradas oficiales por los desarrolladores de YOLO.

Actualmente se presentó la versión YOLOv7 que reduce los parámetros utilizados por el modelo, logra mayor velocidad de inferencia y mayor precisión en la detección (Wang, et al., 2022).

El desarrollo de estos modelos ha sido y continúa siendo muy rápido en el tiempo y en general siempre se persigue incrementar dos aspectos fundamentales que son: i) lograr la mayor precisión, y ii) la menor velocidad de procesamiento. Para poder evaluar las performances de los diferentes MPDO de visión artificial existen ciertos parámetros y medidas de evaluación, las cuales se recomienda que al menos alguna de ellas sea reportada cuando se realizan trabajos asociados a la temática (Koirala et al., 2019b).

2.1.6. Métricas de evaluación de modelos de visión artificial

Cuando se trabaja con detección de objetos, los resultados posibles son:

- i) verdadero y positivo (VP) que es el objeto de interés reconocido como tal;
- ii) falso y positivo (FP) que es el entorno reconocido como el objeto de interés, pero no lo es;
- iii) verdadero y negativo (VN) que es el entorno reconocido como tal;
- iv) falso y negativo (FN) que son los objetos de interés no reconocidos.

A partir de estos resultados surgen métricas como la precisión (Ecuación 1) y el *'recall'* o sensibilidad (Ecuación 2). La primera nos da el número de objetos detectados respecto de la detección total, mientras que la segunda nos dice cuántos de los objetos de interés es capaz de detectar el modelo, es decir, nos da una idea de la exhaustividad del modelo. Se busca obtener alta precisión y sensibilidad cuando existe un alto costo asociado a FP o FN, respectivamente (Koirala et al., 2019b). Además, el área bajo la curva de precisión y sensibilidad puede ser utilizada como una métrica simple para resumir el comportamiento de los modelos y se lo conoce como precisión media o Average Precision (mAP) (Koirala et al., 2019b).

Otra medida muy utilizada es el F1 Score (Ecuación 3) que representa un promedio ponderado de la precisión y la sensibilidad y se calcula como el punto de la curva donde ambas métricas tienen el mayor e igual valor. Puede tomar valores desde cero (peor) a uno (mejor) (Koirala et al., 2019b).

Ecuación 1: fórmula de cálculo de precisión

$$\text{Precisión} = \frac{VP}{VP + FP}$$

Ecuación 2: fórmula de cálculo de sensibilidad o exhaustividad

$$\text{Sensibilidad} = \frac{VP}{VP + FN}$$

Ecuación 3: fórmula de cálculo para el F1 score

$$F1 = 2 * \frac{Precision * Sensibilidad}{Precision + Sensibilidad}$$

Koirala et al., (2019b) sugiere que cuando se trabaja en detección de frutos en ambientes naturales, el éxito de los modelos debería evaluarse considerando la precisión, sensibilidad y F1 score, resaltando la importancia de este último que es de uso común en trabajos del tema y por lo tanto permite realizar comparaciones.

2.1.7. Antecedentes del uso y evaluación de los modelos de visión artificial para la detección de frutos

En la última década se han reportado varios estudios que utilizan diferentes MPDO basados en algoritmos o arquitecturas de DL para la detección de frutos:

- ✓ Sa et al., (2016) trabajaron en la detección de distintos frutos mediante el uso de Faster R-CNN y VGGNet. Utilizan técnicas de transferencia de aprendizaje para entrenar estos MPDO con pocas imágenes (entre 43 y 136) tomadas principalmente de internet. Los frutos utilizados son melón, frambuesa, manzana, naranja, mango, palta y pimienta; y los valores de F1 score obtenidos en cada caso fueron muy satisfactorios, siendo 0,848; 0,948; 0,938; 0,915; 0,942; 0,932 y 0,828 respectivamente.
- ✓ La técnica de Faster R-CNN también fue utilizada por Bargoti y Underwood, (2017) para detectar frutos de manzana, mango y almendra en imágenes de árboles frutales en condiciones naturales con alta intensidad de iluminación logrando altos valores de F1 score: 0,904; 0,908 y 0,775, respectivamente con VGGNet; y 0,892; 0,876 y 0,726 respectivamente con ZFNet.
- ✓ Koirala et., al (2019a) compara seis MPDO para detectar mangos en arboles de montes frutales: Faster R-CNN(VGG), Faster R-CNN(ZF), YOLOv1, YOLO v2, YOLOv3 y SSD, La comparación resulta muy interesante porque utilizan las diferentes técnicas con las mismas bases de datos y características computacionales. Además, desarrollan una técnica que posee características de YOLOv2 (menos capas, más veloz) y

YOLOv3 (múltiples capas para detección, más precisión) la cual llaman MangoYOLO, logrando valores de F1 score de 0,968 y una precisión promedio de 0.983. Además, probaron la técnica con un nuevo conjunto de imágenes de mango logrando un F1 score de 0,89.

- ✓ En mangos, Liang et al. (2018), adapta el modelo SSD para la detección de frutos en imágenes obtenidas de las chacras y además lo comparan con VGG16 y ZFNet de Faster R-CNN. Con este último además utilizan dos resoluciones diferentes de imágenes; ZFNet 0,4 (400 x 400) y ZFNet 0,5 (500 x 500). Los valores obtenidos de F1 score fueron de 0,787; 0,884; 0,798; 0,761 y 0,911 para el modelo SSD original, el SSD adaptado, el ZFNet 0,4, el ZFNet 0,5 y el VGG16, respectivamente. Los valores de precisión obtenidos fueron de 0,947; 0,967 y 0,920 y los de sensibilidad de 0,690; 0,627 y 0,903 para ZFNet 0,4; el ZFNet 0,5 y el VGG16, respectivamente.
- ✓ Chen et al., (2016) desarrollan su propio modelo de detección de frutos basado en DL que utiliza un detector de RI basado en un RN convolucional y un segundo algoritmo basado en otra RN que estima el número de frutos de cada RI, para finalmente hacer la estimación mediante una regresión lineal. La performance del modelo se evaluó en un set de 71 imágenes de naranjas tomadas de día y en un set de 21 imágenes de manzanas tomadas de noche y se obtuvo un F1 score de 0,76. Por un lado, con el set de manzanas se obtuvo una tasa de VP de 0,961 y de FP de 0,033, mientras que para el set de naranjas 0,957 y 0,051 para VP y FP, respectivamente.
- ✓ Bresilla et al., (2019) entrenan YOLOv2 para detección de peras y manzanas logrando resultados de F1score entre 0,79 y 0,89 dependiendo principalmente de la cantidad de fotos utilizadas para el entrenamiento. Las fotos fueron tomadas en diferentes momentos, cultivares, condiciones de luz y con diversas cámaras para generar la mayor variabilidad posible.
- ✓ Tian et al., (2019), utilizaron YOLOv3 para detectar manzanas en diferentes estadios de crecimiento (frutos jóvenes, en crecimiento y maduros) con imágenes tomadas en diferentes condiciones y horarios. Mejoran el modelo agregando un método DenseNet para optimizar algunas capas de la RN de YOLO v3 y lo comparan con YOLOv3 original y con VGG16 de Faster R-CNN. Obtienen valores de F1 score de 0,832; 0,841; 0,864 y 0,817 para manzanas jóvenes, en crecimiento, maduras y totales respectivamente con YOLOv3, y encuentran mejor este modelo respecto al VGG16.
- ✓ Wang y He, (2021), ya utiliza YOLOv5s para detectar ramilletes de frutos en manzanos previos al raleo con imágenes tomadas en diferentes horarios y en días

soleados y nublados. El modelo muestra buenos resultados de precisión, sensibilidad, F1 score y FP (95,8%; 87,6%; 0,915 y 4.2% respectivamente) y además lo comparan con otros modelos de detección (distintas versiones de Faster R-CNN y SSD).

- ✓ Itakura et al., (2021) trabajó en la detección de peras y manzanas a partir de videos utilizando YOLO en sus versiones dos y tres. Se utilizaron videos tomados con teléfonos iPhoneX en diferentes chacras. Para el modelo de detección utilizan YOLOv2 y YOLOv3 y luego aplican un filtro para el conteo de frutos. Considerando los frutos omitidos, los conteos múltiples y las detecciones erróneas, se contaron 226 peras de 234 y 157 manzanas de 170. El F1 score del modelo fue de 0,972 y de 0,929 para pera y manzana respectivamente.

Si bien la detección de frutos a partir del análisis de imágenes ha sido ampliamente estudiada por muchos años, todavía restan aspectos a mejorar. Por un lado, la segmentación o detección de los frutos en las imágenes generalmente se ve muy afectado por las variaciones en las condiciones de luz del medio ambiente (Mao et al., 2009). Además, la detección y localización de la fruta se ve afectada por la estructura de las copas de los árboles que es muy compleja (Karkee y Zhang, 2012) y la precisión de los métodos es limitada por la oclusión de los frutos que ejercen las hojas, ramas e incluso otros frutos presentes en la imagen (Gongal, 2015). Por otra parte, son pocos los estudios realizados en la detección de peras y, si bien, en manzanas se han realizado varios, su mayoría se desarrollan cuando los frutos ya toman color y están cercanos a cosecha. Por este motivo, resulta de gran interés trabajar en la detección y conteo del número de frutos de pera y manzana tempranamente a partir de imágenes en el Alto Valle de Río Negro y Neuquén, con las características propias de los montes frutícolas regionales. Para ello se evaluarán MPDO para la de detección y conteo de frutos a partir de imágenes de árboles de pera y manzana que logren alta precisión, y que se adapten a las condiciones regionales.

2.2. MATERIALES Y METODOS

2.2.1. Características de los montes frutales

2.2.1.1. Pera cv. 'Williams' de la Estación Experimental de INTA

El trabajo se realizó en una parcela de perales cv 'Williams' perteneciente a la Estación Experimental INTA Alto Valle "Ing. Agr. Carlos H. Casamiquela" (Figura 6), implantado en el año 2003 sobre portainjerto franco. La superficie de la misma es de 1,8 hectáreas, la conducción de los árboles es en espaldera modificada (palmeta) y el marco de plantación es de cuatro m entre filas por dos m entre plantas. Hay un total de 60 filas con 36 plantas cada una. El monte frutal recibió cada temporada un manejo similar a los montes frutales comerciales de la región, realizándose las tareas de poda, fertilización, desmalezamiento y raleo, entre otras. El tipo de riego es por manto con una frecuencia de 15 días aproximadamente.

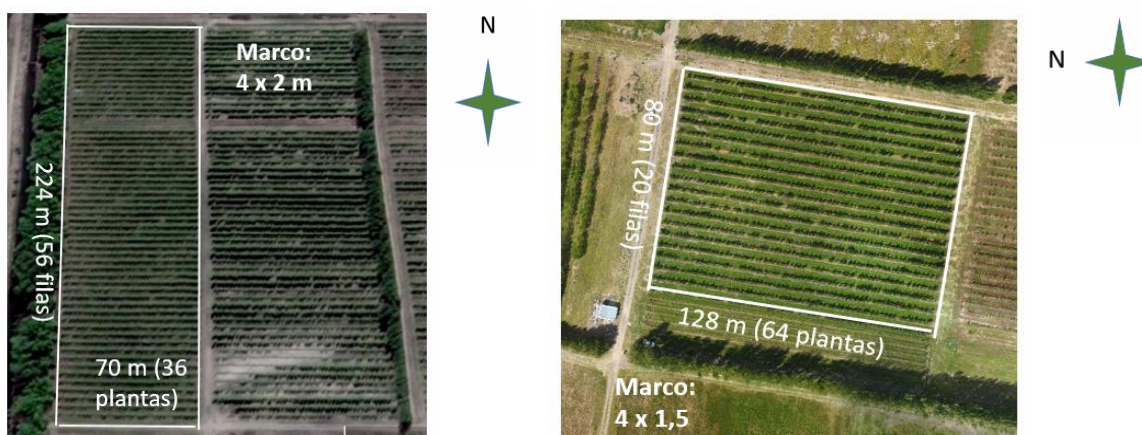


Figura 6: *Izquierda:* Parcela de perales cv 'Williams'. *Derecha:* Parcela de manzana. Ambas de la Estación experimental INTA Alto Valle

2.2.1.2. *Manzana cv. 'Red Chief' de la Estación Experimental de INTA*

También se trabajó en una parcela de manzanas cv 'Red Chief' perteneciente a la Estación Experimental INTA Alto Valle "Ing. Agr. Carlos H. Casamiquela" (Figura 6), implantada en el año 2000 sobre portainjerto MM111 con una superficie es de 0,8 hectáreas. La conducción de los arboles es en espaldera y el marco de plantación es de cuatro m entre filas por 1,5 m entre plantas. Hay un total de 24 filas con 64 plantas cada una. El monte frutal recibió cada temporada un manejo similar a los montes frutales comerciales de la región, realizándose tareas de poda, fertilización, desmalezamiento y raleo, entre otras. El tipo de riego es por micro-aspersión con una frecuencia diaria durante el ciclo productivo.

2.2.1.3. *Montes frutales comerciales.*

Además, se trabajó en cuatro montes frutales comerciales pertenecientes a dos empresas: Patagonian Fruits SRL y Román SRL para poder evaluar el funcionamiento de los modelos ajustados en otras condiciones. Dichos montes se describen en la Tabla 1:

Tabla 1: Características de los montes frutales utilizados para los muestreos

<i>Sitio</i>	<i>Cultivar</i>	<i>Año de plantación</i>	<i>Conducción</i>	<i>Marco (m)</i>	<i>Superficie (ha)</i>
<i>Patagonian Fruits</i>	<i>Red Chieff</i>	<i>2011</i>	<i>Espaldera modificada</i>	<i>4x2m</i>	<i>2,4</i>
<i>Patagonian Fruits</i>	<i>Williams</i>	<i>2005</i>	<i>Espaldera</i>	<i>3,8x2</i>	<i>1,4</i>
<i>Román SRL</i>	<i>Red Chieff</i>	<i>1994</i>	<i>Espaldera</i>	<i>4x1,5</i>	<i>0,15</i>
<i>Román SRL</i>	<i>Williams</i>	<i>1998</i>	<i>Espaldera</i>	<i>4x2</i>	<i>0,71</i>

2.2.2. Obtención de datos para la incorporación de información proveniente de sensores a los modelos predictivos

2.2.2.1. Cámaras digitales RGB

Los datos recolectados para la incorporación de información proveniente de cámaras digitales RGB se tomaron durante cuatro temporadas en la estación experimental INTA (desde 2017 a 2021) y solo en la última temporada se adicionaron los montes frutales comerciales.

Selección de árboles muestra

En la parcela de pera de INTA se seleccionaron 60 árboles representativos del monte frutal y distribuidos de manera regular. De la misma manera, en la parcela de manzana de INTA se seleccionaron 72 plantas.

En las parcelas de peras del monte frutal comercial de Patagonian Fruits SRL se seleccionaron 16 plantas al azar del monte frutal y en la parcela de manzanas de dicha empresa se seleccionaron aleatoriamente 15 plantas.

Por su parte, en la parcela de manzanas del monte frutal comercial de Román SRL se seleccionaron 14 plantas al azar y en la parcela de peras de dicha empresa se seleccionaron aleatoriamente 15 plantas.

Obtención de imágenes digitales de los árboles muestra en reposo invernal

Con una cámara digital de 14 MP se tomaron fotos a todas las plantas de la muestra de INTA en el mes de agosto durante la temporada 2020-2021. Las fotos fueron tomadas al medio día, en condiciones parcialmente nubladas, para evitar efectos indeseados de la luz directa del sol (Dorj et al., 2017). Para cada imagen se colocó un objeto cuadrado de dimensiones conocidas para utilizar como escala o distancia conocida y se utilizó una cortina negra detrás de cada planta para evitar interferencias con las plantas de las filas vecinas. Las fotos se tomaron a 2,5 m de distancia a la planta (Figura 7).

Obtención de imágenes y conteo de flores en árboles muestra

Se tomaron fotos a todos los árboles de la muestra en floración. En general, de dos a tres días previos al estado fenológico de plena floración, lo cual ocurrió para todas las temporadas entre el 24 y el 29 de septiembre.

Las fotos fueron capturadas en dos condiciones diferentes, i) de día siguiendo la metodología utilizada para las fotos de reposo invernal, ii) de noche, ya que las imágenes obtenidas con luz artificial tienen la ventaja de minimizar la sensibilidad que poseen aquellas que se toman de día con diferentes condiciones de luminosidad ambiental (Linker, 2016). A partir de la temporada 2018-2019 se utilizó una cortina negra detrás de cada planta para evitar interferencias con las plantas de las filas vecinas (Figura 7).

El mismo día en que se obtuvieron las imágenes se realizó el conteo de ramilletes florales de cada planta. El conteo fue realizado en forma manual por dos o más operarios, según el caso y se obtuvo un promedio del conteo para cada planta.

Obtención de imágenes y conteo de frutos en arboles muestra

De forma similar a las fotos de floración se tomaron las fotos de las plantas con frutos en el mes de diciembre. Las fotos se tomaron entre los días 03 y 15 de diciembre para todas las temporadas (Figura 7).

El mismo día en que se obtuvieron las imágenes se realizó el conteo de frutos de cada planta. El conteo fue realizado en forma manual por dos o más operarios, según el caso y se obtuvo un promedio del conteo para cada planta.

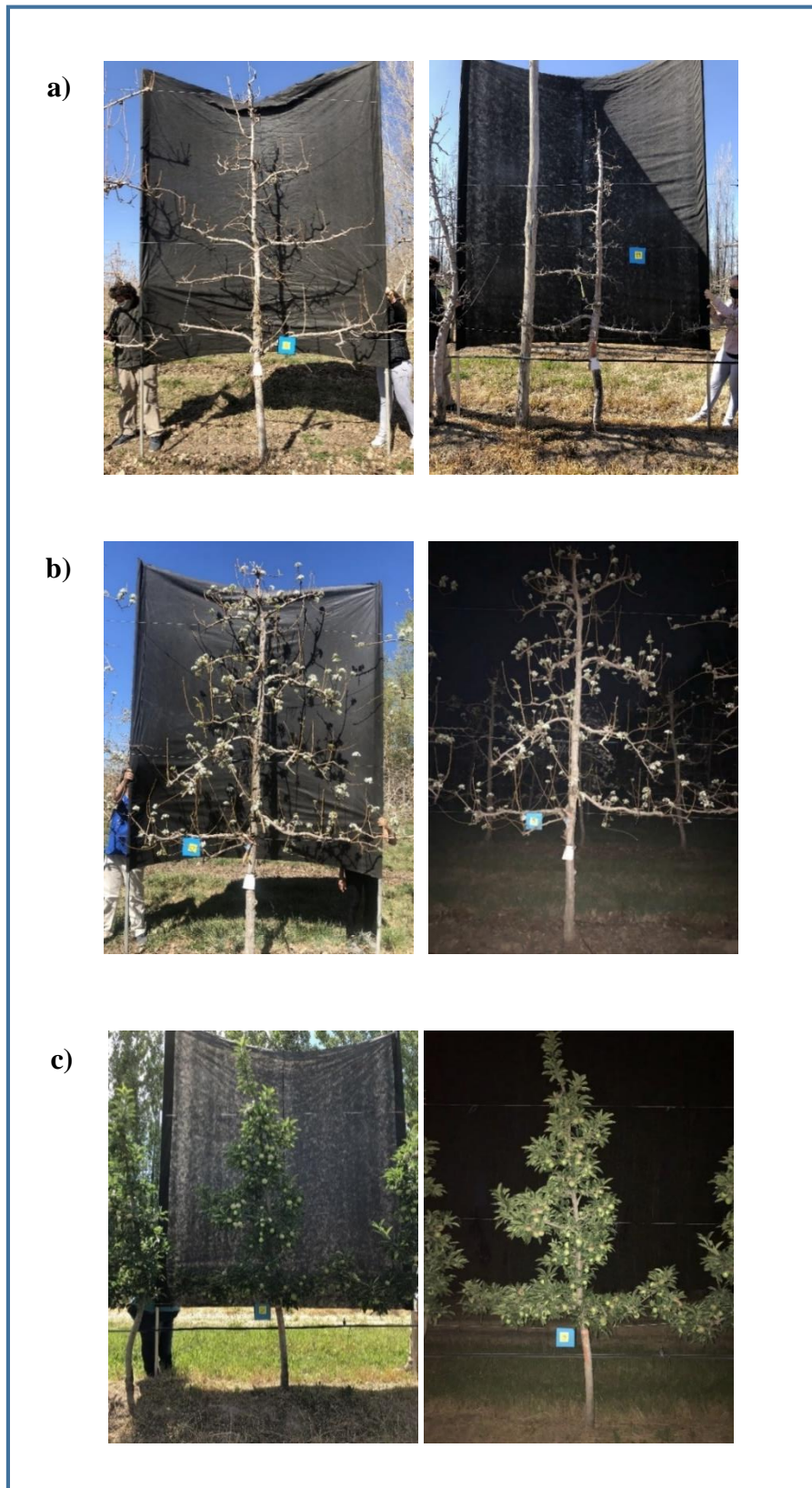


Figura 7: **a)** Fotos de reposo invernal. **Izquierda:** árbol de pera cv William's. **Derecha:** árbol de manzana cv Red chief; **b)** Fotos en floración de un árbol de pera cv 'William's'. **Izquierda:** foto tomada de día. **Derecha:** foto tomada de noche; **c)** Fotos con frutos de un árbol de manzana cv 'Red chief'. **Izquierda:** foto tomada de día. **Derecha:** foto tomada de noche.

2.2.2.2. Cosecha de árboles muestra y tamaño de frutos

Cada una de las plantas seleccionadas fue cosechada en su totalidad y se registró el peso en kilogramos obtenido de cada planta, el número de frutos que presentó cada una y la clasificación los frutos según tamaños comerciales (Tabla 2).

Tabla 2: Tamañado de peras y manzanas. Clases y rangos de pesos y diámetros asociados.

	<i>Peras</i>			<i>Manzanas</i>		
	<i>Peso</i> (g)	<i>Diámetro</i> (mm)	<i>Tamaño</i> <i>comercial</i>	<i>Peso</i> (g)	<i>Diámetro</i> (mm)	<i>Tamaño</i> <i>comercial</i>
<i>Clase 1</i>	>200	>70	80	>220	>80	80
<i>Clase 2</i>	180-199	~67,5	90	200-219	~77,5	80
<i>Clase 3</i>	160-179	~65	100	180-199	~75-76	90
<i>Clase 4</i>	140-159	~62	110	160-179	~74	100
<i>Clase 5</i>	130-139	~60	120	140-159	~70-71	110
<i>Clase 6</i>	120-129	~58	135	130-139	~66-68	125
<i>Descarte</i>	<120	<58	150 o >	<130	<66	135 o >

2.2.3. Análisis de imágenes

2.2.3.1. Detección y conteo de frutos

Para la detección y conteo de frutos se utilizaron varias técnicas de DL, todas pertenecientes a YOLO en diferentes versiones. Los modelos pre-entrenados de YOLO están disponibles en el repositorio github: <https://github.com/AlexeyAB/darknet>.

Los entrenamientos para detección de peras y manzanas se realizaron por separado. Cada modelo utilizó su propio set de datos. En ambos casos el set de datos incluye todas las fotos obtenidas a lo largo de todas las temporadas, tanto las de día como las de noche. El set de datos de manzana quedó constituido por 505 imágenes y 36.260 anotaciones y el de peras por 474 imágenes y 39.052 anotaciones (Figura 8). Las anotaciones se refieren a cada una de las clasificaciones de objetos en las categorías pera o manzana según sea el caso.

Para la creación de cada set de datos, en primera instancia se creó un archivo de Lenguaje de Marcado Extensible (XML por sus siglas en inglés: Extensible Markup Language) por cada imagen con la clasificación de todos los objetos de interés presentes en la misma y con sus respectivas localizaciones. Para ello se utilizó una aplicación de Roboflow con la cual se realizó una clasificación de objeto con las categorías pera o manzana y se señaló en cada imagen todas las veces que apareció dicho objeto mediante un recuadro. Luego cada set de datos se descargó a modo de carpeta de archivo comprimido o enlace para ser utilizado en el entrenamiento.

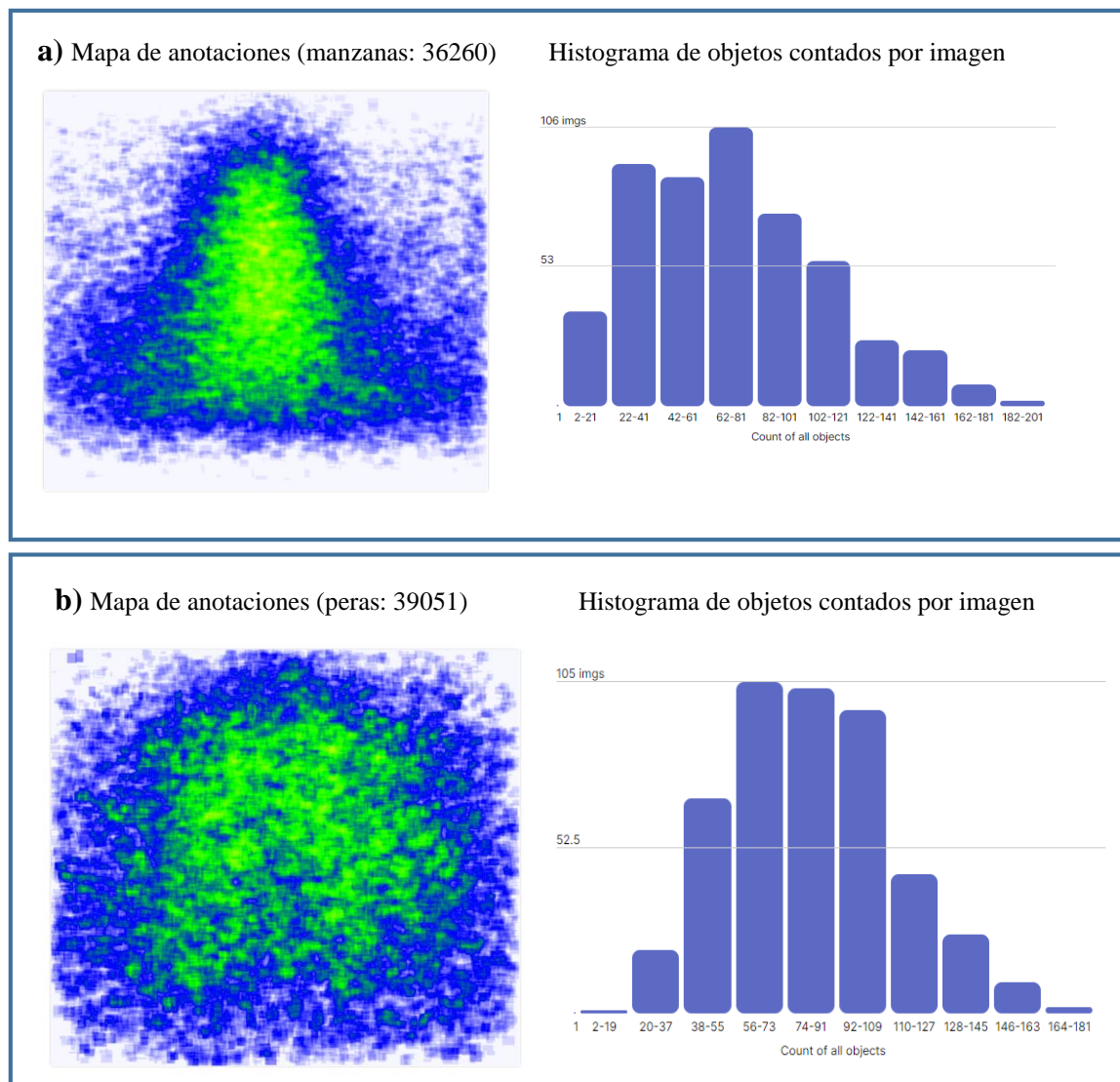


Figura 8: Mapas de anotaciones en set de imágenes e histograma de objetos contados por imagen. A) Manzanas; B) Peras. Tomados de <https://app.roboflow.com/doloresdb-fbagg/williams-lfylv/health>

Los creadores de los modelos recomiendan utilizar imágenes con baja resolución para lograr una mayor velocidad, mientras que para aumentar la precisión es necesario trabajar con una resolución más alta. En este trabajo se utilizaron dos resoluciones: 1024 x 1024 y 1280 x 1280 píxeles.

Además, en todos los casos se realizó un aumento de datos o `augmentation` ya que mediante esta técnica se puede multiplicar la información mediante la transformación de las imágenes y así obtener mejores resultados (Bresilla et al., 2019).

Se calcularon las métricas de evaluación (precisión, sensibilidad, F1 score y mAP) para cada uno de los modelos entrenados.

Por último, se calcularon las diferencias entre los números de frutos detectados por los modelos entrenados (visión artificial) y los números de frutos contados manualmente sobre el mismo lado del árbol para cada árbol. Se realizó además una prueba de hipótesis sobre la igualdad a cero de la media de las diferencias. La prueba utilizada es una prueba t para muestras apareadas en el caso de que los datos presenten distribución normal.

2.2.3.2. *Correlación entre frutos contados vs estimados*

Para corroborar cómo funcionan los MPDO se calcularon los coeficientes de correlación para los diferentes conteos de frutos:

- i) Manual lado imagen: es el número frutos contados manualmente a campo al momento del pronóstico, sólo del lado del árbol en que se tomó la imagen;
- ii) Manual total: es el número total de frutos que surge del conteo de todos los frutos del árbol fotografiado a campo al momento del pronóstico (esto incluye el lado en que se capturó la imagen y el otro lado en el que no se capturó);
- iii) Etiquetados: es el número de frutos etiquetados en las imágenes, es decir, el conteo humano del número de frutos que se visualizan o etiquetan en cada imagen (detectados por el ojo humano);
- iv) Detectados: es el número de frutos estimados por los MPDO, es decir, es el número de frutos detectados por visión artificial a partir de los MPDO;
- v) Cosechados: es el número de frutos obtenido por cada árbol el día de cosecha.

También se realizó una matriz de diagramas de dispersión para cada par de variables.

2.2.3.3. *Estimación del tamaño de frutos*

Se trabajó bajo dos situaciones diferentes, por un lado, se realizó una evaluación en condiciones de laboratorio, y por otro lado se trabajó en condiciones de campo. En el primer caso se trabajó con 106 frutos ya cosechados de diferentes tamaños comerciales. Los frutos fueron colocados en bandejas de celulosa, se numeraron y se midió el diámetro ecuatorial de los mismos de forma manual con un calibre digital (Essex 150 mm, Stainless Hardened). Luego se tomaron las imágenes a cada bandeja con una cámara de fotos de 12 MP, colocando un objeto de dimensiones conocidas en cada una.

A campo se trabajó con 89 frutos de manzana y 98 de pera ubicados en varios árboles. Se enumeraron los frutos, se midió el diámetro ecuatorial de los mismos manualmente con calibre electrónico y se tomaron las imágenes de cada planta colocando un objeto de dimensiones conocidas en cada una.

El análisis de las imágenes se realizó con el software ImageJ (Schneider et al., 2012). Para ajustar la escala en cada imagen, se trazó una línea recta sobre el objeto de dimensiones conocidas y se convirtió el valor de distancia en píxeles arrojado por defecto por ImageJ, a un valor de distancia conocida en mm. Para estimar el diámetro de cada fruto presente en la imagen se trazó una línea sobre la zona ecuatorial del fruto y se determinó la longitud seleccionada (Figura 9). Luego se realizó una correlación de Pearson entre los diámetros obtenidos manualmente y los estimados a partir de imágenes digitales para cada situación analizada.

Además, se realizó el cálculo de las diferencias de los diámetros estimados mediante imágenes versus los medidos manualmente para construir un gráfico de distribución de frecuencias de las mismas. Por último, se realizó la comparación de medias de los diámetros obtenidos de la medición manual y los estimados a partir de imágenes para cada una de las condiciones analizadas. Se corroboró la distribución normal de los datos mediante pruebas Shapiro Wilk y luego se realizó una prueba t en los casos en que se cumplió dicho supuesto. En los casos en los que no se cumplió con la normalidad de los datos se compararon las medias mediante pruebas Wilcoxon.



Figura 9: Imágenes de manzanas y peras numeradas en plantas y bandeja; y de la escala utilizada a campo y en laboratorio.

2.3. RESULTADOS Y DISCUSION

2.3.1. Detección y conteo de frutos

Todos los modelos entrenados para la detección de manzanas lograron valores de precisión y de precisión promedio (mAP) por encima de 0,8 y de sensibilidad por encima de 0,7 (Figura 10). Sin embargo, el mejor modelo fue el YOLOv5n6 con una resolución de imagen de 1280 x 1280 por lo cual se analizaron los resultados en base al mismo.

En el caso de las peras los distintos modelos entrenados lograron diferentes valores de precisión, precisión promedio (mAP) y sensibilidad (Figura 11). Sin embargo, el mejor modelo fue el YOLOv7 con una resolución de imagen de 1280 x 1280 pixeles, por lo cual se analizaron los resultados en base al mismo.

Los modelos analizados lograron una precisión de 0,86 y 0,87 para manzanas y peras respectivamente. Es decir, aproximadamente el 86% de las detecciones fueron realmente frutos (Figura 13). La sensibilidad lograda por los modelos fue de 0,83 para ambas especies. Esto arrojó valores de F1 score 0,83 y 0,84 para manzanas y peras respectivamente (Figura 12).

Los valores obtenidos son menores que los presentados por Sa et al. (2016) para melón, frambuesa, manzana, naranja, mango, palta y pimienta con Faster R-CNN y VGGNet, que logran valores cercanos a 0,9. Ellos trabajaron con fotos propias para pimienta y melón, y fotos de internet para el resto de los frutos para poder demostrar la robustez de los modelos. Los mayores valores obtenidos por dichos autores pueden deberse a que las imágenes contemplan frutos en diferentes estados de madurez y en muchos casos presentan una coloración muy diferente al resto de la planta. También, a que las imágenes solo tienen una parte de plantas y no la planta entera, lo cual hace que los frutos representen una mayor proporción de la imagen. En cambio, en esta tesis se trabajó con imágenes solamente tomadas a campo en condiciones naturales, y de las plantas enteras, lo cual implica que los frutos son pequeños en relación al total de la imagen. Además, el hecho de no utilizar fotos con frutos tomados más de cerca puede empeorar la performance de los algoritmos ya que los bordes de los frutos pueden presentar una menor nitidez.

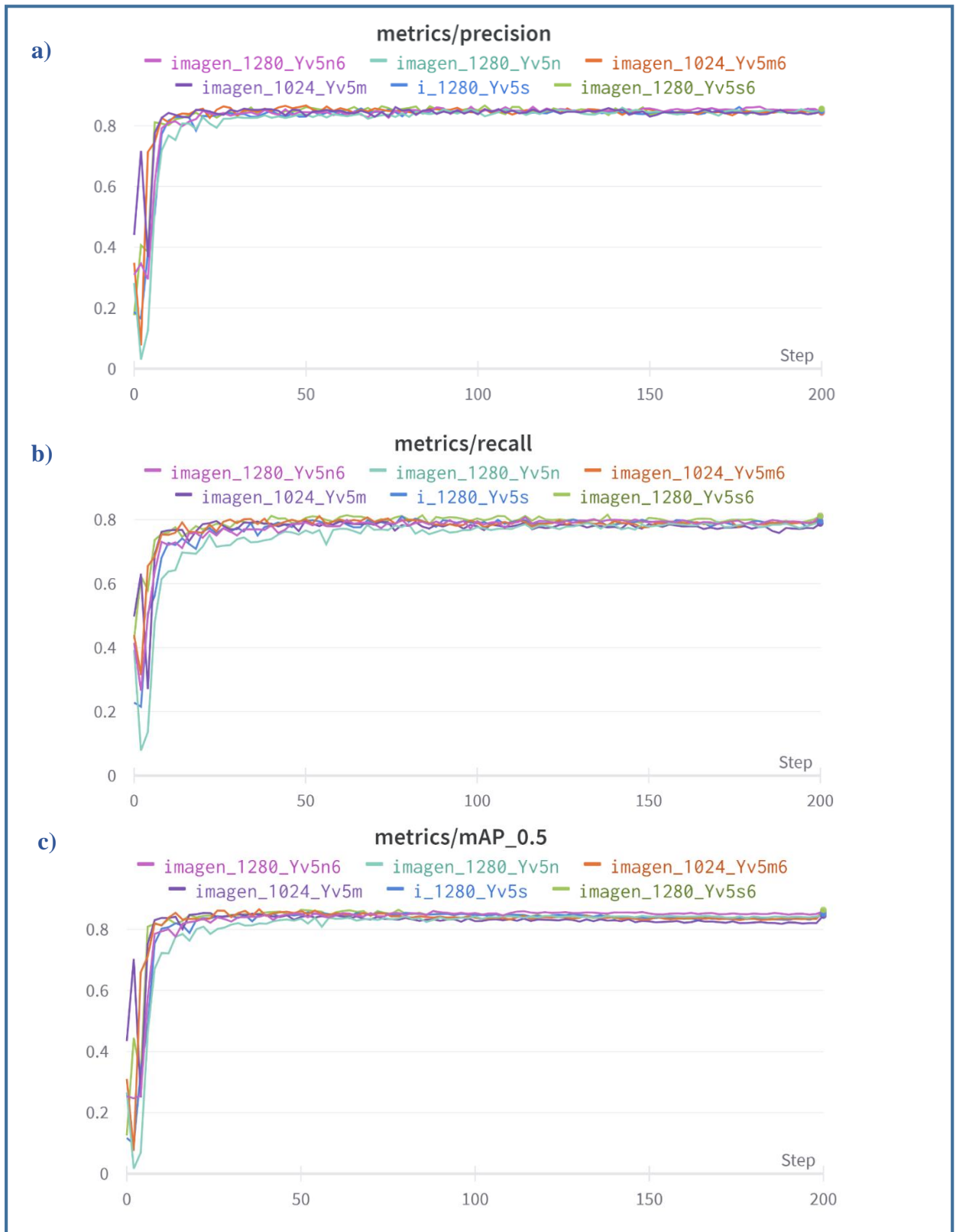


Figura 10: **a)** Gráfico de precisión vs pasos del entrenamiento; **b)** Gráfico de sensibilidad vs pasos del entrenamiento. **c)** Gráfico de precisión promedio 0.5 versus pasos del entrenamiento. Gráficos de los MPDO utilizados para imágenes de manzana. Obtenido de <https://wandb.ai/delbrioyolov5s-W?workspace=user-delbrio>

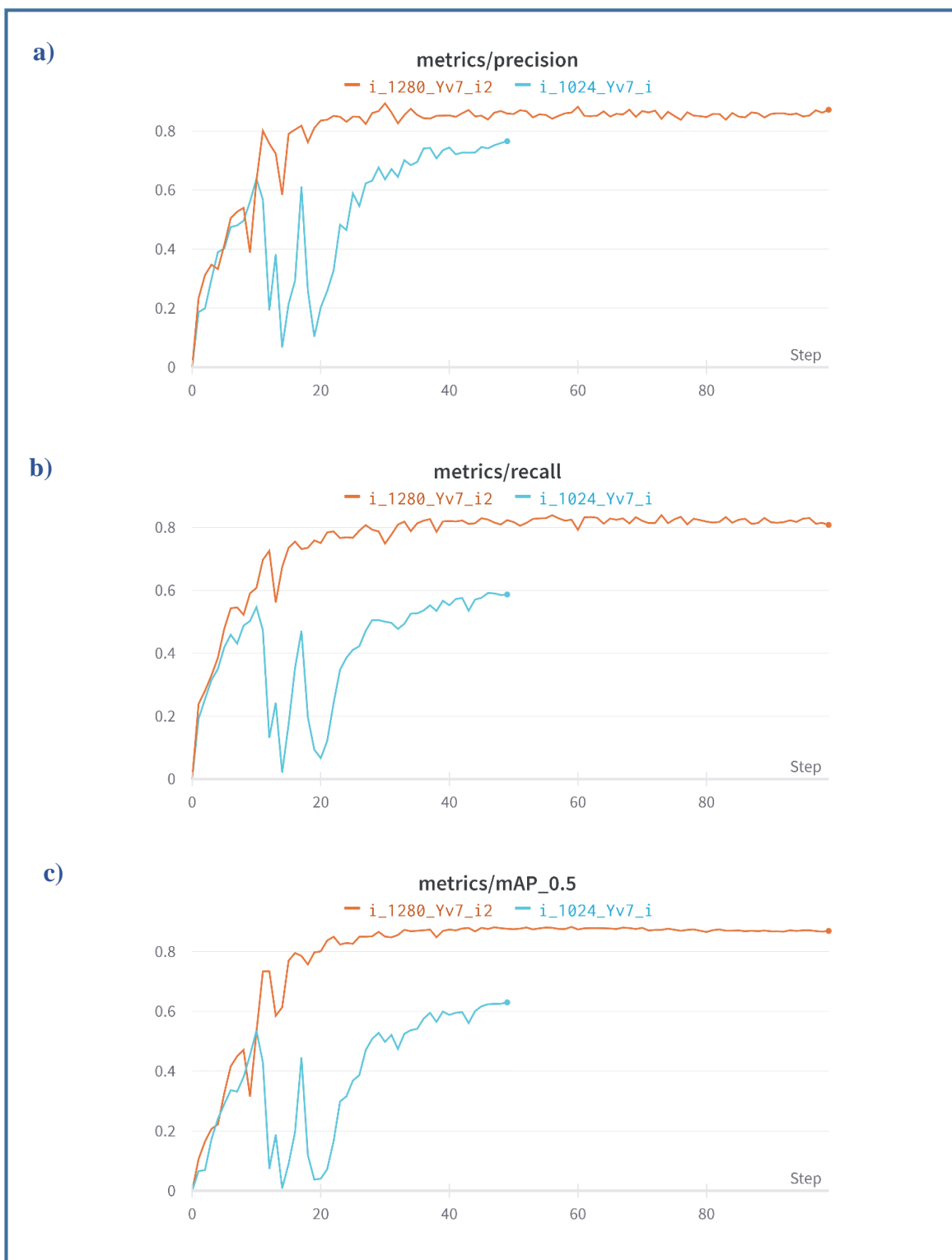


Figura 11: **a)** Gráfico de precisión vs pasos del entrenamiento; **b)** Gráfico de sensibilidad vs pasos del entrenamiento. **c)** Gráfico de precisión promedio 0.5 versus pasos del entrenamiento. Gráficos de los MPDO utilizados para imágenes de pera. Obtenido de <https://wandb.ai/delbrio/yolov5s-W?workspace=user-delbrio>

Bargoti y Underwood, (2017) por su parte, lograron también valores de precisión y de F1 score similares a los obtenidos en esta tesis. Según Koirala et al., (2019b), los valores obtenidos pudieron ser menores a otros de referencia debido a que las imágenes están influenciadas por las condiciones de los cultivos y el clima, y además porque había un gran número de frutos por imagen (100 para mangos y manzanas, y hasta 1000 en el caso de almendras). Ambas situaciones coinciden con lo realizado en este trabajo, tanto el número de frutos que llegó a ser de 170 por imagen en peras y 150 en manzanas, como la captura de las mismas en condiciones naturales. Los autores, por su parte, atribuyen los errores de su modelo principalmente a la incapacidad de detectar todos los frutos de un ramillete o grupo, lo cual también ocurrió en esta tesis.

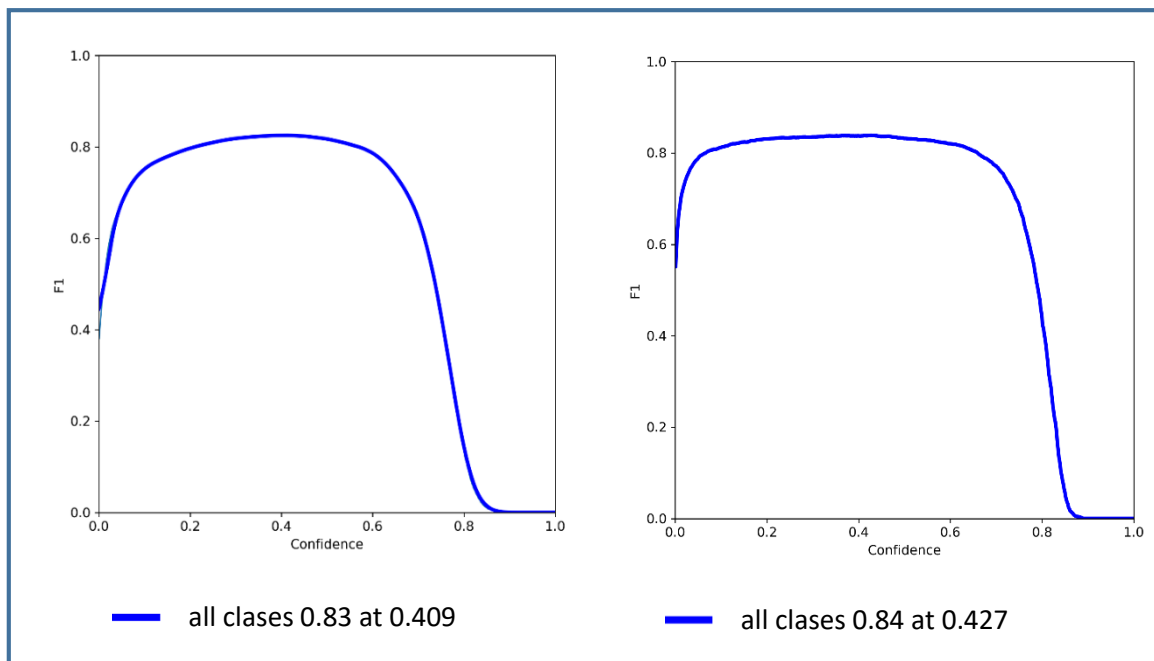


Figura 12: Izquierda: Gráfico de F1 score para el MPDO seleccionado para manzanas. Derecha: Gráfico de F1 score para el MPDO seleccionado para peras. Obtenido de <https://wandb.ai/delbrioyolov5s-W?workspace=user-delbrioyolov5s-W>

Para mangos, también se desarrolló el modelo híbrido llamado MangoYOLO combinando criterios de precisión y velocidad de otras versiones de YOLO, que logró una precisión promedio de 0,98 en su entrenamiento y, utilizado con un nuevo conjunto de imágenes (publicados en la web por otros autores), logró un F1 score de 0,89 (Koirala et., al 2019a). Los mayores valores de precisión logrados en comparación a esta tesis, pueden deberse no solo a que los autores probaron y compararon varios MPDO, sino que también desarrollaron un híbrido. Además, trabajan con 1515 fotos y con aumento de datos, con lo cual logran un

número de imágenes para el entrenamiento muy superior al utilizado en el presente trabajo. Estos resultados mencionados los lograron con imágenes tomadas de noche con barras de luz led para iluminar los árboles, sin embargo, cuando probaron los modelos con imágenes tomadas de día, aumentó significativamente el porcentaje de falsos positivos. En la presente tesis se utilizaron tanto las fotos de día como las de noche, por lo cual las fotos de día con sus condiciones diferenciales de luminosidad, nubosidad, etc. pueden estar generando una disminución en la precisión de los modelos.

Liang et al., (2018) por su parte, logró valores de precisión de 0,97 y 0,92 y de sensibilidad de 0,63 y 0,90 para la detección de mangos con los modelos ZFNet 0,5 y VGG16, respectivamente. Los valores de F1 fueron entre 0,79 y 0,91. Estos autores también trabajaron con técnicas de aumento de datos, para lograr mayor robustez. Entre las dificultades mencionadas por los autores para la detección de frutos, se encontraron las diferentes densidades de hojas en la copa de los árboles, los distintos ángulos con que se capturaron las imágenes y la diferencia de las imágenes en sí. Como así también la oclusión de los frutos por las hojas, ramas y otros frutos. Estos inconvenientes también se vieron en el desarrollo de esta tesis, sobre todo en los árboles de pera, que poseen muchas hojas muy verdes que tapan una gran cantidad de frutos.

Otros resultados reportados en manzana que se aproximaron a los obtenidos en esta tesis fueron los de Chen et al., (2016) cuyo modelo arrojó un F1 score de 0,76 evaluado en un set de 71 imágenes de naranjas tomadas de día y en un set de 21 imágenes de manzanas tomadas de noche. Los valores son algo menores a los obtenidos en esta tesis, probablemente por el menor número de fotos utilizadas. Este aspecto es fundamental según Bresilla et al., (2019), quienes reportan F1 scores entre 0,79 y 0,89 según la cantidad de datos utilizados en imágenes de peras y manzanas con YOLOv2.

También los resultados obtenidos por Tian et al., (2019) han sido similares. Ellos evaluaron la performance de YOLOv3 en la detección de manzanas jóvenes, en crecimiento, maduras y totales (0,83; 0,84; 0,86 y 0,82 respectivamente). Los valores para manzanas en crecimiento se aproximan mucho a los obtenidos en la presente tesis. Los mayores valores los obtienen para manzanas maduras y esto puede deberse a que la fruta en este estado ya posee mayor tamaño y un color que se diferencia más del resto de la planta. Sin embargo, a los fines del pronóstico es muy importante el hecho de que la detección pueda realizarse en estadios de crecimiento, cuando uno esperaría realizar la detección, por ello en esta tesis se trabajó con manzanas y peras jóvenes. Estos autores trabajaron con un número similar de imágenes y con

aumento de datos. A medida que los datos o imágenes aumentaron, reportan que también lo hizo la performance del modelo.

Itakura et al., (2021) trabajó en la detección de peras y manzanas a partir de videos con YOLO en sus versiones dos y tres. La precisión promedio de los distintos modelos probados va desde 0,88 a 0,97. Uno de los motivos que podría explicar la mejor performance de estos modelos sería que muchas imágenes se tomaron desde ángulos diferentes con los cuales la fruta queda más expuesta a la vista, contrariamente a lo realizado en esta tesis que siempre se tomaron las imágenes desde la misma distancia y ángulo.

Cultivo	N	Media1 (manual)	Media2 (visión artificial)	Diferencia De medias	P –valor ($H_0:\mu dif=0$)
Manzana	45	53,86	56,90	3,05	0,3642
Pera	45	68,53	74,04	5,51	0,0762

Tabla 3: Pueba t para muestras apareadas sobre la igualdad a cero de la media de las diferencias entre los números de frutos detectado por los modelos entrenados (visión artificial) y los números de frutos contados manualmente sobre el mismo lado del árbol para cada árbol.

La media de las diferencias entre los números de frutos detectados por los modelos entrenados (visión artificial) y los números de frutos contados manualmente sobre el mismo lado del árbol para cada árbol no mostraron diferencias significativas respecto de cero (Tabla 3) para ambos cultivos. Estos resultados son muy alentadores, y más adelante en el capítulo cuatro de esta tesis se analizará el uso de dichos conteos de frutos obtenidos por visión artificial para realizar pronósticos de producción a partir de imágenes

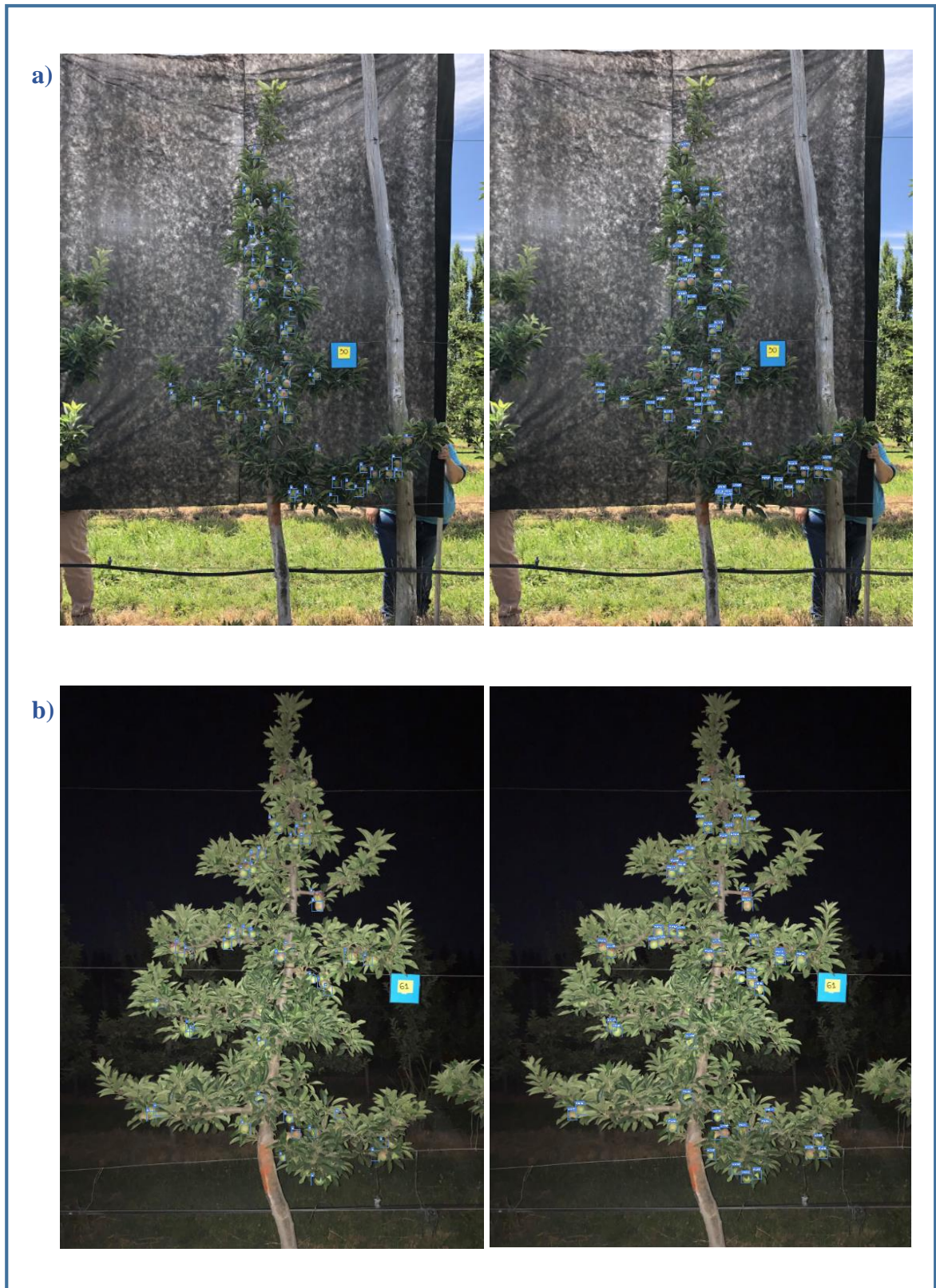


Figura 13 : a) Izquierda: foto de árbol de manzana tomada de día con los frutos etiquetados manualmente; Derecha: misma foto con frutos detectados por MPDO; b) Izquierda: foto de árbol de manzana tomada de noche con frutos etiquetados manualmente; Derecha: misma foto con frutos detectados por MPDO

2.3.2. Correlación entre frutos contados vs estimados

En la tabla 4 se pueden observar que las correlaciones obtenidas entre los frutos contados manualmente y los etiquetados o cosechados fueron diferentes en las temporadas. Durante las dos últimas temporadas de trabajo las correlaciones fueron mayores, lo cual puede deberse a que en la temporada 2018 los conteos los realizó un operario diferente (debido a una licencia por maternidad) y quizás la falta de experiencia pudo conducir a errores de conteo. Por otra parte, durante la temporada 2017 las imágenes se tomaron con una cámara de fotos digital, y algunas imágenes no lograron la nitidez deseada, por lo que las siguientes temporadas se cambió por un teléfono celular con una mayor resolución.

La correlación entre frutos etiquetados y cosechados alcanzó valores de 0,90 en el caso de manzanas, lo cual sugiere que las imágenes podrían dar una aproximación cercana de los frutos a cosechar a la hora de realizar los pronósticos de producción. En el caso de las peras los valores de correlación fueron menores, lo cual sugiere que sería más difícil realizar una predicción a partir de imágenes. Esto podría deberse a que las plantas utilizadas de pera tienen una estructura mayor a los árboles de manzana, con ramas más grandes y más cantidad de follaje. Además, los frutos al momento de la captura de las imágenes todavía son pequeños y su coloración es verde lo cual dificulta en gran medida la visión de los mismos. Esto coincide con los resultados presentados por Zhou et.al., (2012), quienes también concluyen en sus estudios realizados en manzana que si la predicción es temprana pueden quedar muchos frutos ocultos por su tamaño y por su color.

Manzanas	2017-2018	2018-2019	2019-2020	2020-2021
<i>Manual_Total vs cosechados</i>	0,712	0,902	0,948	0,870
<i>Manual_Lado_Imagen vs cosechados</i>	0,702	0,858	0,918	0,860
<i>Manual_lado_Imagen vs Etiquetados</i>	0,779	0,798	0,902	0,949
<i>Manual_total vs etiquetados</i>	0,761	0,819	0,883	0,958
<i>Etiquetados vs cosechados</i>	0,654	0,803	0,897	0,824
<i>Detectados vs Manual_Lado_Imagen</i>			0,887	
Peras	2017-2018	2018-2019	2019-2020	2020-2021
<i>Manual_Total vs cosechados</i>	0,827	0,715	0,759	0,672
<i>Manual_Lado_Imagen vs cosechados</i>	0,701	0,598	0,569	0,531
<i>Manual_lado_Imagen vs Etiquetados</i>	0,790	0,720	0,850	0,843
<i>Manual_total vs etiquetados</i>	0,725	0,757	0,763	0,808
<i>Etiquetados vs cosechados</i>	0,629	0,609	0,611	0,574
<i>Detectados vs Manual_Lado_Imagen</i>			0,684	

Tabla 4: Correlaciones entre frutos contados manualmente al momento del pronóstico del lado de la imagen (Manual_lado_imagen); los frutos contados manualmente al momento del pronóstico de todo el árbol (Manual_Total); los detectados por el ojo humano en las imágenes (Etiquetados), los detectados por los MPDO (Detectados) y los cosechados (Cosechados).

Otro aspecto interesante es que las correlaciones de los frutos cosechados versus los conteos manuales del total de la planta siempre fueron mayores que los valores obtenidos entre los frutos cosechados versus los conteos manuales de la planta sólo del lado de la imagen. Esto pone en evidencia la necesidad de capturar imágenes en ambos lados de la planta para lograr mejores predicciones. Esto se contrapone a lo hallado por Linker (2016) quien concluye que utilizar fotos de un solo lado del árbol no empeora los resultados. Dorj et al., (2017), utilizan tres y hasta cuatro fotos por árbol para detectar y contar frutos a partir de imágenes en cítricos, logrando R^2 de 0,92. Es decir, el mayor número de imágenes por árbol podría mejorar este aspecto.

Por otra parte, el coeficiente de correlación entre el conteo manual de frutos por árbol del mismo lado en que se tomó la imagen y los detectados por el modelo se calculó para todas las temporadas en conjunto en el caso de manzanas alcanzando un valor de 0,887. Esto demuestra

que el modelo evaluado presentó una gran capacidad de detección de manzanas en las imágenes analizadas, representando una alternativa muy interesante a los métodos tradicionales de conteo manual. Črtomir et al., (2012), en su modelo híbrido de predicción de cosecha logró resultados similares a los obtenidos en este trabajo para manzanas cv ‘Golden Delicious’ (entre 0,73 y 0,83), y para el cv ‘Braeburn’ mostraron resultados más variables, (entre 0,51 y 0,78). Por otra parte, Farjon et al., (2019), que trabajaron con redes neuronales convolucionales para la detección de flores en manzana cv. ‘Golden Delicious’ lograron correlaciones entre las estimaciones del algoritmo y las realizadas por un grupo de personas expertas en la temática de entre 0,78 y 0,93. Otros autores que trabajaron con imágenes de manzanas cv ‘Golden Delicious’ y ‘Gala’ tomadas en cinco momentos diferentes del periodo de crecimiento y maduración de los frutos fueron Stajnko et al., (2005), y lograron correlaciones entre 0,71 y 0,89 en ‘Golden Delicious’ y 0,73 y 0,91 en ‘Gala’, según el momento de adquisición de las imágenes. Considerando que los mayores valores corresponden con los frutos más cerca de la madurez, se puede decir que estos resultados son similares a los obtenidos en esta tesis. Hočevcar et al., (2014) trabajaron en la estimación de ramilletes florales en manzanas a partir de imágenes tomadas de día y de noche con dos tipos de cámaras, digital e industrial. Los R^2 obtenidos para las flores contadas y estimadas a partir de las fotos nocturnas fueron entre 0,59 y 0,53 y para las fotos diurnas entre 0,52 y 0,20 para la cámara industrial y digital respectivamente. Los bajos valores obtenidos por dichos autores pueden deberse a que utilizaron algoritmos desarrollados con métodos tradicionales de programación, que incluyeron transformación del espacio de colores y procesos de binarización. Este tipo de procedimientos es más sensible a cualquier cambio de iluminación o condición en la que se tomaron las imágenes.

En el caso de las peras, los coeficientes de correlación entre el conteo manual de frutos por árbol del mismo lado en que se tomó la imagen y los detectados por el modelo fue de 0,694. Este valor es menor al obtenido para manzanas y esto puede deberse, por un lado, a que el conteo manual en los arboles de pera es más difícil de realizar por el tamaño del árbol, y por el follaje que es más abundante y de un color muy similar a los frutos. Por otro lado, otro aspecto a considerar, es algunas imágenes tienen ramas de árboles vecinos y los frutos que están sobre dichas ramas son detectados también por el modelo, pero no fueron contemplados en el conteo manual.

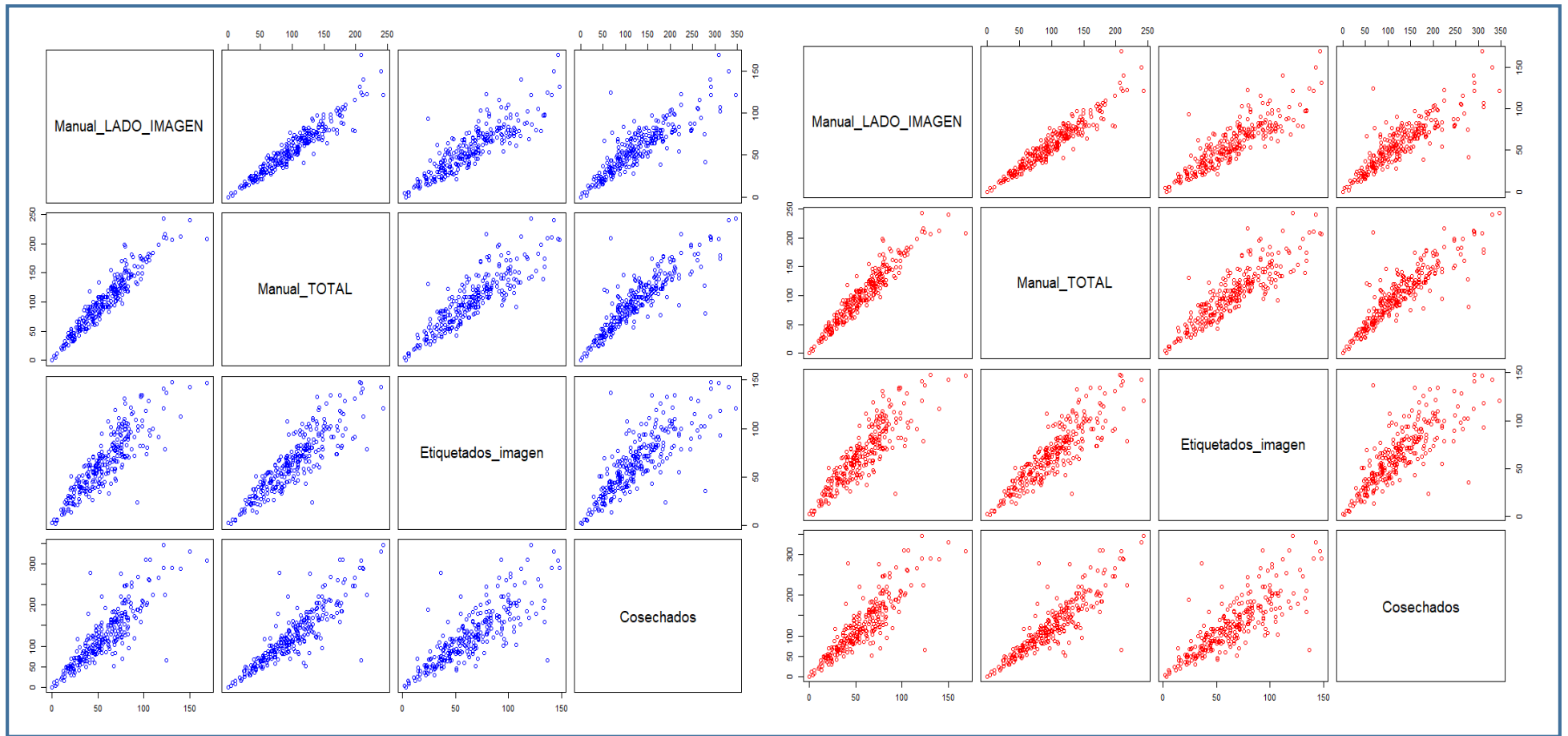


Figura 14: Diagramas de dispersión entre frutos contados manualmente del lado que se tomó la imagen, los contados manualmente en el total de la planta, los etiquetados en la imagen y los cosechados. **Izquierda:** Manzanas. **Derecha:** Peras

Si bien los gráficos y coeficientes de correlación muestran que existe una relación entre cada par de variables, dichas relaciones podrían evaluarse además mediante ajustes de modelos de regresión que consideren que todas las variables son aleatorias y permitan realizar estimaciones. Sin embargo, en esta tesis para realizar la estimación del número medio de frutos para realizar pronósticos de producción, se utilizará el cálculo del coeficiente de fruto oculto, el cual se desarrollará en el capítulo cuatro.

2.3.3. Estimación del tamaño de frutos

Las correlaciones obtenidas entre los diámetros medidos manualmente y los estimados a partir de imágenes fueron de 0,89 para los frutos de laboratorio, 0,73 para las manzanas a campo y 0,80 para las peras a campo (Figura 15). En los frutos analizados en laboratorio se observó una correlación mayor que en los frutos analizados a campo, lo cual puede deberse a que la ubicación y la posición de los frutos es mucho más variable en las plantas en condiciones naturales que en condiciones de laboratorio. Además, la distancia en el árbol de los frutos que se posicionan lejos del objeto de dimensiones conocidas utilizado como escala, puede ser una variable que aumente el error, por lo cual sería útil en futuras investigaciones medir solamente aquellos frutos que estén cercanos a dicho objeto, o bien, colocar más objetos de dimensiones conocidas por imagen.

En las mediciones a campo, los resultados obtenidos para peras fueron un poco superiores a los hallados para manzanas. Esto podría deberse a que el diámetro medido manualmente es el resultado del promedio de dos mediciones y en manzanas es muy común que exista mucha diferencia entre éstas, mientras que en la imagen solo se considera un solo diámetro sin contemplar dicha diferencia.

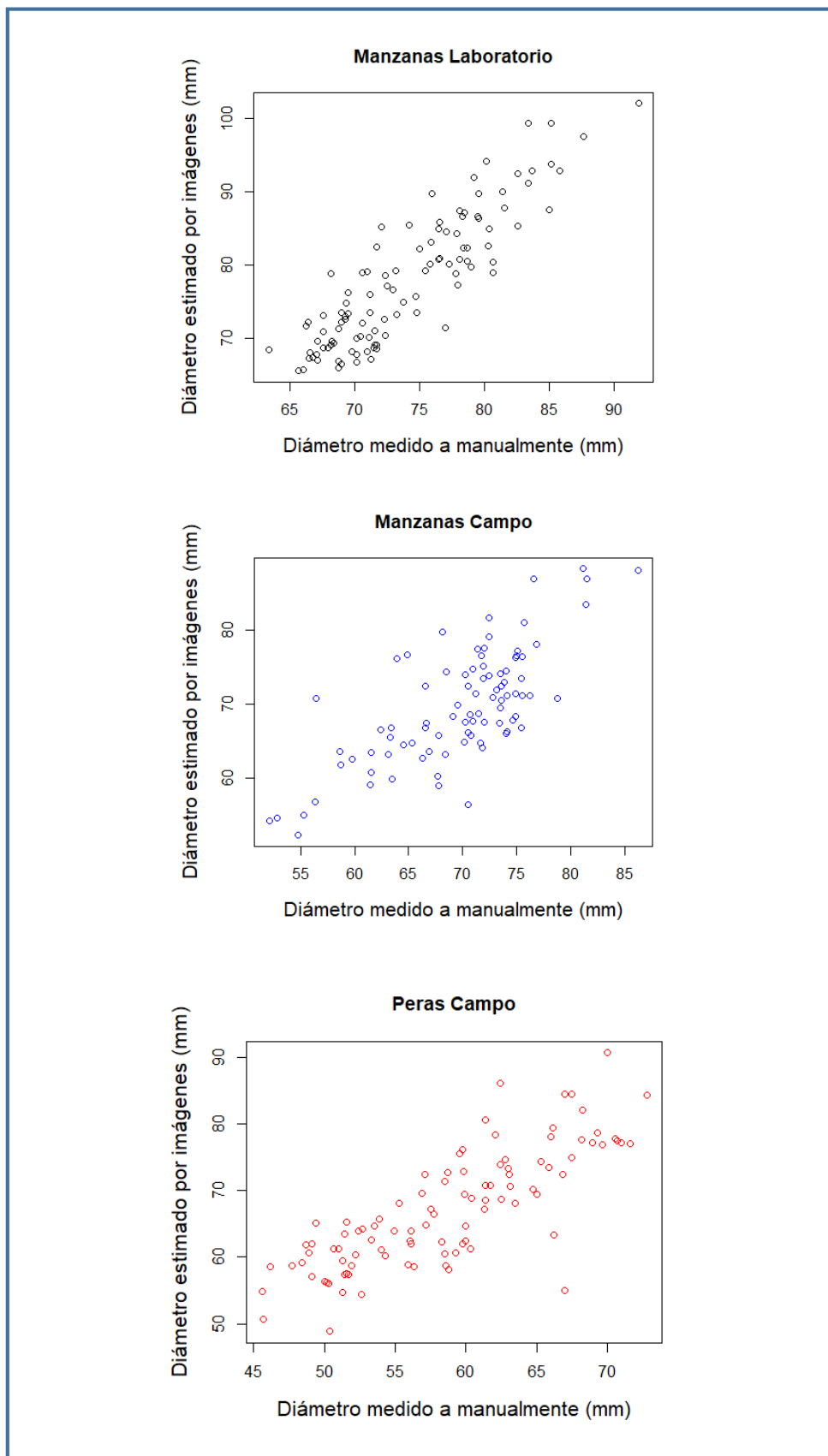


Figura 15: Diagramas de dispersión de las mediciones manuales de los diámetros de los frutos versus sus estimaciones a partir de análisis de imágenes.

Stajanko et al., (2005) también han trabajado en la estimación de diámetros de frutos en diferentes momentos a partir de imágenes en manzanos cv. ‘Golden Delicious’ y ‘Gala’ y las correlaciones que obtuvieron fueron muy variables, con valores que fueron desde 0,19 a 0,79 para ‘Golden Delicious’ y de 0,34 a 0,88 para ‘Gala’. Los diferentes valores se corresponden a los distintos momentos del crecimiento de los frutos (comenzando luego del raleo químico y finalizando a cosecha) y a las distintas temporadas. Una de las razones por la cual los coeficientes de correlación fueron bajos es que el algoritmo utilizado por los autores se basa en la medición del segmento más largo y además consideraron que el raleo tuvo un efecto nivelador de los frutos en los árboles, por lo cual pequeñas desviaciones del diámetro resultaron en una gran caída de la correlación. Algunos de los valores hallados por estos autores fueron similares a los hallados en esta tesis para manzanas y peras.

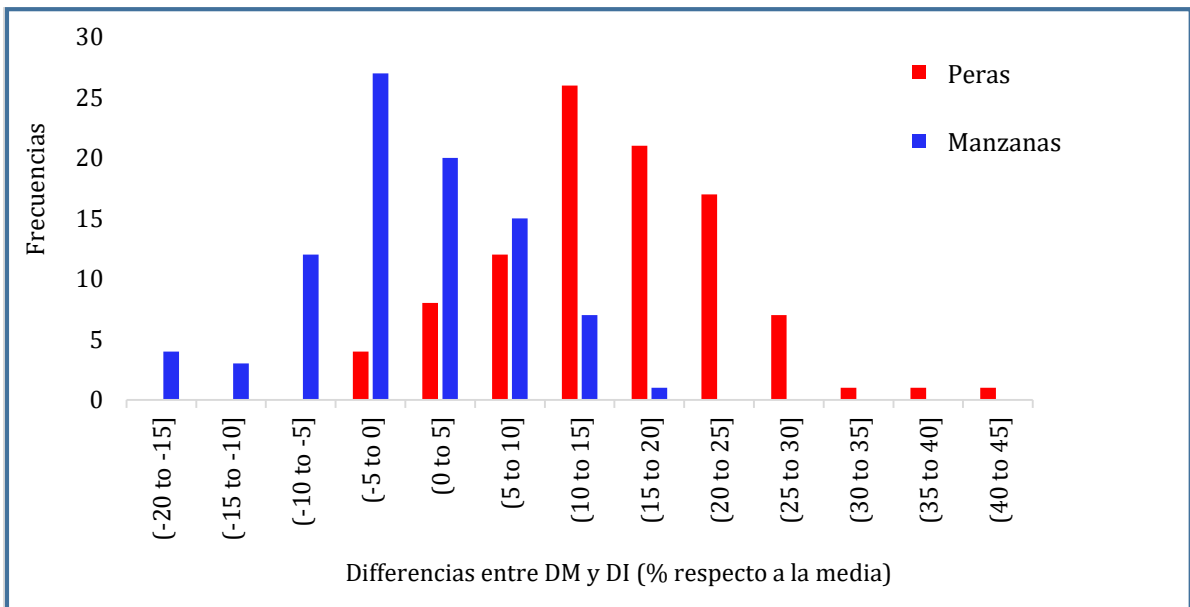


Figura 16: Gráfico de frecuencias absolutas para las diferencias entre los diámetros obtenidos por imágenes (DI) y los medidos manualmente con calibre digital (DM).

Stajanko et al., (2009) también encontraron que en manzanas cv. ‘Jonagold’ y ‘Golden Delicious’ existe una sobreestimación de los diámetros cuando las imágenes son tomadas con los frutos más jóvenes, mientras que a medida que los frutos crecen, existe una subestimación de los diámetros. Esto no concuerda con lo hallado en esta tesis para peras, ya que se utilizaron frutos maduros, listos para ser cosechados y las estimaciones fueron siempre mayores a las mediciones manuales, excepto en tres frutos. En manzanas, en cambio, los resultados fueron

más variables, mostrando tanto subestimación como sobreestimación de los diámetros en proporciones similares (Figura 16).

Tabla 5: Prueba t de comparación de medias de los diámetros de frutos medidos manualmente (grupo 1) y los estimados a partir de imágenes (grupo 2) para manzanas en campo.

	<i>N1</i>	<i>N2</i>	<i>Media1</i>	<i>Media 2</i>	<i>T</i>	<i>P (bilateral)</i>
Manzanas Campo	89	89	69,98	69,74	0,21	0,8366

Tabla 6: Pruebas Wilcoxon de comparación de medias de los diámetros de frutos medidos por imágenes (grupo 1) y los medidos manualmente (grupo 2), para manzanas en laboratorio y peras a campo.

	<i>N1</i>	<i>N2</i>	<i>Media1</i>	<i>Media2</i>	<i>W</i>	<i>P (bilayteral)</i>
Manzanas	106	106	77,98	74,13	12.585,5	0,0037
Laboratorio						
Peras campo	98	98	67,13	58,55	12.201,0	<0,0001

Por otra parte, las comparaciones entre las medias de los diámetros estimados por imágenes y los medidos manualmente resultaron significativos para manzanas en laboratorio y peras a campo y no significativos para manzanas a campo (Tabla 5 y Tabla 6). Esto quiere decir que la media de los diámetros estimados por imágenes y medidos manualmente en manzanas a campo no fueron diferentes. Este resultado concuerda con los obtenidos por Stajniko et al. (2005), que encontraron que los diámetros promedios estimados por imágenes para los cv ‘Golden Delicious’ y ‘Gala’ fueron prácticamente iguales a los obtenidos manualmente. Por su parte, los diámetros estimados por imágenes para peras a campo y manzanas en laboratorio presentaron medias diferentes a la de los diámetros reales, y en este caso los resultados se contraponen con los obtenidos para manzanas a campo y los hallados por Stajniko et al., (2005).

2.4. CONCLUSIONES

El número de frutos presentes en los árboles, tanto de peras como de manzanas, pudo ser estimado a partir de imágenes con una precisión aproximada de 0,85. Estos resultados pueden mejorarse si se aumenta la cantidad de datos de entrada a los modelos utilizados con más imágenes y etiquetas de frutas tomadas en la mayor diversidad de condiciones posibles (diferentes ángulos, distancias al foco, estados de crecimiento de los frutos).

La posibilidad de estimar los frutos a partir de imágenes, reduce el tiempo de mano de obra destinada a dicha tarea, y esto permitiría aumentar el número de árboles muestreados para realizar los pronósticos.

La estimación del tamaño medio de manzanas realizada a partir de imágenes de frutos a campo no presentó diferencias significativas respecto a la media real de los diámetros medidos manualmente. Contrariamente, en peras, la estimación del tamaño de los frutos a partir de imágenes no resultó muy precisa, variando mucho en función de la cercanía de los frutos al elemento de dimensiones conocidas, y también según la posición que tome el fruto en el árbol y cuan oculto quede por otras partes de la planta.

Los resultados obtenidos en este capítulo responden a la hipótesis uno de la presente tesis, la cual plantea que es posible contar y medir el tamaño de los frutos de pepita en tiempo real, a partir de sensores proximales. En este capítulo se observó que sí fue posible contar los frutos de pepita, pero no fue tan precisa la medición de tamaños a partir de sensores.

3.CAPITULO III: CURVAS DE CRECIMIENTO

3.1. INTRODUCCION

3.1.1. Estimación del tamaño final de los frutos

Como se mencionó previamente, para poder realizar estimaciones de la producción, es necesario estimar el tamaño final que alcanzarán los frutos al momento de cosecha. Es decir, el rendimiento de un árbol, no solo depende del número de frutos que éste posee, sino que también depende del peso que tendrán los mismos al momento de la cosecha (Jiménez y Royo Días, 2003).

Estimar el tamaño de la fruta con algunas semanas de anticipación es de gran utilidad ya que permite definir estrategias de manejo en pos cosecha de las tareas de empaque, de conservación, y de comercialización ya que existe un mercado diferencial de precio de la fruta según su tamaño (Giménez y Tassile, 2015). Diferencias en algunos gramos en el peso de los frutos pueden representar grandes pérdidas en los retornos por exportación (Mcartney et al., 1996). Es por ello que la predicción de la distribución final de tamaños de la fruta es una práctica difundida en muchas regiones de producción de frutales de pepita (Tassile, 2020).

Existen dos alternativas diferentes para predecir el tamaño de los frutos al momento de la cosecha: las curvas de crecimiento y los modelos estocásticos. Estos últimos se basan en el supuesto de que existe una relación entre la carga de las plantas y el tamaño final de los frutos. La carga se puede medir de forma sencilla con antelación a la cosecha y por lo tanto se podrían realizar inferencias del tamaño final de los frutos. Estos modelos, no solo requieren la estimación de la relación peso del fruto y carga del árbol, sino que además es necesario conocer la distribución de probabilidades de las cargas para obtener el modelo de distribución de probabilidades de los pesos (Avanza, 2010). En esta región resulto difícil determinar la relación peso carga debido a la alta heterogeneidad en el manejo los montes frutales, por lo cual las estimaciones a partir de modelos estocásticos no resultaron tan precisas (Bramardi et al., 2005). Por su parte, las curvas de crecimiento intentan describir el patrón de crecimiento de los frutos a través de modelos de regresión generalmente no lineales del tipo sigmoideo.

3.1.2. Curvas de crecimiento

Los pronósticos de producción que utilizan curvas de crecimiento han demostrado ser muy útiles ya que además de la predicción cuantitativa de la producción total potencial, pueden estimar en parte la calidad de la misma al separarla por tamaños comerciales (Giménez, 2012). En manzanas se han presentado diversos patrones para describir el crecimiento de frutos. Se han utilizado curvas dobles sigmoideas en manzanas cv ‘Golden Delicious’ y ‘Cox’s Orange Pippin’ para representar el crecimiento de los frutos con tres fases bien definidas; fase I: en la cual se produce la máxima tasa de crecimiento, la cual se corresponde al periodo de división celular; fase II: en la cual el crecimiento se reduce marcadamente; y fase III: se produce el alargamiento y crecimiento de las células (Magein, 1989). Se ha propuesto también el modelo expolineal de crecimiento de Goudriaan, y Monteith, (1990) en los cultivares ‘Empire’ y ‘Golden Delicious’ con tres parámetros definidos: la tasa máxima de crecimiento relativo, la tasa máxima de crecimiento absoluto y la intersección de la fase lineal de crecimiento (Lakso et al., 1995). Por otra parte, se ha propuesto una función exponencial negativa con diferentes parámetros según el cultivar en “Gala Brookfield”, “Golden Delicious”, “Braeburn” y “Fuji Kiku 8” (Zadravec et al., 2014). Este modelo presenta un fuerte incremento en el diámetro de los frutos al inicio, seguido por un crecimiento menor hacia el final del desarrollo de los frutos. También se han ajustado funciones cuadráticas que utilizan el diámetro y el cuadrado del mismo para la estimación del tamaño de frutos en cultivares “Jerseymac”, “Galaxy Gala” y “Braeburn” (Atay, et al., 2009). En el Alto Valle de Río Negro el crecimiento de manzanas cv “Red Delicious” y “Granny Smith” ha sido descrito mediante patrones sigmoideos de crecimiento ajustados a partir de mediciones sucesivas del diámetro de frutos (Stangafarro et al., 2001).

Por su parte, en pera ‘William’s’, Rodríguez (2011) ajustó modelos matemáticos del tipo exponenciales para describir el peso de frutos de en función de los días después de plena floración (DDPF) para las condiciones del Alto Valle de Río Negro y Neuquén. Otros autores han ajustado modelos de crecimiento para la misma región, pero utilizando regresiones logísticas en pera cv “Williams” (Bramardi et al., 1997; Tassile, 2020); “Red Sensation” (Garriz et al., 2006); “Beurre D’Anjou” (Gimenez, 2020) pero utilizando el diámetro de los frutos en función de los DDPF. Para representar el crecimiento de peras cv “Williams” el mejor modelo ha sido la tercera parametrización según Ratkowsky (1983) Este modelo tiene la ventaja de que sus parámetros tienen una interpretación práctica directa, por lo cual, son modelos ampliamente utilizados en estudios de crecimiento de frutos (Fernández et al., 2014;

Bóbeda, 2022). Por lo tanto, los modelos que se ajustarán en esta tesis para describir el crecimiento de frutos tanto de peras como manzanas, estarán todos basados en la Ecuación 4 pero contemplarán diferentes factores fijos y aleatorios y/o covariables.

Ecuación 4: curva logística de crecimiento de frutos

$$Y = \frac{1}{\beta_1 + \beta_2 * \beta_3^{DDPF}}$$

donde:

β_1 : es el parámetro relacionado inversamente a la asíntota superior (asociado al tamaño final que pueden alcanzar los frutos)

β_2 : es el parámetro que relaciona la asíntota superior (S) con la asíntota inferior (I) mediante la función: $e^{I/S}$

β_3 : es el parámetro relacionado con la tasa de crecimiento desde los valores iniciales (β_2) y los valores finales (β_1)

DDPF: días después de plena floración

3.1.3. Incorporación de variables climáticas e índices verdes a las curvas de crecimiento

Cada ajuste modela patrones de crecimiento para predecir tamaños para una variedad en una zona determinada (Bramardi et al., 2006; Bóbeda, 2022), por lo cual no sería apropiado extrapolar dichos modelos en diferentes condiciones de cultivo. Esto se debe a que el tamaño final esperado de los frutos varía según el cultivar, y también depende en gran medida de otros factores como las condiciones climáticas, las técnicas de manejo y el número de frutos que hay en el árbol (Jiménez y Royo Días, 2003).

El clima es uno de los factores más importantes a considerar y se relaciona tanto directa como indirectamente con los cultivos regulando el crecimiento, el rendimiento y la calidad de la producción (de Oliveira Aparecido et al., 2016). Se han reportado diferentes relaciones entre las condiciones climáticas y ambientales y los rendimientos de cultivos. En granos por ejemplo, las temperaturas máximas y mínimas influyen en los rendimientos (Kumar et al., 2014); en olivos se ha propuesto un modelo para realizar estimaciones tempranas de la cosecha

por árbol, basado en la temperatura, las precipitaciones mensuales y la emisión de polen que se utilizó como un bioindicador (Galán et al., 2004); en naranjos se utilizaron variables meteorológicas para ajustar un modelo de estimación del número de frutos por planta (Pasqua et al., 2007); en arándanos se ajustaron modelos lineales múltiples y la incorporación de variables climáticas mejoró las estimaciones (Salvo et al., 2012); entre otros. Por su parte, en manzanas cv 'Royal Gala', Stanley et al., (2000) consideran que tanto la carga del árbol como la temperatura de los primeros 50 DDPF determinan el tamaño final de los frutos. En pera cv 'William's', Rodríguez (2011) concluye que la temperatura posee un papel regulador en el crecimiento de los frutos entre los 10 y 20 DDPF (período de máxima multiplicación celular) y en los primeros 60 DDPF actúa linealmente sobre la velocidad de crecimiento del fruto. Además, ha surgido un gran interés en el uso de índices vegetales que pueden aportar información interesante y se calculan a partir de información espectral obtenida de sensores remotos, siendo el NDVI uno de los más utilizados (Gemtos et al., 2013). Este índice surge de la relación entre bandas del rojo y del infrarrojo (Ecuación 5) y es una medida de la reflectancia, la cual es sensible al color de vegetación sana. Su intervalo de valores posibles es entre -1 y 1. Los valores negativos generalmente se asocian a zonas de agua o nieve, valores positivos próximos al cero se asocian a zonas rocosas o suelo desnudo y puede llegar a valores próximos a 0,3 a medida que adquiere algo de vegetación. Luego de este valor ya se asocia a la presencia de vegetación, aumentando a medida que las plantas son más frondosas y sanas hasta adquirir valores próximos a uno.

Ecuación 5: cálculo del NDVI

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

Donde:

NIR: representa la banda del infrarrojo cercano

RED: representa el valor obtenido para la banda del rojo

Esta información puede ser utilizada en las estimaciones de rendimientos debido al hecho de que la reflectancia espectral de las copas de los árboles está determinada por las propiedades biofísicas y bioquímicas del cultivo que afectan la fotosíntesis (Asner, 1998). Un cambio en

el vigor, densidad, salud o productividad de las plantas, afecta las propiedades ópticas de las copias, y por lo tanto el desarrollo y crecimiento de los cultivos podría ser monitoreado mediante el uso de imágenes satelitales (Nagy et al, 2018). En uvas, se han ajustado modelos de estimación de rendimientos a partir del vigor de las copas (obtenido a partir del índice NDVI), y de variables asociadas a la fertilidad del viñedo como el número de yemas, brotes y racimos, obteniendo altas correlaciones (Martínez-Casasnovas y Bordes, 2005). En maíz y trigo también se han desarrollado metodologías de estimación de cosecha basadas en NDVI y además se evaluó si dicho modelo es capaz de estimar las pérdidas causadas por la sequía (Nagy et al, 2018). También se han evaluado la precisión de diferentes metodologías de obtención del NDVI comparándose con valores de clorofila obtenidos con SPAD (soil plant analysis development) en la predicción del estrés en plantas de manzana y cítricos (Glenn y Tabb, 2019). En estos últimos también se ha trabajado en la delimitación de zonas de producción a partir de indicadores como rendimiento a cosecha, volumen de copa de los árboles, elevación del suelo, conductividad aparente y NDVI. Como resultado el rendimiento se correlacionó positivamente con el volumen de copa, el NDVI y la conductividad aparente, y negativamente con la elevación (Mann et al., 2011). Por su parte, Gemtos et al., (2013) compara mapas de rendimiento y mapas de NDVI en cultivos de manzana y vid y encontró semejanzas en la variabilidad de ambos mapas, es decir, altos valores de NDVI se corresponden con áreas de altos rendimientos.

Otro índice que se puede obtener de información proveniente de imágenes satelitales es el índice de estrés por humedad (MSI por sus siglas en inglés: Moisture Stress Index). Este, es una medida de reflectancia sensible al aumento de contenido de agua de las hojas. A medida que el contenido de agua en la copa de las plantas aumenta, la fuerza de la absorción alrededor de 1599 nm aumenta. La absorción a 819 nm casi no es afectada por el cambio de contenido de agua, por lo que se utiliza como la referencia. El MSI tiene relación inversa con los otros índices de vegetación de agua, los valores más altos indican un mayor estrés hídrico y por lo tanto un menor contenido de agua. El valor de este índice varía de cero a más de tres y se calcula mediante la Ecuación 6. El intervalo común para la vegetación verde es de 0,4 a 2 (Hunt y Rock, 1989; Ceccato, 2001).

Ecuación 6: cálculo del MSI

$$MSI = \frac{SWIR}{NIR}$$

Donde:

NIR: representa la banda del infrarrojo cercano

SWIR: representa el valor del infrarrojo medio

Este índice ha sido utilizado en algunos estudios, por ejemplo, Villatoro (2020), generó una metodología para usar índices de humedad y monitorear esta variable como indicador del momento óptimo de cosecha de caña de azúcar; otros autores llegaron a la conclusión de que los índices relacionados con la humedad obtenidos de Sentinel-2 están fuertemente relacionados con el grado de degradación de bosques (Quang et al., 2022).

En este sentido resulta interesante incluir en los modelos de crecimiento de frutos con fines de pronóstico, aquellas covariables asociadas a condiciones climáticas e índices de vegetación que potencialmente puedan aportar información valiosa a los mismos. Dichas variables pueden ser contempladas mediante el uso de modelos mixtos no lineales (NLMM por sus siglas en inglés: No Linear Mixed Models), donde las covariables ambientales e índices verdes surgen como candidatas para combinar los efectos de cada sitio y de cada ciclo de crecimiento estudiados.

3.1.4. Modelos mixtos no lineales

Un NLMM es aquel en el cual la función de regresión de los factores fijos y aleatorios no es lineal. Esto generalmente implica que dicho modelo utilice menos parámetros que sus modelos competitivos polinomiales, volviéndolos más flexibles, parsimoniosos y más fáciles de interpretar. Además, las predicciones obtenidas de un modelo no lineal se extienden de manera más confiable fuera del rango observado de datos. Por otra parte, una desventaja de los modelos no lineales es que no tienen una expresión de forma cerrada. Como consecuencia, es necesario utilizar una función de probabilidad aproximada para la estimación de parámetros, lo que lleva a un cálculo más intensivo de algoritmos de estimación y resultados de inferencia menos fiables. A diferencia de los modelos lineales, los no lineales requieren estimaciones iniciales para los coeficientes de efectos fijos (Pinheiro y Bates, 2000).

Por otra parte, al ser mixtos permiten modelar conjuntos de datos en los que las observaciones no son independientes (Balzarini et al., 2004). Permiten incluir tanto parámetros de efectos fijos, que pueden estar asociados con una o más covariables continuas o categóricas, como también efectos aleatorios según diferentes niveles de agrupamiento, es decir, tienen la ventaja

de analizar estructuras de datos agrupados y modelar la matriz de covarianza de datos correlacionados (Tassile, 2020). Además, la inclusión de covariables a NLMM generalmente permite explicar la variabilidad entre grupos simplificando los efectos aleatorios del modelo y permitiendo una mejor comprensión de los mecanismos de respuesta. Asimismo, permiten contemplar la correlación que existe cuando se trabaja con mediciones repetidas de la variable respuesta sobre un mismo individuo a través del tiempo. Esto puede hacerse de forma explícita, a través de la estimación de una matriz de covarianza, o implícitamente, mediante la adición de efectos aleatorios en el modelo incorporando el efecto aleatorio de la unidad experimental (Balzarini et al., 2004). De esta manera se busca evitar que las varianzas estimadas del error de los parámetros resulten sesgadas, ya que estos estimadores son frecuentemente usados en diferentes pruebas de hipótesis para seleccionar el mejor conjunto de covariables, por lo tanto, el ajuste final del modelo resulta sensible a la especificación de la estructura de covarianza del error (Tassile, 2020).

Por lo cual, partir de la aplicación de NLMM, se pretende, no sólo caracterizar el crecimiento típico de los frutos; sino también contemplar la incorporación de efectos aleatorios a diferentes niveles de agrupamiento en los coeficientes (fruto, planta, sitio, temporada), modelar la correlación de los errores, e incorporar covariables asociadas a condiciones ambientales e índices verdes que resulten de interés para caracterizar a la temporada de crecimiento. En pera cv ‘William’s’ se han propuesto NLMM para montes frutales del Alto Valle de Río Negro. Estos modelos consideraron información de naturaleza longitudinal y transversal, y además incorporaron covariables ambientales representativas de distintos momentos críticos del crecimiento de los frutos, resultando favorable en todos los casos evaluados dicha incorporación (Tassile, 2020).

3.1.5. Aspectos estadísticos del ajuste y evaluación de los modelos mixtos no lineales

La construcción de un modelo mixto involucra algunas cuestiones diferentes a los modelos lineales y no lineales, entre ellas: determinar qué efectos deben asociarse a componentes aleatorias y cuáles deben ser puramente fijos; el uso de covariables que expliquen la variabilidad modelada en los efectos aleatorios y el uso de matrices de varianza-covarianza estructuradas para reducir la cantidad de parámetros. Si bien existen diferentes NLMM (paramétricos, no paramétricos, semiparamétricos y de modelado bayesiano), en esta tesis se

utilizarán los NLMM paramétricos con efectos aleatorios y términos de error distribuidos normalmente, siguiendo la metodología propuesta por Tassile (2020), cuya expresión general es la siguiente:

Ecuación 7: modelo a nivel de individuo

$$Y_{ij} = f(z_{ij}, \beta_i) + e_{ij}, \quad i = 1, \dots, m \quad e = 1, \dots, n_i$$

Ecuación 8: modelo a nivel de población

$$\beta_i = d(a_i, \beta, b_i)$$

Con e_{ij} y b_i variables aleatorias con:

$$E(e_{ij}) = E(b_i) = 0$$

$$\text{VAR}(e_{ij}) \sim N(0, R_i(\beta_i, \mathcal{E}, z_i))$$

$$\text{VAR}(b_i) \sim N(0, D)$$

En el modelo general b_i representa los efectos correspondientes a individuos particulares, y el modelo a nivel de individuo (Ecuación 7) describe la respuesta para individuos en particular, por lo que para caracterizar el comportamiento del individuo en interés dependerá tanto de β como de b_i . En esta tesis, para la estimación de parámetros fijos, siguiendo la metodología propuesta por Tassile (2020), se utilizarán particularmente dos variaciones del método de aproximación a la integral de Laplace: la versión máxima verosímil (REM) propuesta por Pinheiro y Bates (2000) que trata la similitud marginal de los modelos como una integral solo de los efectos aleatorios; y la versión de máxima verosimilitud restringida (REML) que integra sobre los efectos aleatorios y fijos asumiendo una densidad no informativa previa para la verosimilitud marginal del modelo. Por otro lado, los b_i son vectores aleatorios y la inferencia sobre ellos se vuelve más complicada. Caracterizar b_i implica predecir el valor que tomara una variable aleatoria de un individuo seleccionado al azar de una población, por lo cual se realiza una caracterización de b_i dado un Y_i . Esto se puede realizar a partir de los mejores

estimadores lineales insesgados (BLUPs) propuestos por Lindstrom y Bates (1990), conocido como método de linealización EBLUP que se puede encontrar en la función NLME de R (R core Team, 2018). Para cada nivel del factor de efecto aleatorio, el BLUP realiza una predicción de la ubicación dentro de la distribución normal de la que proviene el efecto aleatorio asociado al sujeto (Balzarini et al., 2004).

Para la selección de modelos candidatos para los datos es necesario evaluar la precisión de las estimaciones de los distintos términos del mismo, o poder realizar comparaciones entre diferentes modelos candidatos. Para los modelos lineales, en los cuales los efectos fijos son lineales, existen pruebas cuyos estadísticos tienen una distribución conocida para cualquier tamaño muestral. Contrariamente, en los modelos mixtos no existe una única prueba, sino que se usan aproximaciones relacionadas a las distribuciones asintóticas de dichos modelos (Tassile, 2020). El cociente de verosimilitud (LRT) y la prueba de Wald son dos pruebas asintóticas que se basan en el principio de verosimilitud. El LRT solo puede ser utilizado si el modelo se ajusta con el método máximo verosímil, mientras que Wald puede utilizarse con método máximo verosímil y máximo verosímil restringido.

También se utiliza los criterios de información o de verosimilitud penalizada como el Criterio de Información de Akaike (AIC) y el Criterio de Información Bayesiano (BIC). Si bien ambos criterios dependen del valor de verosimilitud del modelo, el AIC penaliza por el número de parámetros del modelo mientras que el BIC penaliza no solo por la cantidad de parámetros sino también por el logaritmo de la cantidad de datos de manera que este afecta más a los modelos con menor cantidad de datos (Giménez, 2020). Estos siguen el criterio de más pequeño es mejor, es decir, si se comparan dos o más modelos, se escogerá aquel que posea menor valores de AIC y BIC, pudiendo afirmar que un valor de dos a cinco unidades menor de criterios de información para un modelo, implica que tiene altas chances de resultar más adecuado (Konishi y Kitagawa, 2007). Algunos autores como Sober (2002), sugieren que AIC mide la capacidad predictiva mientras que BIC lo hace con la bondad de ajuste, por lo que, al proveer estimaciones de métricas diferentes no se deben considerar como criterios en competencia. En esta tesis los modelos a comparar tanto para peras como para manzanas, estarán todos basados en la Ecuación 4, pero contemplarán diferentes efectos fijos, aleatorios y covariables.

Por otra parte, antes de realizar inferencias, es necesario examinar el modelo a los fines de comprobar si los supuestos distribucionales son válidos para los datos. Estos supuestos son:

i) que los errores dentro de los individuos son independientes, normales e idénticamente distribuidos, con media 0 y varianza σ^2 , y son independientes de los efectos aleatorios.

ii) los efectos aleatorios se encuentran normalmente distribuidos, con media 0 y matriz de covarianza D e independientes para los diferentes individuos.

El primer supuesto se refiere al estudio de las dependencias entre los errores que son susceptibles de ser modeladas a través de estructuras de correlación. Se puede verificar mediante gráficos de residuales versus la variable independiente, valores observados versus predichos, gráficos de probabilidad normal para los residuales considerados junto a correlogramas y qqplots de residuales para evaluar correlaciones.

El segundo supuesto se puede verificar a partir de gráficos de probabilidad normal y de la identificación de valores extremos, y de los efectos aleatorios versus factores de diseño y variable independiente para el chequeo de supuestos de homogeneidad en la estructura de varianza de los b_i .

Una vez verificados los supuestos se pueden analizar los criterios de capacidad predictiva de los modelos. Estos criterios usan los valores predichos para construir métricas de capacidad predictiva y entre ellos se pueden mencionar: el coeficiente de determinación (R^2) que tiene la ventaja de proveer una interpretación simple e intuitiva (Liu et al., 2008); la raíz del cuadrado medio del error (RMSE); el sesgo medio (SM); el sesgo medio relativo (SMR); sesgo medio absoluto (SMA) y el sesgo medio relativo absoluto (SMRA). Todos ellos se utilizarán en esta tesis para evaluar la capacidad predictiva de los modelos ajustados.

3.2. MATERIALES Y METODOS

3.2.1. Obtención de los datos

3.2.1.1. Temporada 2017/2018

En la Estación Experimental del INTA seleccionaron al azar 15 plantas de pera y 5 plantas de manzana. En cada planta el día 27/10 se seleccionaron 8 frutos chicos (menores a 15mm), 8 frutos medianos (entre 15 y 20 mm) y 8 frutos grandes (mayores a 20 mm), resultando un total de 480 frutos. Dicha fecha se corresponde con 29 y 30 DDPF para manzanas y peras, respectivamente. Para determinar las categorías de tamaño se midieron los frutos más pequeños y los más grandes y luego se dividió el rango observado. Se midió semanalmente el diámetro de los mismos utilizando un calibre digital (Essex 150 mm, Stainless Hardened). El diámetro se midió en la zona ecuatorial de los frutos en dos lados diferentes y se obtuvo la media de ambas mediciones. Los frutos caídos fueron reemplazados por otro que tenga un diámetro similar. Las mediciones finalizaron una semana después de la cosecha.

3.2.1.2. Temporadas 2018/2019 y 2019/2020

Se seleccionaron 5 plantas al azar de pera y 5 plantas de manzana. En cada planta se seleccionaron 8 frutos chicos, 8 frutos medianos y 8 frutos grandes, resultando un total de 240 frutos. Se continuó con la misma metodología utilizada en la temporada anterior. Las mediciones comenzaron el 25/10 (se corresponde con 27 y 24 DDPF para peras y manzanas respectivamente en la temporada 2018/2019; y con 34 y 30 DDPF para peras y manzanas, respectivamente en la temporada 2019/2020). Las mediciones finalizaron el día de la cosecha.

3.2.1.3. Temporada 2020/2021

Durante esta temporada de crecimiento se adicionaron a las mediciones de la estación experimental de INTA, los cuatro montes frutales comerciales, resultando un total de tres montes frutales por especie (Tabla 1). En cada uno de ellos se utilizó la misma metodología descrita para la temporada anterior, es decir, 5 árboles por monte frutal y 24 frutos por árbol discriminados en tamaños grandes, medianos y pequeños. Las mediciones comenzaron el día

27 de octubre (se corresponde con 25 y 22 DDPF para peras y manzanas respectivamente) y finalizaron el día de la cosecha. Esta temporada además se midieron los diámetros al azar de 250 frutos aproximadamente en cada monte frutal a principios de diciembre para realizar la estimación del diámetro medio de fruto a cosecha para utilizar en los pronósticos y las estimaciones de distribuciones de frecuencias de tamaños según clases.

3.2.2. Obtención de datos meteorológicos

Para todas las temporadas de trabajo se contó con el registro de los datos meteorológicos obtenidos a partir de una estación automática modelo DAVIS Vantage Pro 2, inalámbrica, Mb5. Dicha estación registra con una frecuencia de diez minutos las temperaturas ambiente, máxima y mínima; la humedad ambiente; el rocío; la velocidad, intensidad y dirección del viento; las lluvias, la radiación solar y ultravioleta; entre otras.

Con estos datos se crearon algunos índices de acumulación de temperatura descritos por Tassile, (2020). Los índices calculados fueron los siguientes: i) TEMP_1: acumulación de temperaturas diarias que superan 16 grados centígrados desde el día 15 de Setiembre hasta el 15 de noviembre inclusive; ii) TEMP_2: acumulación de temperaturas diarias que superan los 16 grados centígrados a partir del 30 de setiembre hasta el 15 de noviembre inclusive; iii) TEMP_3: acumulación de temperaturas diarias que superan los 22 grados centígrados desde el 01 al 31 de diciembre inclusive.

3.2.3. Obtención de índices verdes a partir de imágenes satelitales

Para descargar y procesar las imágenes satelitales se utilizó la plataforma Google Earth Engine. Dicha plataforma posee varios petabytes y un conjunto de datos espaciales que permiten al usuario calcular y analizar una amplia variedad de geo procesos mediante algunos scripts específicos. Se utilizaron coordenadas geográficas EPSG 4326 para determinar el área de estudio, la cual está conformada por las plantas seleccionadas cada temporada para medir semanalmente los diámetros ecuatoriales de los frutos. En primer lugar, se cargó el área de estudio, y luego para dicha área se descargaron las imágenes satelitales del Sentinel 2 (agencia espacial europea) disponibles desde el 01 de noviembre al

01 de febrero de cada temporada. Dichas imágenes poseen una resolución de 10 m por pixel. Luego se calculó el NDVI y MSI (Ecuación 5 y Ecuación 6.) para cada uno de los casos.

Por último, se crearon 4 covariables asociadas a estos índices para incorporar a los modelos de crecimiento de cada especie. Dichas covariables son los valores de NDVI y MSI para los meses de noviembre y diciembre. Las covariables se nombraron NDVI_NOV, NDVI_DIC, MSI_NOV y MSI_DIC, respectivamente.

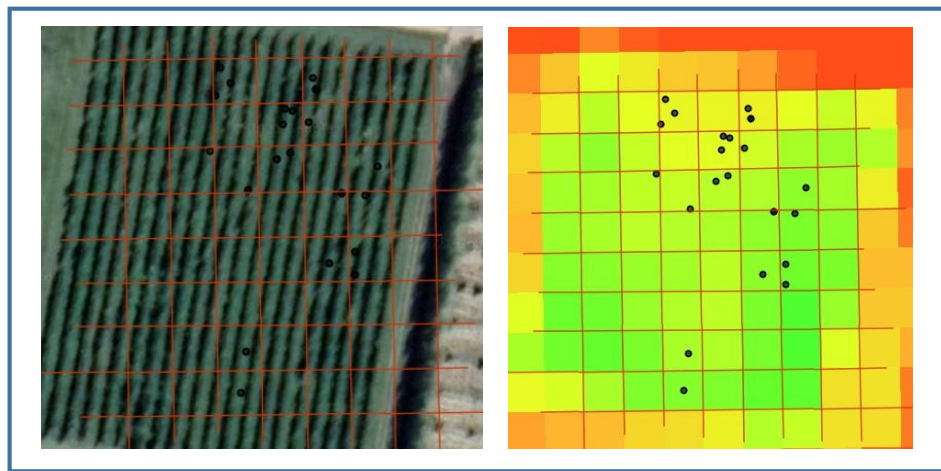


Figura 17: Izquierda: Imagen Sentinel 2 del cuadro de manzanas cv `Red chief` con grilla que indica el tamaño de resolución de pixel para el cálculo del NDVI. Derecha: índices NDVI para cada pixel de la grilla.

3.2.4. Ajuste de curvas de curvas de crecimiento

Una vez obtenidos todos los datos, se realizaron diagramas de dispersión de diámetros versus los DDPF para cada temporada y sitio para observar el comportamiento de los datos. Luego, se procedió a realizar las estimaciones de los modelos a través del paquete estadístico NLME del software R. Este método de estimación obtiene la verosimilitud a partir de una expansión de Taylor de acuerdo al método de linealización BLUP propuesto por Lindstrom y Bates, (1990) para obtener las estimaciones de los NLMM con efectos aleatorios. Los modelos obtenidos por NLME presentan efectos anidados o una estructura de efectos cruzados muy simples, presentado la posibilidad de modelar directamente la heterocedasticidad y la correlación de los residuales (Tassile, 2020).

Para cada especie se ajustó un modelo diferente. En ambos casos se fueron ajustando modelos sigmoideos mencionados, contemplando diferentes factores aleatorios en los distintos niveles de agrupamiento, para luego escoger el mejor modelo que se adapta a los datos. Los factores considerados fueron: *Tamaño de fruto*, *Sitio* (constituido por la combinación temporada y lugar), *Fruto dentro de Sitio*. Luego, se incorporaron las diferentes covariables asociadas a condiciones meteorológicas e índices verdes, para ver si realizan un aporte significativo a los modelos ajustados con el objetivo de que los modelos describan mejor el crecimiento en cada sitio. Para la comparación de modelos se utilizaron tanto criterios de información como pruebas de razón de verosimilitud, según el caso. Por último, una vez obtenidos los modelos finales, se realizó la calibración y validación de los mismos con datos no contemplados en el proceso de modelado y se calcularon indicadores de bondad de ajuste y criterios de predicción. En el proceso de calibración, tanto para la estimación de efectos aleatorios como para la estimación del diámetro predicho, se utilizó el método de BLUP.

3.3. RESULTADOS Y DISCUSIÓN

3.3.1. Análisis exploratorio de los datos

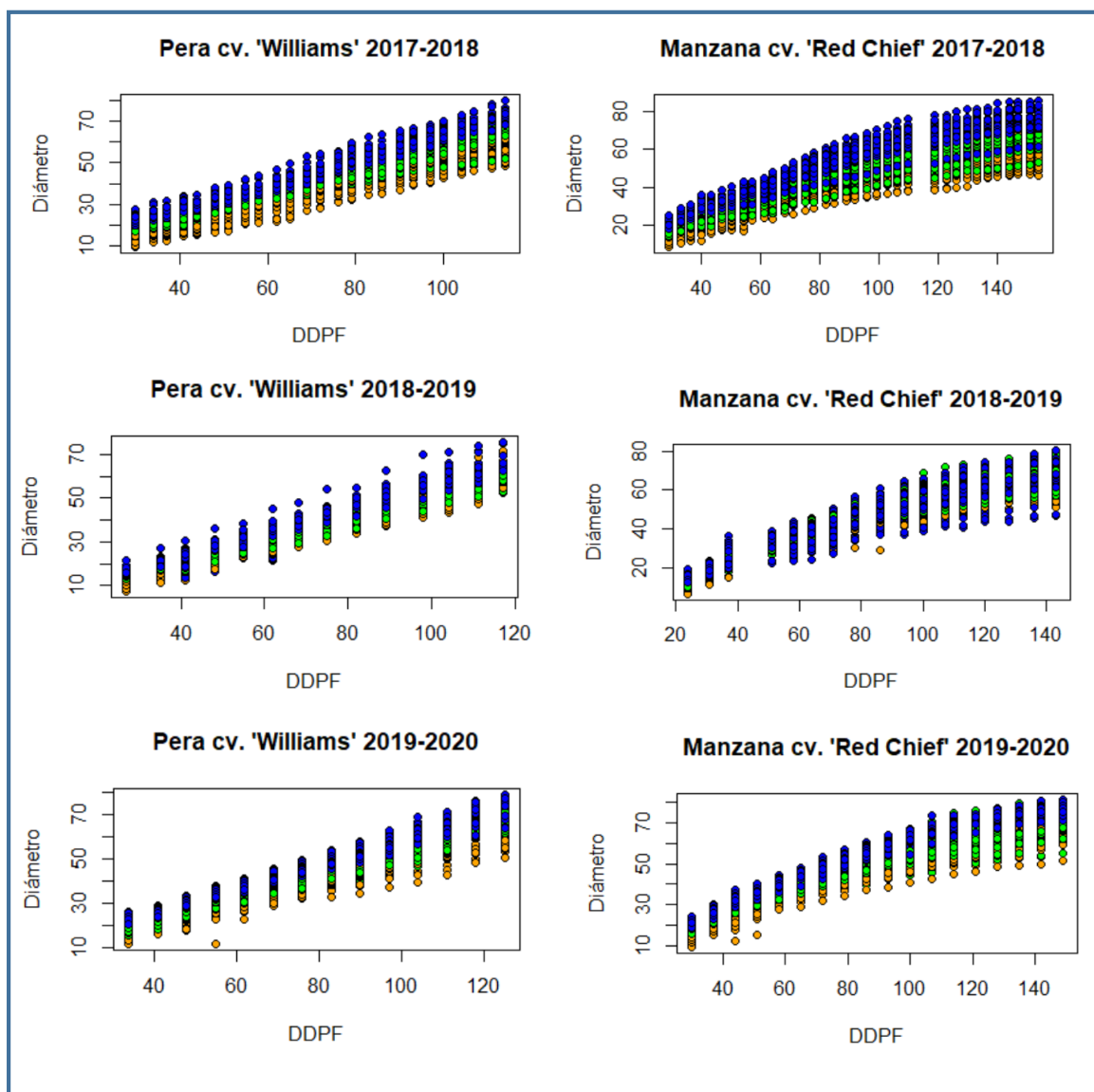


Figura 18: Diagramas de dispersión de los diámetros de los frutos a lo largo de los DDPF para cada especie durante las temporadas 2018-2018; 2018-2019 y 2019-2020 en INTA.. Los colores representan el tamaño de frutos: azul corresponde a grandes, verde a medianos y naranja a pequeños.

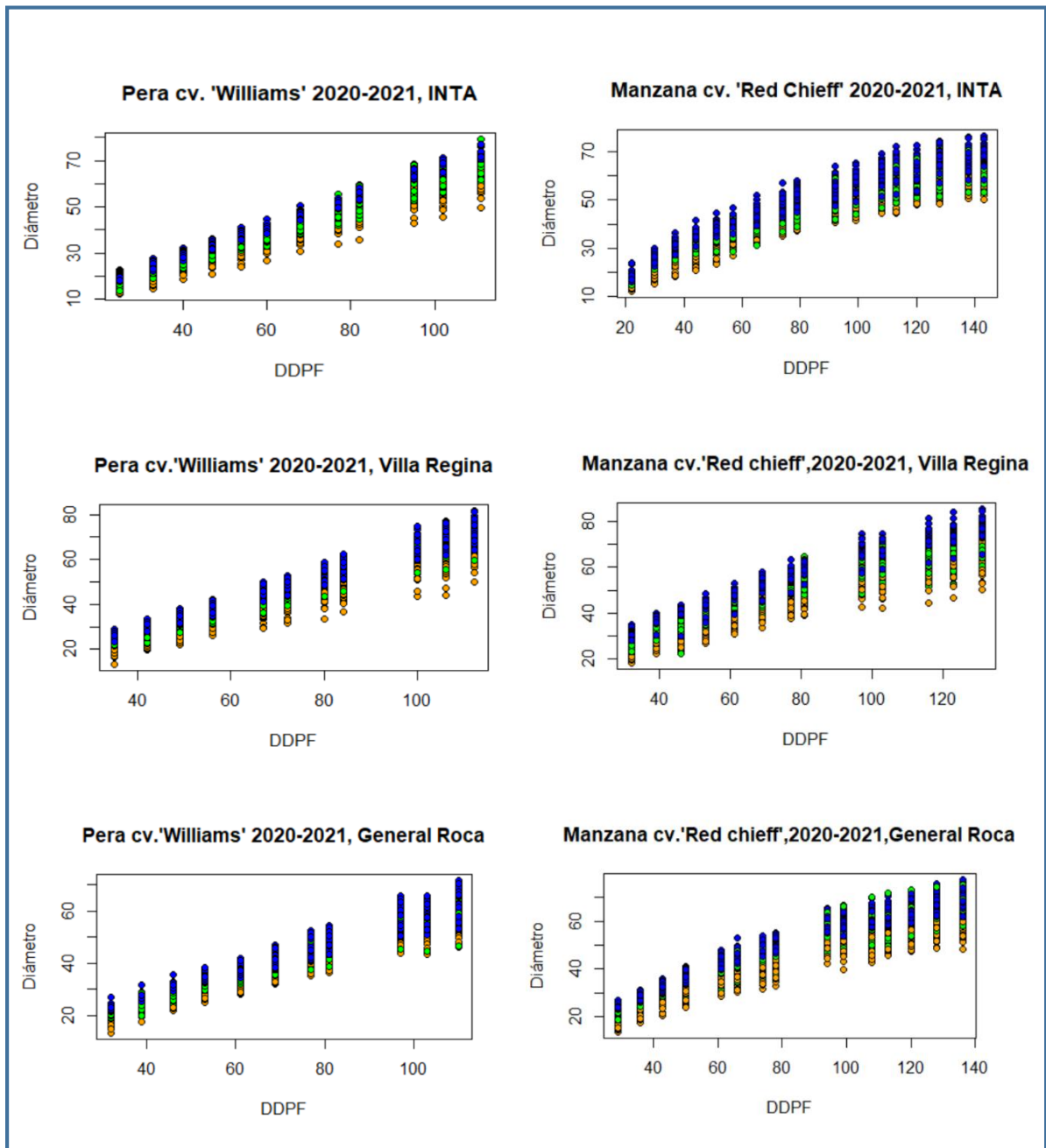


Figura 19: Diagramas de dispersión de los diámetros de los frutos a lo largo de los DDPF para cada sitio y especie de la temporada 2020-2021. Los colores representan el tamaño de frutos: azul corresponde a grandes, verde a medianos y naranja a pequeños.

3.3.2. Resultados de los ajustes de los modelos de crecimiento para peras

En primer lugar, se buscó la componente estructural más adecuada para el conjunto de datos bajo estudio. Esta expresa el diámetro ecuatorial en función de los días después de plena floración (DDPF), y la función propuesta se mencionó previamente en la Ecuación 4 (Bramardi et al. 1997). Con el objetivo de mejorar las condiciones de convergencia de los algoritmos de optimización, se utilizó la reparametrización (Ecuación 9) realizada por Tassile (2020).

Ecuación 9: modelo de crecimiento reparametrizado (Tassile, 2020)

$$Y = \frac{1}{\beta_1 * 0,01 + (\beta_2 * 0,01) (\beta_3 * 0,1)^{DDPF}}$$

Donde:

β_1 : es el parámetro relacionado inversamente a la asíntota superior (asociado al tamaño final)..

β_2 : es el parámetro que relaciona la asíntota superior (S) con la asíntota inferior (I) mediante la función: e^I/S

β_3 : es el parámetro relacionado con la tasa de crecimiento desde los valores iniciales (β_2) y los valores finales (β_2 :)

DDPF: días después de plena floración

Por lo tanto, para definir la componente estructural se plantearon distintos modelos que contemplaron los diferentes efectos evaluados a lo largo de las temporadas en estudio, es decir, el factor *sitio* (como combinación entre temporada y localidad), *fruto* y *tamaño de fruto*. Este último se consideró como factor fijo en todos los modelos analizados, y lo que se fue modificando fue la parte aleatoria. Se consideró el factor aleatorio fruto ya que permite contemplar, implícitamente, la correlación que existe cuando se trabaja con mediciones repetidas sobre un mismo individuo a través del tiempo (Balzarini et al., 2004) y el efecto sitio porque ya ha sido demostrada su implicancia en los modelos de crecimiento de peras (Bramardi, 1995; Tassile, 2020).

Los modelos ajustados fueron los siguientes:

✓ *Pera_1*: consideró como factor fijo el *tamaño de fruto* y como factor aleatorio el *fruto*.

- ✓ **Pera_2:** consideró como factor fijo el *tamaño de fruto* y como factores aleatorios los factores *fruto* anidado a *sitio* y *sitio*. Las matrices de varianzas y covarianzas de los efectos aleatorios se condicionaron a cero a para todos los parámetros (β_1 , β_2 y β_3).
- ✓ **Pera_3:** consideró como factor fijo el *tamaño de fruto* y como factores aleatorios los factores *fruto anidado a sitio* y *sitio*. Las matrices de varianzas-covarianzas de los factores aleatorios no se condicionaron a cero en ninguno de sus parámetros (β_1 , β_2 y β_3).

Tabla 7: Comparación de modelos de crecimiento de pera que contemplan los diferentes efectos aleatorios en estudio

	<i>AIC</i>	<i>BIC</i>	<i>logLik</i>	<i>Test</i>	<i>L.Ratio</i>	<i>p-value</i>
Pera_1	47881,98	47957,04	-23930,99			
Pera_2	48104,64	48179,70	-24042,32			
Pera_3	47244,93	47365,02	-23606,46	2 vs 3	871,70	<0,0001

Entre los modelos ajustados se escogió aquel que arrojó menores valores para los criterios de información AIC y BIC (Tabla 7). El modelo seleccionado fue **Pera_3**, el cual contempló como efectos fijos el *tamaño de fruto*, y como efectos aleatorios el *fruto* anidado al *sitio* y el *sitio* y las matrices de varianzas-covarianzas de los factores aleatorios no se condicionaron a cero en ninguno de sus parámetros

El modelo obtenido puede expresarse de la siguiente manera:

$$Y_{klm} = \frac{1}{(\beta_{1,TAM_j} + b_{1,kl}) * 0.01 + ((\beta_{2,TAM_j} + b_{2,kl}) * 0.01) * ((\beta_{3,TAM_j} + \beta_{3,kl}) * 0.1)^{DDPF}} + \varepsilon_{klm}$$

con

$$b_{i,klm} = b_{i,SITIO_k} + b_{i,FRUTO_l}$$

dónde:

- ✓ β_{i,TAM_j} efecto del j-ésimo tamaño para el i-ésimo parámetro del efecto fijo (i: $\beta_1, \beta_2, \beta_3$; k: pequeño, mediano y grande);

- ✓ $b_{i,SITIO_k}$ efecto aleatorio del k-ésimo sitio para el i-ésimo parámetro del efecto aleatorio con $b_{i,SITIO_k} \sim N([0,0,0], D_{SITIO})$;
- ✓ $b_{i,FRUTO_{l(k)}}$ efecto del m-ésimo fruto para el i-ésimo parámetro del efecto aleatorio con $b_{i,FRUTO_{l(k)}} \sim N([0,0,0], D_{FRUTO})$;
- ✓ ϵ_{klm} error aleatorio con $\sim N(0, \sigma^2)$

con D_{SITIO} y D_{FRUTO} matrices de varianza-covarianza de los efectos aleatorios en sus correspondientes niveles.

Una vez definida la estructura de efectos fijos y aleatorios del modelo se prosiguió con la incorporación de las covariables ambientales e índices verdes a los diferentes parámetros del modelo para ver si las mismas permiten explicar la variabilidad asociada a los sitios y temporadas.

Los modelos ajustados con covariables ambientales fueron los siguientes:

- ✓ ***Pera_3_amb_1***: Se basó en el modelo ***Pera_3*** y se agregó la covariable ambiental TEMP_1 al parámetro β_3 .
- ✓ ***Pera_3_amb_2***: Se basó en el modelo ***Pera_3*** y se agregó la covariable ambiental TEMP_3 al parámetro β_3 .
- ✓ ***Pera_3_amb_3***: Se basó en el modelo ***Pera_3*** y se agregó la covariable ambiental TEMP_2 al parámetro β_2 y TEMP_3 al parámetro β_3

Los modelos ajustados con índices verdes fueron los siguientes

- ✓ ***Pera_3_verde_1***: Se basó en el modelo ***Pera_3*** y se agregó la covariable MSI_NOV al parámetro β_3 .
- ✓ ***Pera_3_verde_2***: Se basó en el modelo ***Pera_3*** y se agregó la covariable MSI_NOV al parámetro β_2 .
- ✓ ***Pera_3_verde_3***: Se basó en el modelo ***Pera_3*** y se agregó la covariable MSI_NOV al parámetro β_1 .
- ✓ ***Pera_3_verde_4***: Se basó en el modelo ***Pera_3*** y se agregó la covariable NDVI_NOV al parámetro β_3 .
- ✓ ***Pera_3_verde_5***: Se basó en el modelo ***Pera_3*** y se agregó la covariable NDVI_NOV al parámetro β_2
- ✓ ***Pera_3_verde_6***: Se basó en el modelo ***Pera_3*** y se agregó la covariable NDVI_NOV al parámetro β_1

- ✓ ***Pera_3_verde_7***: Se basó en el modelo ***Pera_3*** y se agregó la covariable MSI_DIC al parámetro β_3
- ✓ ***Pera_3_verde_8***: Se basó en el modelo ***Pera_3*** y se agregó la covariable MSI_DIC al parámetro β_2
- ✓ ***Pera_3_verde_9***: Se basó en el modelo ***Pera_3*** y se agregó la covariable MSI_DIC al parámetro β_1
- ✓ ***Pera_3_verde_10***: Se basó en el modelo ***Pera_3*** y se agregó la covariable NDVI_DIC al parámetro β_3
- ✓ ***Pera_3_verde_11***: Se basó en el modelo ***Pera_3*** y se agregó la covariable NDVI_DIC al parámetro β_2
- ✓ ***Pera_3_verde_12***: Se basó en el modelo ***Pera_3*** y se agregó la covariable NDVI_DIC al parámetro β_1

Tabla 8: Comparación de los modelos con diferentes covariables basados en Pera_3

	<i>AIC</i>	<i>BIC</i>	<i>logLik</i>
<i>Pera_3</i>	47243,14	47340,72	-23608,57
<i>Pera_3_amb_1</i>	46356,37	46506,49	-23158,18
<i>Pera_3_amb_2</i>	46352,86	46502,98	-23156,43
<i>Pera_3_amb_3</i>	46347,34	46504,97	-23152,67
<i>Pera_3_verde_1</i>	46350,66	46500,78	-23155,33
<i>Pera_3_verde_2</i>	46351,41	46501,53	-23155,70
<i>Pera_3_verde_3</i>	46354,36	46504,48	-23157,18
<i>Pera_3_verde_4</i>	46345,09	46495,21	-23152,54
<i>Pera_3_verde_5</i>	46355,44	46505,56	-23157,72
<i>Pera_3_verde_6</i>	46345,58	46495,70	-23152,79
<i>Pera_3_verde_7</i>	46353,63	46503,75	-23156,81
<i>Pera_3_verde_8</i>	46355,84	46505,96	-23157,92
<i>Pera_3_verde_9</i>	46355,63	46505,75	-23157,81
<i>Pera_3_verde_10</i>	46343,26	46493,38	-23151,63
<i>Pera_3_verde_11</i>	46355,66	46505,78	-23157,83
<i>Pera_3_verde_12</i>	46352,89	46503,01	-23156,44

De todos los modelos ajustados con covariables, aquellos que presentaron menores valores de criterios de información fueron *Pera_3_amb_3* y *Pera_3_verde_4*, *Pera_3_verde_6* y *Pera_3_verde_10* (Tabla 8). En el primer caso el modelo incluye las covariables TEMP_2 y TEMP_3 en los parámetros β_2 y β_3 , respectivamente; el segundo modelo incluye la covariable NDVI_NOV en el parámetro β_3 , el tercero NDVI_NOV en el parámetro β_1 y el último incluye NDVI_DIC al parámetro β_3 .

Finalmente, se ajustaron tres modelos diferentes que contienen al mismo tiempo aquellas covariables ambientales y de índices verdes que mostraron los mejores resultados. Los modelos ajustados fueron los siguientes:

- ✓ *Pera_3_MIX_1*: Se basó en el modelo *Pera_3* y se agregó la covariable TEMP_2 al parámetro β_2 , y la covariable NDVI_NOV al parámetro β_1 .
- ✓ *Pera_3_MIX_2*: Se basó en el modelo *Pera_3* y se agregó la covariable TEMP_2 al parámetro β_2 y las covariables TEMP_3 y MSI_DIC al parámetro β_3 .
- ✓ *Pera_3_MIX_3*: Se basó en el modelo *Pera_3* y se agregó la covariable TEMP_2 al parámetro β_2 , la covariable NDVI_NOV al parámetro β_1 y la covariable TEMP_3 al parámetro β_3 .

Tabla 9: Comparación de los modelos candidatos al momento y aquellos modelos que contemplan ambos tipos de covariables en peras.

	<i>AIC</i>	<i>BIC</i>	<i>logLik</i>
<i>Pera_3</i>	47243,14	47340,72	-23608,57
<i>Pera_3_amb_3</i>	46347,34	46504,97	-23152,67
<i>Pera_3_verde_4</i>	46345,09	46495,21	-23152,54
<i>Pera_3_verde_6</i>	46345,58	46495,70	-23152,79
<i>Pera_3_verde_10</i>	46343,26	46493,38	-23151,63
<i>Pera_3_MIX_1</i>	46340,08	46497,71	-23149,04
<i>Pera_3_MIX_2</i>	46339,90	46505,04	-23147,95
<i>Pera_3_MIX_3</i>	46338,38	46503,51	-23147,19

A partir de los resultados se observó que el modelo *Pera_3_MIX_3* arrojó los menores valores para los criterios de información (Tabla 9). A los fines descriptivos se presentan a continuación las estimaciones de dicho modelo (Tabla 12,

Tabla 13 y Tabla 12) y los gráficos de residuales se presentan en el anexo.

Tabla 10: Estimaciones de efectos aleatorios a nivel de sitio para el modelo Pera_3_MIX_3

	$\beta_1(\text{intercepto})$	$\beta_2(\text{intercepto})$	$\beta_3(\text{intercepto})$
<i>stdDev</i>	0,0816	0,4465	0,0089

Tabla 11: Estimaciones de efectos aleatorios a nivel de fruto anidado a sitio para el modelo Pera_3_MIX_3

	<i>stdDev</i>	<i>corr</i>		
$\beta_1(\text{intercepto})$	0,1327	$\beta_1(\text{in})$		
$\beta_2(\text{intercepto})$	0,9361	-0,268	$\beta_2(\text{in})$	
$\beta_3(\text{intercepto})$	0,0286	-0,681	-0,271	$\beta_3(\text{in})$
<i>Residual</i>	1,0619			

Tabla 12: Estimaciones de efectos fijos para el modelo Pera_3_MIX_3

	Value	Std.error	DF	t-value	p-value
β_1 (intercepto)	1,3942	0,1007	12589	13,8466	<0,0001
β_1 tamaño M	0,1142	0,0130	12589	8,7726	<0,0001
β_1 tamaño P	0,2636	0,0130	12589	20,1640	<0,0001
β_1 NDVI_NOV	-0,4702	0,1407	12589	-3,3425	<0,0001
β_2 (intercepto)	9,6436	0,6525	12589	14,7785	<0,0001
β_2 tamaño M	0,8268	0,0858	12589	9,6374	<0,0001
β_2 tamaño P	1,8428	0,0883	12589	20,8731	<0,0001
β_2 TEMP_2	-0,0354	0,0091	12589	-3,8659	<0,0001
β_3 (intercepto)	9,6826	0,0278	12589	348,5057	<0,0001
β_3 tamaño M	-0,0231	0,0029	12589	-7,8891	<0,0001
β_3 tamaño P	-0,0463	0,0030	12589	-15,6602	<0,0001
β_3 TEMP_3	0,0020	0,0009	12589	2,2767	0,0228

Tabla 13: Incidencia de los efectos aleatorios en los diferentes niveles considerados

NIVEL	β_1	β_2	β_3
SITIO	0,0816	0,4465	0,0089
FRUTO: SITIO	0,1127	0,9361	0,0286

Se observó que en todos los parámetros del modelo predominó el efecto aleatorio *fruto* (Tabla 13). En el caso del parámetro asociado al tamaño inicial de los frutos (β_2) el efecto aleatorio *sitio* tuvo mayor incidencia respecto de los otros dos parámetros.

Por último, se calcularon también algunos indicadores de bondad de ajuste y de capacidad predictiva de los modelos candidatos al momento (Tabla 14) para terminar de decidir cuál será el modelo final. Para ello, se sometieron los modelos a un proceso de calibración y validación en el cual se utilizaron datos no contenidos en los procedimientos de ajuste a fin de garantizar métricas predictivas más realistas (Yang y Huang, 2013; Tassile, 2020). En primer lugar, se realizó la calibración (que predice los nuevos efectos aleatorios (bi) a partir de datos no

contenidos en el ajuste) y luego, con los modelos calibrados se hicieron predicciones de futuras observaciones para cada uno de los frutos considerados. Por último, con los valores predichos se evaluó la performance predictiva de los modelos a través de las métricas de validación SM, SMR, SMRA, RMSE y Varianza SM. Para dichos cálculos se utilizaron dos funciones creadas por Tassile (2020).

Tabla 14: indicadores de bondad de ajuste y capacidad predictiva obtenidos para los modelo candidatos

	R^2	R^2_{cond}	SM	SMR	SMA	SMRA	RMSE	Varianza SM
<i>Pera_3</i>	0,99	0,99	-5,33	5,75	-8,29	8,93	6,43	3,58
<i>Pera_3_amb_3</i>	0,99	0,99	-1,64	2,35	-2,64	3,71	3,86	3,49
<i>Pera_3_verde_6</i>	0,99	0,99	0,45	1,99	0,55	3,09	3,45	3,42
<i>Pera_3_MIX_3</i>	0,99	0,99	-1,24	2,13	-2,03	3,36	3,69	3,47

Se observó que tanto el coeficiente de determinación como el coeficiente de determinación condicionado es el mismo en todos los modelos. Por su parte, el SM, el SMR y el SRMA fueron menores para los modelos *Pera_3_MIX_3* y *Pera_3_verde_6*, por lo cual dichos modelos fueron los modelos seleccionados para la sección de discusión de este capítulo y para realizar los pronósticos en el capítulo siguiente.

Cabe destacar que, si bien este modelo busca describir lo mejor posible el crecimiento de los frutos asociados a cada sitio bajo estudio, a los fines del pronóstico un modelo que incluya índices verdes, para ser práctico y poder llevarse a cabo, debería venir acompañado de un paquete tecnológico apropiado que permita obtener la información de manera automática o semi-automática. Es decir, por ejemplo, cuando el operario mide el calibre ecuatorial de un fruto al azar, se debería contar con algún sensor o dispositivo que registre la coordenada geográfica de donde se ubica dicho fruto para poder asociar su valor de cada índice y contar con ese dato de entrada que es requerido por el modelo.

3.3.3. Resultados de los ajustes de los modelos de crecimiento para manzanas

Tal como se realizó en peras, en primer lugar, se buscó definir la componente estructural del modelo, para lo cual se plantearon distintos modelos que contemplaron los diferentes efectos evaluados a lo largo de las temporadas en estudio, es decir, el factor *sitio* (como combinación entre temporada y localidad), *fruto* y *tamaño de fruto*. Este último se consideró como factor fijo en todos los modelos analizados, y lo que se fue modificando fue la parte aleatoria de los mismos.

Los modelos ajustados fueron los siguientes:

- ✓ **Manzana_1:** consideró como factor fijo el *tamaño de fruto* y como factor aleatorio el *fruto*.
- ✓ **Manzana_2:** consideró como factor fijo el *tamaño de fruto* y como factores aleatorios los factores *fruto anidado a sitio* y *sitio*. Las matrices de varianzas y covarianzas de los efectos aleatorios se condicionaron a cero a para todos los parámetros (β_1 , β_2 y β_3).
- ✓ **Manzana_3:** consideró como factor fijo el *tamaño de fruto* y como factores aleatorios *el fruto anidado a la planta* y *la planta anidada al sitio*. Las matrices de varianzas-covarianzas de los factores aleatorios no se condicionaron a cero en ninguno de sus parámetros (β_1 , β_2 y β_3).

Tabla 15: Comparación de modelos de crecimiento de manzana que contemplan los diferentes efectos aleatorios en estudio

	<i>AIC</i>	<i>BIC</i>	<i>logLik</i>	<i>Test</i>	<i>L ratio</i>	<i>p-value</i>
Manzana_1	62883,81	62936,85	-31434,91			
Manzana_2	62208,31	62284,08	-31094,16			
Manzana_3	61926,62	62047,84	-30947,31	2 vs 3	293.69	<0,001

Entre los modelos ajustados se escogió **Manzana_3**, ya que arrojó menores valores para los criterios de información (Tabla 15). Este modelo contempló como efectos fijos el *tamaño de fruto*, y como efectos aleatorios el *fruto anidado al sitio* y *el sitio*; y las matrices de varianzas-covarianzas de los factores aleatorios no se condicionaron a cero en ninguno de sus parámetros.

Una vez definida la estructura de efectos fijos y aleatorios del modelo se prosiguió con la incorporación de las covariables ambientales e índices verdes a los diferentes parámetros del modelo para ver si las mismas permiten explicar la variabilidad asociada a los sitios y temporadas.

Los modelos ajustados con covariables ambientales fueron los siguientes:

- ✓ ***Manzana_3_amb_1***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable ambiental TEMP_1 al parámetro β_3 .
- ✓ ***Manzana_3_amb_2***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable ambiental TEMP_2 al parámetro β_2 .
- ✓ ***Manzana_3_amb_3***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable ambiental TEMP_2 al parámetro β_2 y TEMP_3 al parámetro β_3 .

Los modelos ajustados con índices verdes fueron los siguientes:

- ✓ ***Manzana_3_verde_1***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable MSI_NOV al parámetro β_3 .
- ✓ ***Manzana_3_verde_2***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable MSI_NOV al parámetro β_2 .
- ✓ ***Manzana_3_verde_3***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable MSI_NOV al parámetro β_1 .
- ✓ ***Manzana_3_verde_4***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable NDVI_NOV al parámetro β_3 .
- ✓ ***Manzana_3_verde_5***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable NDVI_NOV al parámetro β_2 .
- ✓ ***Manzana_3_verde_6***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable NDVI_NOV al parámetro β_1 .
- ✓ ***Manzana_3_verde_7***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable MSI_DIC al parámetro β_3 .
- ✓ ***Manzana_3_verde_8***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable MSI_DIC al parámetro β_2 .
- ✓ ***Manzana_3_verde_9***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable MSI_DIC al parámetro β_1 .
- ✓ ***Manzana_3_verde_10***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable NDVI_DIC al parámetro β_3 .
- ✓ ***Manzana_3_verde_11***: Se basó en el modelo ***Manzana_3*** y se agregó la covariable NDVI_DIC al parámetro β_2 .

- ✓ **Manzana_3_verde_12**: Se basó en el modelo **Manzana_3** y se agregó la covariable NDVI_DIC al parámetro β_1

Tabla 16: Comparación de los modelos con diferentes covariables basados en **Manzana_3**

	<i>AIC</i>	<i>BIC</i>	<i>logLik</i>
Manzana_3	61926,62	62047,84	-30947,31
Manzana_3_amb_3	61219,02	61393,27	-30586,51
Manzana_3_amb_2	61209,75	61384,01	-30581,88
Manzana_3_amb_3	61199,55	61381,38	-30575,77
Manzana_3_verde_1	61222,13	61396,39	-30588,07
Manzana_3_verde_2	61221,31	61395,56	-30587,65
Manzana_3_verde_3	61216,30	61390,55	-30585,15
Manzana_3_verde_4	61222,13	61396,39	-30588,07
Manzana_3_verde_5	61221,12	61395,38	-30587,56
Manzana_3_verde_6	61217,11	61391,36	-30585,56
Manzana_3_verde_7	61218,80	61393,05	-30586,40
Manzana_3_verde_8	61218,47	61392,72	-30586,24
Manzana_3_verde_9	61221,63	61395,89	-30587,82
Manzana_3_verde_10	61221,07	61395,33	-30587,54
Manzana_3_verde_11	61220,16	61394,42	-30587,08
Manzana_3_verde_12	61221,11	61395,36	-30587,55

De todos los modelos ajustados con covariables, aquellos que presentaron menores valores de criterios de información fueron **Manzana_3_amb_3**, **Manzana_3_verde_3** y **Manzana_3_verde_6** (Tabla 16). En el primer caso el modelo incluye las covariables TEMP_2 y TEMP_3 en los parámetros β_2 y β_3 , respectivamente; el segundo incluye la covariable MSI_NOV al parámetro β_1 , y el último agrega la covariable NDVI_NOV al parámetro β_1 .

Finalmente, se ajustaron diferentes modelos que contienen al mismo tiempo aquellas covariables ambientales y de índices verdes que mostraron los mejores resultados.

Los modelos ajustados fueron los siguientes:

- ✓ **Manzana_3_MIX_1:** Se basó en el modelo **Manzana_3** y se agregó la covariable TEMP_3 al parámetro β_3 , la covariable TEMP_2 al parámetro β_2 y la covariable NDVI_NOV al parámetro β_1 .
- ✓ **Manzana_3_MIX_2:** Se basó en el modelo **Manzana_3** y se agregó la covariable NDVI_NOV al parámetro β_1 , la covariable TEMP_2 al parámetro β_2 y las covariables TEMP_3 al parámetro β_3 .
- ✓ **Manzana_3_MIX_3:** Se basó en el modelo **Manzana_3** y se agregó la covariable TEMP_2 al parámetro β_2 y las covariables MSI_DIC y TEMP_3 al parámetro β_3 .

Tabla 17: Comparación de los modelos candidatos al momento y aquellos modelos que contemplan ambos tipos de covariables en manzanas.

	AIC	BIC	logLik
Manzana_final	61926,62	62047,84	-30947,31
Manzana_3_amb_final	61199,55	61381,38	-30575,77
Manzana_3_verde_final	61216,30	61390,55	-30585,15
Manzana_6_verde_final	61217,11	61391,36	-30585,56
Manzana_1_MIX_final	61187,45	61376,85	-30568,72
Manzana_2_MIX_final	61187,50	61376,91	-30568,75
Manzana_3_MIX_final	61192,31	61381,71	-30571,15

De los modelos ajustados se observó que aquel que arrojó menores valores de criterios de información entre los modelos que contemplan solo dos efectos aleatorios fue el modelo **Manzana_3_MIX_1** (Tabla 17). A continuación, se presentan los coeficientes obtenidos para dicho modelo.

Del mismo modo que lo realizado en peras, a los fines descriptivos se presentan a continuación las estimaciones de dicho modelo (Tabla 20, Tabla 21, Tabla 21) y los gráficos de residuales se presentan en el anexo.

Tabla 18: Estimaciones de efectos aleatorios a nivel de sitio para el modelo *Manzana_3_MIX_1*

	β_1 (intercepto)	β_2 (intercepto)	β_3 (intercepto)
<i>stdDev</i>	0,0502	0,3173	0,0124

Tabla 19: Estimaciones de efectos aleatorios a nivel de fruto anidado a sitio para el modelo *Manzana_3_MIX_1*

	<i>stdDev</i>	<i>corr</i>
β_1 (intercepto)	0,1322	β_1 (in)
β_2 (intercepto)	0,8661	-0,502 β_2 (in)
β_3 (intercepto)	0,0235	-0,090 -0,586 β_3 (in)
<i>Residual</i>	1,6707	

Tabla 20: Estimaciones de efectos fijos para el modelo *Manzana_3_MIX_1*

	<i>Value</i>	<i>Std.error</i>	<i>DF</i>	<i>t-value</i>	<i>p-value</i>
β_1 (intercepto)	1,0794	0,0471	13663	22,9162	<0,0001
β_1 tamaño <i>M</i>	0,0989	0,0124	13663	7,9528	<0,0001
β_1 tamaño <i>P</i>	0,2291	0,0125	13663	18,2850	<0,0001
β_1 MSI_NOV	0,4552	0,0754	13663	6,0144	<0,0001
β_2 (intercepto)	8,5154	0,1739	13663	48,8202	<0,0001
β_2 tamaño <i>M</i>	0,7298	0,0905	13663	8,0670	<0,0001
β_2 tamaño <i>P</i>	1,7328	0,0951	13663	18,2113	<0,0001
β_2 TEMP_2	-0,0466	0,0014	13663	-33,1806	<0,0001
β_3 (intercepto)	9,6445	0,0083	13663	1167,011	<0,0001
β_3 tamaño <i>M</i>	-0,0156	0,0029	13663	-5,3893	<0,0001
β_3 tamaño <i>P</i>	-0,0312	0,0029	13663	-10,5891	<0,0001
β_3 TEMP_3	0,0020	0,0002	13663	9,6881	<0,0001

Además, se evaluó la incidencia que tuvo cada efecto aleatorio en cada uno de los parámetros del modelo.

Tabla 21: coeficientes de variación para los efectos aleatorios en los diferentes niveles considerados.

NIVEL	β_1	β_2	β_3
SITIO	0,0502	0,3172	0,0124
FRUTO:PLANTA:SITIO	0,1322	0,8661	0,0235

Se observó que en todos los parámetros del modelo predominó el efecto *fruto*, mientras que *sitio* tuvo una menor incidencia (Tabla 21).

Por último, tal como se realizó en peras, se calcularon también los indicadores de bondad de ajuste y de capacidad predictiva de los modelos candidatos al momento.

Tabla 22: indicadores obtenidos en los modelos candidatos para el análisis de la capacidad predictiva

	R^2	R^2_{cond}	SM	SMR	SMA	SMRA	Varianza SM
Manzana_3	0,99	0,99	-0,03	-0,76	1,06	3,34	1,56
Manzana_3_amb_3	0,99	0,99	-0,03	-0,74	1,07	3,35	1,57
Manzana_3_verde_3	0,99	0,99	-0,03	-0,75	1,07	3,35	1,57
Manzana_3_verde_6	0,99	0,99	-0,03	-0,75	1,07	3,35	1,57
Manzana_3_mix_1	0,99	0,99	-0,03	-0,75	1,07	3,35	1,57

Se observó que tanto el coeficiente de determinación como el coeficiente de determinación condicionado es el mismo para todos los modelos. Por su parte, el SM, el SMR, el SMRA y el SD fueron prácticamente iguales también para todos modelos, siendo el modelo *Manzana_3_amb_3* el que presentó el menor SMR. Por lo cual cualquiera de los modelos podría ser seleccionado para el pronóstico (Tabla 22).

3.3.4. Discusión acerca de los modelos de crecimiento obtenidos

Los modelos obtenidos expresan el tamaño del fruto como diámetro ecuatorial en función de los días después de plena floración, a diferencia de otros autores que consideran la necesidad de medir el peso o el volumen de los frutos (Lakso et al., 1995). Debido a que medir el peso o volumen implica técnicas destructivas que impiden medir los mismos frutos a lo largo del tiempo, en este trabajo se utilizó la medición de diámetros y luego la relación peso-diámetro (De Silva et al., 1997).

En esta tesis, el crecimiento de manzanas cv ‘Red Chief’ y peras cv ‘Williams’ en el Alto Valle de Rio Negro fue descrito por la curva logística en su tercera parametrización, lo cual concuerda con lo presentado para peras cv. ‘William’s’ y ‘Packham’s Triumph’ para la misma región (Bramardi et al., 1997; Giménez, 2012; Tassile, 2020) y con lo presentado por Atay et al., (2010) en manzanas en Turquía. Otros autores, a diferencia del modelo logístico, han ajustado para manzanas curvas dobles sigmoideas (Magein, 1989), modelos expolineales (Lakso et al., 1995), funciones exponenciales negativas (Zadravec, et al., 2014), o funciones cuadráticas (Atay y Pirlak, 2009).

Los resultados obtenidos en los modelos para ambas especies ponen en evidencia que existe un efecto del factor tamaño de fruto (pequeño, mediano y grande) sobre los calibres ecuatoriales obtenidos y la necesidad de considerar efectos aleatorios a nivel de fruto y sitio para todos los parámetros considerados. El uso de NLMM multiniveles presentó la ventaja de poder cuantificar distintas fuentes de variabilidad aleatorias, además de la variabilidad del error. Esto concuerda con los resultados obtenidos por otros autores como Hall y Bailey (2001); Avanza (2010); Jiang y Li (2010); Fu, et al. (2013); Tassile (2020). Además, la introducción de coeficientes de efectos aleatorios a nivel de fruto en los modelos permitió modelar la covarianza entre observaciones provenientes de un mismo fruto brindando flexibilidad a dichos modelos para contemplar la variabilidad existente entre individuos. Resultados similares fueron presentados por Tassile et al., (2005); Avanza (2010); Giménez y Tassile (2015); Tassile (2020) y Bóveda (2022). Tanto en manzanas como en peras, el efecto aleatorio de fruto fue el que tuvo mayor incidencia en todos los parámetros, lo cual concuerda con los hallados por Tassile (2020) para peras cv “William’s”. En manzanas, particularmente, el efecto sitio tuvo muy poca incidencia y esto podría explicarse porque la incorporación de covariables ambientales e índices verdes podrían estar explicando el efecto asociado a cada temporada y lugar.

En ambas especies, la inclusión de covariables del tipo ambientales e índices verdes mejoró notablemente la performance de los modelos obtenidos. Esto concuerda con el trabajo realizado por Tassile, et al., (2005) que demostraron el efecto diferencial que generan los índices de termo acumulación según el momento de ciclo en el cual se encuentran los frutos. Por un lado, altas temperaturas producidas en etapas tempranas del crecimiento de los frutos dan lugar a mayores calibres, mientras que, por el otro, si éstas ocurren en etapas más tardías (diciembre) dan lugar a frutos con menores calibres comerciales. Rodríguez (2011), también concluyó que la temperatura tiene un efecto regulador tanto en el tamaño final de los frutos como en la tasa de crecimiento de pera cv 'William's. Estudios posteriores realizados por Tassile, (2020), también concuerdan con lo obtenido en la presente tesis. Dicho autor propone NLMM para describir el crecimiento de peras cv 'Williams' con la incorporación de diversos índices de termo acumulación en distintos momentos críticos del crecimiento de los frutos, mostrando resultados favorables para dicha incorporación. También concluye en su investigación que las performances predictivas de los modelos que no incluyen covariables ambientales, dependen en gran medida del momento en el cual se realiza la predicción, mientras que los modelos que incorporan las covariables ambientales muestran buena capacidad predictiva en todas las situaciones estudiadas.

La acción que ejercen las temperaturas sobre los procesos fisiológicos involucrados en la expansión celular a lo largo del mes de diciembre y, por ello, en el tamaño final de los frutos, pueden explicar la adecuada performance de los índices ambientales o de termo acumulación en los modelos. Esto también concuerda con los resultados obtenidos por Tassile et al., (2005); Giménez (2012) y Tassile (2020). En frutos de Lichi los elementos meteorológicos también influenciaron las tasas de crecimiento, particularmente la variable déficit de agua afectó el desarrollo de los frutos; y las variables evapotranspiración potencial y evapotranspiración actual, tuvieron influencia en la materia fresca y seca, en el volumen y tamaño de fruto y en el número de frutos (de Oliveira Aparecido et al., 2016). En arándanos la incorporación de variables ambientales mejoró las estimaciones de los modelos obtenidos, presentando las unidades de frío significancia en los modelos múltiples ajustados, lo que sugirió que las condiciones del invierno pueden afectar el desarrollo de las yemas florales (Salvo et al., 2012). En naranjas, las variables asociadas a temperaturas del aire máximas y mínimas y a lluvias, fueron útiles para predecir el número de frutos por árbol a partir de modelos de regresión lineal múltiples (Pasqua Paulino et al., 2007). En manzanas por su parte, los grados día se asocian con las variables respuesta diámetro, peso y calidad de fruto, la humedad relativa se asocia a

la calidad del fruto y la evapotranspiración potencial se asocia al diámetro de fruto (Kaack y Pedersen, 2010).

El NDVI también tuvo un buen comportamiento en los modelos cuando se lo asoció al parámetro β_3 , es decir a la tasa de crecimiento. Otros autores como Martínez-Casasnovas y Bordes, (2005), han encontrado relación entre los rendimientos de vid y el vigor de las plantas obtenido a partir del índice NDVI y otras variables asociadas a la fertilidad del viñedo; y Gemtos et al., (2013) que concluyen que altos valores de NDVI se corresponden con altos rendimientos en de manzana y vid.

Los R^2 obtenidos en los modelos finales propuestos fueron muy cercanos a uno por lo cual se puede decir que la variable dependiente fue explicada por los modelos ajustados. La capacidad predictiva de los modelos fue razonable mostrando valores de SM cercanos a un mm en peras y bastante menores en manzanas y de SMR cercanos al dos por ciento en peras y al uno por ciento en manzanas cuando se validaron con datos no contemplados en la modelación. Tassile (2020), obtuvo valores similares de SM en su trabajo en el cual ajustó NLMM en peras para la misma región en estudio. Por su parte, los valores de RSME obtenidos en esta tesis fueron de 3,69 para peras y 3,37 para manzanas. Bóbeda, (2022) ajustó modelos no lineales para describir el crecimiento de naranjos y obtuvo valores de 1,00 a 1,24 de RSME, los cuales son menores a los hallados en la presente tesis.

3.4. CONCLUSIONES

El uso de MMNL multiniveles presentó la ventaja de poder discernir y cuantificar fuentes de variabilidad a diferentes niveles. El efecto aleatorio que mayor influencia tuvo en todos los parámetros de los modelos fue *fruto*.

La incorporación de variables ambientales e índices verdes mejoró el comportamiento de los modelos obtenidos, permitiendo contemplar la variabilidad asociada a cada temporada y sitio analizado. Si bien los modelos lograron describir bien el crecimiento de los frutos, a los fines del pronóstico no resultan prácticos si no están acompañados de un paquete tecnológico apropiado, por lo cual para hacer pronósticos se utilizarán los modelos sin los índices verdes.

La capacidad predictiva de los modelos cuando se calculó con datos no contemplados en el proceso modelado fue razonable, con valores de SM cercanos a 1 mm en peras y menores para manzanas; y de SMR cercanos al 2% y 1% para peras y manzanas respectivamente.

4.CAPITULO IV: PRONÓSTICOS DE PRODUCCIÓN

4.1. INTRODUCCION

4.1.1. Pronósticos de producción

La estimación de los rendimientos continúa siendo un gran desafío en la mayoría de los montes frutales, ya que dicha información, como se mencionó previamente, es muy valiosa tanto para los productores como para las empresas que brindan servicios de empaque y poscosecha para poder administrar los insumos, el equipamiento y la mano de obra necesaria de manera más efectiva (Linker, 2016).

Son diversas las metodologías llevadas a cabo en la actualidad para realizar estimaciones de producción en frutales. Estas pueden clasificarse según los métodos utilizados en subjetivas y objetivas. Las primeras son mucho más simples, y se basan, generalmente, en la opinión de personas expertas en la temática y de los productores; como así también, en los volúmenes comercializados de fruta para una zona en particular. Estos métodos dependen mucho de la experiencia de quienes lo realizan y tienen como ventaja que poseen un costo muy bajo (Avanza, 2010). Los métodos objetivos, por su parte, son aquellos que utilizan modelos estadísticos o matemáticos para predecir el comportamiento de la producción, y pueden estar basados tanto en información histórica de distintos aspectos productivos, como en mediciones a campo. A su vez, los métodos objetivos pueden clasificarse según la forma de estimar el tamaño de los frutos en aquellos que utilizan curvas de crecimiento o aquellos basados en modelos estocásticos. Estos últimos están basados en el supuesto de que existe una relación entre el peso de los frutos y la carga de los árboles. En la Unión Europea, el Centro Servizi Ortofrutticoli, utiliza estos modelos para realizar pronósticos de producción en frutales de pepita. En Argentina, en cambio, no se ha logrado establecer esta relación entre el peso y la carga, por lo que ha resultado más conveniente utilizar curvas de crecimiento (Bramardi et al., 2005). Durante muchos años 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, ha realizado pronósticos de producción a nivel regional de las principales variedades cultivadas mediante el relevamiento de información y, a partir del año 2005, mediante el uso de curvas de crecimiento. En esta tesis, se trabajará siguiendo la metodología utilizada por dicho pronóstico regional.

Otros antecedentes de pronósticos de producción basados en mediciones objetivas son los modelos ajustados por Miranda Jiménez & Royo Díaz, (2004) en España que estiman la producción en manzano cv 'Gala' y 'Golden Delicious' a partir de diferentes variables medidas en los árboles como el diámetro de tronco a la altura del pecho, el espacio designado a cada árbol, la densidad de flores, el número de ramilletes frutales, número de frutos por ramillete, entre otras. Mediante regresión múltiple ajustaron modelos para el número de frutos por centímetro cuadrado y para el rendimiento medio por ramillete frutal. Luego con estas variables obtuvieron las estimaciones de los rendimientos que se alejaron aproximadamente el 15% de los valores reales. En peras cv 'Blanquilla' y 'Conference' realizaron un trabajo similar y el error observado entre los rendimientos esperados y observados fue de aproximadamente 23% lo cual consideraron aceptable para el propósito perseguido (Miranda Jiménez & Royo Díaz, 2003). También varios autores han trabajado en la estimación de los rendimientos de montes de manzana mediante uso de imágenes como Aggelopoulou et al., (2011) que estiman cosecha a partir de imágenes de árboles tomadas en plena floración en Grecia. El algoritmo desarrollado predice la cosecha mediante el análisis de la textura de las imágenes, y combinado con datos históricos, permite realizar una predicción temprana de los rendimientos del año. Otros autores lo hacen a partir de imágenes tomadas a los árboles con frutos en distintos momentos de su ciclo de crecimiento, como Linker, (2016) en Israel, que utiliza regresión lineal múltiple y obtiene desviaciones estándar de hasta el 27% de los rendimientos estimados promedio por árbol; Zhou et al., (2012) que logra coeficientes de determinación entre 0,58 y 0,71 para las estimaciones de cosecha en Alemania, mejorando los resultados en función del momento en que se toman las imágenes; Stajniko et al., (2009) logran coeficientes de determinación de 0,96, existiendo una subestimación de la cosecha en casi todos los momentos, a excepción de los primeros en Eslovenia; Crotmir et al., (2012) con algoritmos de redes neuronales logra correlaciones de 0,51 a 0,83 entre la cosecha estimada y real en función del momento y cultivar.

En esta tesis, además de realizar las estimaciones de producción utilizando la metodología del pronóstico regional, se evaluará la incorporación de información proveniente de imágenes, tal como los últimos ejemplos citados en el párrafo anterior.

4.1.2. Comparación de modelos de predicción.

Una vez obtenidos los modelos de predicción basados en el uso de imágenes y en la obtención de datos de forma tradicional, es fundamental poder comparar estas metodologías con el fin de conocer la precisión de las mismas (respecto de la cosecha real) y el costo asociado a cada una de ellas. Esta información resulta de interés para poder optar por aquella metodología que resulte más conveniente a cada productor o empresa en particular.

Para la comparación se puede realizar un presupuesto parcial para cada una de las metodologías, que es básicamente un formato para la planificación y toma de decisiones utilizado para comparar costos y beneficios asociados a distintas alternativas a las que se enfrenta una empresa. Un presupuesto de este tipo se centra solamente en los cambios que ocurrirán en los ingresos y gastos que se derivan de aplicar una alternativa específica (Harper et al., 2014). En otras palabras, se llaman parciales, porque solamente contemplan los costos asociados con la decisión de usar o no un tratamiento o metodología. Estos son los costos que permiten diferenciar una alternativa de la otra (costos que varían), y el resto de los costos no se verán afectados por la decisión de usar o no una alternativa en particular, y permanecerán constantes (costos fijos). Generalmente se desarrollaron para formular recomendaciones a partir de datos agronómicos, por lo tanto, el proceso de aplicación de este tipo de presupuesto debe generar una conclusión de recomendación o decisión de manejo (Hernández, 2001).

Este tipo de presupuestos se ha utilizado por ejemplo para comparar el efecto de diferentes productos químicos para controlar *Polyphagotarsonemus latus* en plantas de *Jatropha curcas* y determinar cuál es el mejor tratamiento considerando la efectividad y el costo de cada uno (Avalos-Cerda y Villalobos-Mongue, 2018); para evaluar diferentes marcas comerciales de prostaglandina F2 α (nacionales e importadas) y comparar los porcentajes de presentación de celo y preñez posterior a los tratamientos y los costos de cada uno para poder realizar recomendaciones (Ferreya, 2018), entre otros. Por lo cual realizar un análisis de presupuesto parcial para la comparación de las metodologías de predicción, se vuelve interesante para comparar ventajas, desventajas, precisión y costos de las mismas, con el objetivo de dar una recomendación de selección de alguna de ellas.

4.1.3. Estimación de la distribución de tamaños

El conocimiento anticipado de los calibres de los frutos a obtener en cosecha constituye una herramienta de gran utilidad. Por un lado, por que ayuda a la programación de tareas a cosecha, teniendo en cuenta la necesidad de mano de obra, el espacio de almacenamiento y el aprovisionamiento de materiales de embalaje del tamaño apropiado. Por otro lado, para la negociación, ya que define, en gran medida, el mercado de destino de comercialización de la fruta y también su precio; ya que cada mercado consumidor tiene su preferencia por frutos de distinto tamaño (Bóbeda, 2022).

Mayorano et al. (2006) presentaron un modelo para la estimación del crecimiento de frutas en manzanas cv “Granny Smith” basado en cadenas de Markov, que fue validado con éxito. El modelo se realizó como soporte a las decisiones comerciales y logísticas, para realizar pronósticos de las distribuciones de tamaños. Por su parte, Marine et al. (2019), midieron diámetros de manzanas a 60 DDPF y los utilizaron 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 estos pesos se compararon con las distribuciones de pesos estimadas tempranamente, encontrándose diferencias entre ambas distribuciones para todos los árboles estudiados. En su trabajo, miden entre el 7 y 30% de los frutos presentes en 93 árboles, y concluyen que una estimación más precisa de la distribución del tamaño de la fruta requeriría medir aproximadamente el 50% de los frutos de cada árbol.

Giménez (2020) realizó predicciones de tamaños comerciales a cosecha en peras utilizando la técnica de máquinas de vectores soporte. Construyó una matriz de confusión contrastando los valores predichos por dicho método y los observados a campo y la exactitud global obtenida (expresada como frutos clasificados correctamente) fue de 0,60. También redujo la cantidad de categorías de tamaños de frutos y así logró mejorar la exactitud, alcanzando un valor de 0,77 con un intervalo de confianza bootstrap de 0,68 a 0,86.

En la región del Alto Valle de Río Negro y Neuquén, la Facultad de Ciencias Agrarias ha trabajado en la construcción de tablas de crecimiento de frutos para las variedades de peras y manzanas más representativas de la región, las cuales contienen estimaciones de los tamaños comerciales que alcanzarán los frutos al momento de cosecha, basados en la medición de diámetros y en los DDPF. Estas tablas se basan en los modelos de crecimiento ajustados para cada variedad, utilizan la relación peso-diámetro de los frutos y los tamaños comerciales en base a las cajas que se emplean para la comercialización de la fruta. Son tablas de doble entrada (Figura 20) donde la primera columna hace referencia al tamaño comercial al

momento de la cosecha y las demás columnas indican el diámetro (en mm) alcanzado por los frutos en diferentes momentos expresados en DDPF. Estas tablas describen el comportamiento medio de los frutos para toda la región y constituyen una herramienta práctica, que permite pronosticar para cada diámetro observado en un determinado DDPF, el tamaño que dicho fruto alcanzará a cosecha. Regionalmente se las conoce también como tablas de raleo y son muy utilizadas para mejorar esta práctica en montes frutales.

		Días Posteriores a Plena Floración (DPPF)																			
		Raleo Manual					Cosecha														
		40	45	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120	125	130	135
Tamaño Comercial Envase 18 kg	90	32,0	35,1	38,3	41,6	44,9	48,2	51,4	54,6	57,6	60,5	63,1	65,7	68,0	70,1	72,0	73,8	75,3	76,7	78,0	79,1
	100	31,4	34,2	37,0	39,9	42,8	45,8	48,7	51,6	54,4	57,2	59,8	62,4	64,8	67,1	69,2	71,1	73,0	74,6	76,2	77,6
	110	30,5	33,5	36,5	39,7	42,8	46,0	49,1	52,1	54,9	57,7	60,3	62,7	65,0	67,0	68,9	70,6	72,1	73,5	74,7	75,8
	120	30,0	32,8	35,6	38,5	41,5	44,4	47,3	50,2	52,9	55,6	58,2	60,6	62,9	65,0	66,9	68,7	70,4	71,9	73,2	74,4
	135	27,9	30,7	33,5	36,4	39,3	42,2	45,2	48,0	50,8	53,5	56,1	58,5	60,7	62,8	64,7	66,5	68,1	69,6	70,9	72,0
	150	27,6	30,3	33,1	35,9	38,8	41,7	44,6	47,4	50,1	52,7	55,2	57,5	59,7	61,7	63,6	65,2	66,8	68,2	69,4	70,5

Figura 20: Tabla de crecimiento de frutos, tomada de Reeb et al., 2018.

Por ejemplo, si el raleo manual del monte frutal se realiza 60 DDPF y se pretende dejar solamente los frutos que alcanzarán al menos un diámetro de 70,6 mm (tamaño comercial 110) al momento de cosecha (115 DPPF), debemos eliminar todos los frutos con diámetro inferior a 42,8 mm.

En esta tesis se realizarán estimaciones de distribuciones de tamaños para la temporada 2020-2021, clasificando la fruta en tres clases (i: descarte y fruta chica; ii: fruta mediana; iii: fruta grande) las cuales se describirán más adelante en la metodología.

4.2. MATERIALES Y METODOS

4.2.1. Pronostico de producción

La construcción del modelo general de predicción, como se ha mencionado previamente, se basó en mediciones objetivas a campo que incluyen tres elementos:

- i) la predicción de la carga frutal: se obtuvo por métodos de conteo manual y de conteo a partir de imágenes como se explicó en el capítulo dos de esta tesis;
- ii) la estimación del tamaño que alcanzarán los frutos al momento de cosecha: lo cual se obtuvo mediante mediciones de calibres de frutos elegidos al azar realizadas al momento del pronóstico, luego con dichas mediciones se realizaron las estimaciones de calibres a cosecha a partir de las curvas de crecimiento ajustadas en el capítulo tres de esta tesis; y, por último, los calibres se llevaron a peso de fruto en gramos a partir de la relación peso diámetro de frutos (Bramardi et al., 1997; Giménez, 2020) para luego promediarlos.
- iii) el número de árboles presentes en monte frutal: esta información es obtenida por los registros del productor.

En la Figura 21 se puede observar el modelo general de rendimiento basado en el utilizado por la Facultad de Ciencias Agrarias, Universidad Nacional del Comahue para el pronóstico de producción de peras y manzanas para las provincias de Río Negro y Neuquén, pero con la diferencia de que en la presente tesis las estimaciones serán a nivel de parcela y, además, se incorpora el uso de análisis de imágenes como alternativa a los métodos tradicionales de conteo de frutos.

El pronóstico se realizó para cada uno de los tres montes de manzana y de pera para la temporada 2020 y 2021. En las temporadas anteriores no se realizó la medición de diámetros de frutos al azar y por lo tanto no se pudo calcular el peso medio de fruto esperado.

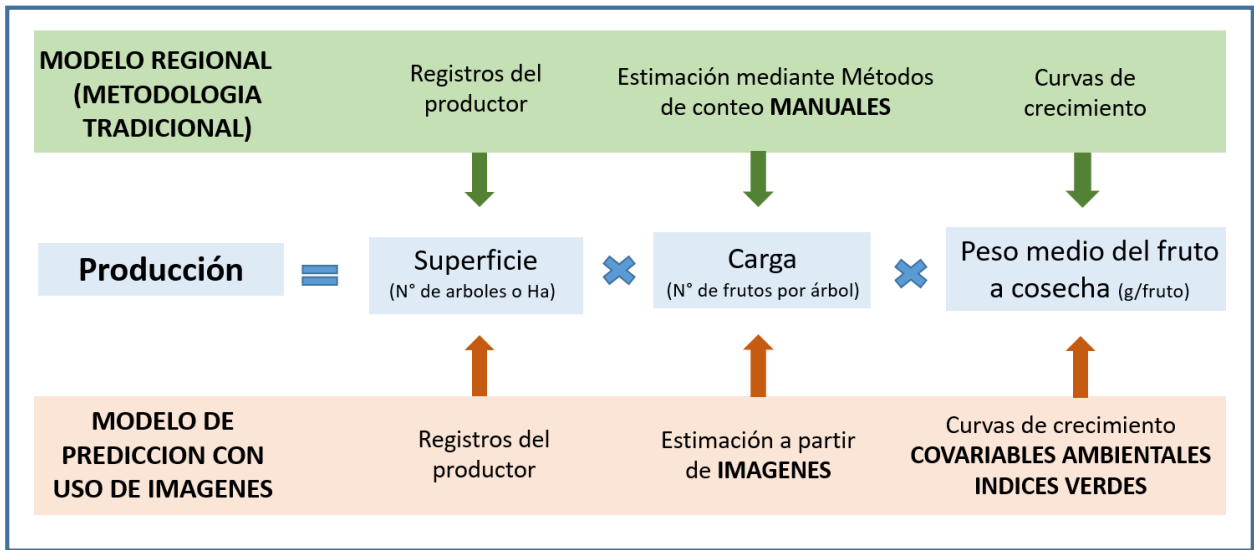


Figura 21: esquema del modelo general del pronóstico de producción regional

4.2.1.1. Metodología tradicional

En este caso se siguió la metodología utilizada en pronóstico de cosecha de Río Negro y Neuquén. Para estimar el número de frutos presentes en los árboles se realizó el recuento total de frutos por árbol en un número de árboles representativo del monte frutal, aproximadamente 45 días previos al momento del inicio de cosecha en pera y 70 días en manzana y los valores de carga se corrigen según un coeficiente de fruto oculto.

El número de plantas presentes en el monte frutal y/o la superficie implantada fue brindado por los productores y el INTA. Con dichos datos, la estimación de producción de cada especie se calculó a partir de la Ecuación 10.

Ecuación 10: modelo utilizado en el pronóstico regional de las provincias de Río Negro y Neuquén.

$$Pa = \frac{NF * PMF * NA}{k}$$

Donde:

NF = número medio de frutos por árbol

PMF = peso estimado del fruto medio a cosecha

NA = número de árboles

k = factor menor a uno, que depende de cada especie y variedad, y tiene por finalidad corregir el error que se produce durante el conteo de cargas por fruto oculto. Se calculó como la relación entre el número de frutos contados y el número de frutos cosechados.

Para calcular el peso medio de fruto a cosecha se estimaron los diámetros a cosecha de los frutos medidos al azar a principios de diciembre. Luego, estos diámetros estimados se pasaron a peso, para ello se utilizaron los modelos potenciales ajustados por Bramardi (1995) para describir la relación peso-diámetro de frutos de pera cv “William’s” que alcanzó coeficientes de determinación (R^2) del 99%, y los modelos potenciales ajustados por Giménez (2020) para manzanas cv “Red Delicious”. Por último, se calculó el peso medio de fruto a cosecha.

4.2.1.2. Mediante incorporación de imágenes

El único elemento que se modificó respecto del pronóstico anterior es el conteo del número de frutos, que se realizó a partir del uso de imágenes como se describió en el capítulo dos de la presente tesis. Luego, con toda la información se realizaron los cálculos de estimación de rendimientos de la misma manera que para la metodología tradicional.

4.2.2. Cálculo de presupuestos parciales y comparación

En primer lugar, se realizó la suma de las horas de trabajo necesarias para llevar a cabo las dos metodologías de pronóstico: tradicional y mediante el uso de imágenes considerando un solo monte frutal y tomando como muestra 60 plantas (como lo realizado en INTA para esta tesis).

Luego a partir del valor del jornal determinado para el año 2023 por la comisión nacional de trabajo (resolución 126, del 12 de agosto de 2022) se calculó el valor de la hora de trabajo dividiéndolo por ocho. Luego, se multiplicó el valor de la hora por las horas requeridas en

cada tarea para calcular el costo asociado a las mismas. Por último, se comparó solo la parte diferencial del costo asociado a cada metodología y la precisión en la estimación de cada una.

4.2.3. Estimación de la distribución de tamaños comerciales

Para la estimación de la distribución de tamaños comerciales se utilizaron los modelos ajustados en el capítulo tres de esta tesis. Dichos modelos se ajustaron en base al diámetro de frutos, y para poder realizar pronósticos de cosecha expresados en kilos o toneladas por hectárea, deben transformarse a pesos. Para ello se utilizaron los modelos potenciales ajustados por Bramardi (1995) y por Giménez (2020).

Luego, con dichos pesos y diámetros se realizó la estimación de la distribución de tamaños esperados, clasificándolos en tres clases (Tabla 24 y Tabla 24), para cada sitio y especie. A su vez se realizaron pruebas Chi cuadrado para comparar la distribución esperada con la observada y se realizaron histogramas de las distribuciones de tamaños observados y esperados.

Tabla 23: Clases creadas para la estimación de la distribución de tamaños a cosecha en peras

<i>Peras</i>		
	<i>Diámetro</i>	<i>Clases Tamañadora</i>
<i>Fruta chica y descarte</i>	<62 mm	Descarte y clases 5 y 6
<i>Fruta mediana</i>	62 a 66 mm	Clases 3 y 4
<i>Fruta grande</i>	67 mm o más	Clases 1 y 2

Tabla 24: Clases creadas para la estimación de la distribución de tamaños a cosecha en manzanas

<i>Manzanas</i>		
	<i>Diámetro</i>	<i>Clases Tamañadora</i>
<i>Descarte</i>	<70 mm	Descarte y clases 5 y 6
<i>Fruta chica</i>	70 a 76 mm	Clases 3 y 4
<i>Fruta grande</i>	77 o más mm	Clases 1 y 2

4.3. RESULTADOS

4.3.1. Pronóstico de producción

En la Tabla 25 a Tabla 28 se muestran los rendimientos obtenidos en cada monte frutal, los resultados obtenidos a partir del modelo de estimación de la producción con información obtenida con métodos tradicionales y mediante la incorporación de imágenes con sus respectivos errores.

Tabla 25: Rendimientos obtenidos y estimados mediante pronóstico tradicional en tn/ha para manzanas

Manzanas	Rendimiento real (tn/ha)	Pronostico tradicional	Error (tn/ha)	Error relativo
INTA	17,20	18,02	0,82	4,77%
ROMAN SRL	26,66	27,06	0,40	1,50%
PATAGONIAN SRL	29,16	30,63	1,47	5,04%

Tabla 26: Rendimientos obtenidos y estimados mediante pronóstico por imágenes en tn/ha para manzanas

Manzanas	Rendimiento real (tn/ha)	Pronostico imágenes	Error (tn/ha)	Error relativo
INTA	17,20	16,63	-0,57	3,31%
ROMAN SRL	26,66	29,50	2,84	10,65%
PATAGONIAN SRL	29,16	27,40	-1,76	6,03%

Tabla 27: Rendimientos obtenidos y estimados mediante pronóstico tradicional en tn/ha para peras

Peras	Rendimiento real (tn/ha)	Pronostico tradicional	Error (tn/ha)	Error relativo
INTA	28,50	25,75	-2,75	9,64%
ROMAN SRL	25,35	23,73	-1,62	6,39%
PATAGONIAN SRL	20,50	19,18	-1,32	6,44%

Tabla 28: Rendimientos obtenidos y estimados mediante pronóstico por imágenes en tn/ha para peras

Peras	Rendimiento real (tn/ha)	Pronostico imágenes	Error (tn/ha)	Error relativo
INTA	28,50	25,77	-2,73	9,58%
ROMAN SRL	25,35	31,56	6,21	24,49%
PATAGONIAN SRL	20,50	21,33	0,83	4,04%

En el caso de las manzanas el pronóstico realizado de forma tradicional presentó errores menores a 1,5 tn/ha, y errores relativos de hasta 5%, mientras que el pronóstico por imágenes arrojó errores un poco más altos de hasta 2,84 tn/ha y errores relativos desde tres a diez por ciento. En el caso de las peras, el pronóstico realizado de forma tradicional presentó errores de hasta 2,75 tn/ha, y errores relativos menores al diez por ciento. Por su parte, el pronóstico por imágenes arrojó errores más diversos, desde 0,83 toneladas en uno de los montes frutales (que representa un error relativo del 4%) hasta 6,21 tn/ha en otro de los huertos (que representa un error relativo de casi el 25%).

Entre el pronóstico realizado tradicionalmente y a partir de imágenes se observó que el primero arrojó errores de estimación menores que el segundo. Sin embargo, de todas las estimaciones realizadas a partir de imágenes solo un monte de peras presentó un error más alto, y el resto de los montes frutales mostraron errores relativos de hasta 10%. Esto puede explicarse porque las imágenes utilizadas para el conteo de número de frutos por árbol, en la mayor parte de las ocasiones muestran más de un árbol, lo cual puede estar generando un número de frutos diferente al real. Se debe trabajar a futuro en este aspecto para determinar

los límites del árbol al que se pretende contar sus frutos para evitar este error. Otro aspecto a considerar en este pronóstico también, es que la imagen solo capturó un plano del árbol y los montes frutales utilizados en el ensayo, si bien son espalderas, no forman una pared o muro totalmente horizontal, sino que muchas ramas cruzan en otras direcciones, por lo cual mucha fruta puede no salir en la imagen. En este caso se puede seguir avanzando en el trabajo planteando más de una imagen por árbol en diferentes posiciones o ángulos respecto del árbol como lo hace Itakura et al., (2021). El último de los aspectos a considerar es que el pronóstico se realizó tempranamente, por lo cual en dicho momento los frutos son pequeños, de un color similar a las hojas por lo cual la detección presentó algunas dificultades, tal como se explicaron en el capítulo dos de esta tesis.

Por su parte, los resultados de los pronósticos realizados mediante la metodología tradicional resultaron menores o iguales a los obtenidos en las estimaciones del pronóstico regional de cosecha de las provincias de Río Negro y Neuquén. Un aspecto a tener en cuenta que hace diferente estas estimaciones es que este último realiza la estimación para toda la región, mientras en esta tesis se realizó a nivel de parcela, por lo cual se planteó un muestreo mayor en cuanto al número de plantas por parcela e incluyó variables ambientales e índices verdes propios de cada una de ellas. Esto requiere contar con más información al momento de realizar las predicciones. Las estimaciones del pronóstico de producción en miles de toneladas (Tabla 29, construida con datos tomados de Giménez, (2020)) muestran diferencias desde tres hasta 80 mil tn/ha en estimación de peras y desde 41 a 291 mil tn/ha en la de manzanas según la temporada, lo cual en términos de errores relativos representa hasta el 16 y 36% en peras y manzanas, respectivamente. Giménez (2020), explicó que el mayor error observado en la estimación de manzanas puede deberse a una sobreestimación del método, o quizás existe un volumen de producción que, o no se retira del monte frutal por daños, o no se comercializa por canales de comercialización registrados por la secretaría de fruticultura de la provincia. En general, resulta de interés tanto para los productores como para las distintas empresas, que la estimación regional sea expresada en términos de variación anual de la producción total por lo cual se calculan las variaciones interanuales como diferencia de una temporada respecto a la anterior (Tassile et al., 2013). En este caso los errores observados fueron desde 1,78 a 5,2 con un error promedio de 3,46 entre las temporadas 2002 y 2008 (Giménez, 2020), lo cual no se aleja de lo hallado en esta tesis.

Tabla 29: Comparación de la producción estimada por el pronóstico de producción regional de las provincias de Rio Negro y Meuquen, respecto a la producción total registrada por la secretaria de fruticultura de Rio Negro (en miles de toneladas)

Peras	Pronóstico	Producción total	Error	Error relativo
Año 2004	407	487	-80	16,42%
Año 2006	637	622	15	2,41%
Año 2007	622	626	-4	0,64%
Año 2008	603	600	-3	0,50%

Manzanas	Pronóstico	Producción total	Error	Error relativo
Año 2004	1087	796	291	36,55%
Año 2006	949	899	50	5,56%
Año 2007	946	905	41	4,53%
Año 2008	923	836	87	10,40%

Por su parte, Bóveda (2022), estimó la producción de mandarinas también con un modelo similar de pronóstico y obtuvo errores en los volúmenes de producción por huerto que oscilaron entre 1,72% y 10,43%. Estos porcentajes de error tampoco se alejan de los obtenidos en esta tesis mediante los pronósticos tradicionales.

Avanza (2010) por su parte estudio diferentes formas de estimación de carga en naranjos para predecir volúmenes de frutos a cosecha y los mejores resultados los obtuvo con el método de recuento total de frutos, con errores de estimación entre 3 y 15%. Estos resultados tampoco se alejan de los obtenidos en la presente tesis.

Miranda Jiménez & Royo Díaz, (2004) estimaron la producción en manzano cv 'Gala' y 'Golden Delicious' a partir de diferentes variables medidas en los árboles y las estimaciones de los rendimientos se alejaron aproximadamente el 15% de los valores reales. En peras cv 'Blanquilla' y 'Conference' realizaron un trabajo similar observando un error de aproximadamente el 23% entre los rendimientos esperados y observados, lo cual consideraron aceptable para el propósito perseguido (Miranda Jiménez & Royo Díaz, 2003). Los resultados obtenidos en esta tesis mediante el pronóstico por imágenes se acercan a los presentados por dichos autores, mientras que los resultados obtenidos mediante la metodología tradicional son mejores.

Otros autores que han trabajado en la estimación de los rendimientos de montes de manzana mediante uso de imágenes como Zhou et al., (2012) lograron coeficientes de determinación entre 0,58 y 0,71 para las estimaciones de cosecha a partir de frutos inmaduros y maduros, respectivamente. Por su parte, Linker (2016) utilizó regresión lineal múltiple y obtuvo desviaciones estándar de hasta el 27% de los rendimientos estimados promedio por árbol, mejorando los resultados en función del momento en que se toman las imágenes. Crotmir et al., (2012) con algoritmos de redes neuronales lograron correlaciones de 0,51 a 0,83 entre la cosecha estimada y real en función del momento y cultivar. Los resultados obtenidos por estos autores no superan en ningún caso los obtenidos en esta tesis.

Sin embargo, a futuro se continuara trabajando en mejorar el pronóstico de producción a partir de imágenes aumentando el número de árboles fotografiados por monte frutal (para obtener más datos de números de frutos por árbol), determinando la sección del árbol fotografiado para no incluir ramas de árboles vecinos, aumentando la cantidad de datos con los que se entrenarán los modelos, entre otros aspectos ya mencionados en el capítulo dos de esta tesis.

4.3.2. Comparación de costos y precisión de ambos pronósticos

EL valor del jornal establecido por la comisión nacional del trabajo para el 2023 fue de 4.733,28\$ para un peón general. Con dicho valor se calculó el valor de la hora de trabajo lo cual resulto 591,66\$ y fue el utilizado para construir el presupuesto ya que las actividades se contabilizaron en horas de trabajo.

Para calcular el presupuesto parcial solo se tuvo en cuenta el gasto derivado de aplicar una metodología específica, es decir, solamente se contempló el costo que se diferenció en una alternativa respecto de la otra (costos variables), y el resto de los costos no fueron afectados al cálculo (Harper et al., 2014).

Tabla 30: tareas y horas de trabajo asociados a cada metodología de pronóstico.

Tareas	Horas de trabajo		Costo asociado	
	Tradicional	Imágenes	Tradicional	Imágenes
<i>Conteo de frutos</i>	8	0	4.733,28	0
<i>Captura de imágenes</i>	0	2	0	1.183,32
<i>Medición diámetros azar</i>	4	4	2.366,64	2.366,64
<i>Curvas de crecimiento</i>	60	60	35.499,6	35.499,6
COSTO Total	72	66	42.599,52	39.049,56
Costo fijo en ambos			37.866,24	37.866,24
Costo variable			4.733,28	1.183,32

A partir de los resultados se puede ver que el costo asociado a la realización del pronóstico con la metodología tradicional suma 4.733,28\$ fuera del gasto en común que conlleva ambos pronósticos, mientras que la metodología de pronóstico a partir de imágenes implica 1.183,32\$. Se puede ver que la diferencia es de 3.549,96\$, es decir, la metodología tradicional es tres veces más costosa que la de imágenes.

Por su parte la precisión del pronóstico por imágenes fue un poco menor que la del pronóstico tradicional pero aceptable excepto en uno de los montes frutales, por lo cual a la hora de tomar la decisión se deberían considerar ambos aspectos.

A futuro se continuará trabajando en mejorar varios aspectos de la metodología a partir de imágenes en busca de lograr precisiones similares y disminuir los costos del pronóstico.

4.3.3. Estimación de la distribución de tamaños

A continuación, se presentan los gráficos obtenidos a partir de las frecuencias relativas observadas y esperadas para los tamaños de manzanas y peras en cada monte frutal.

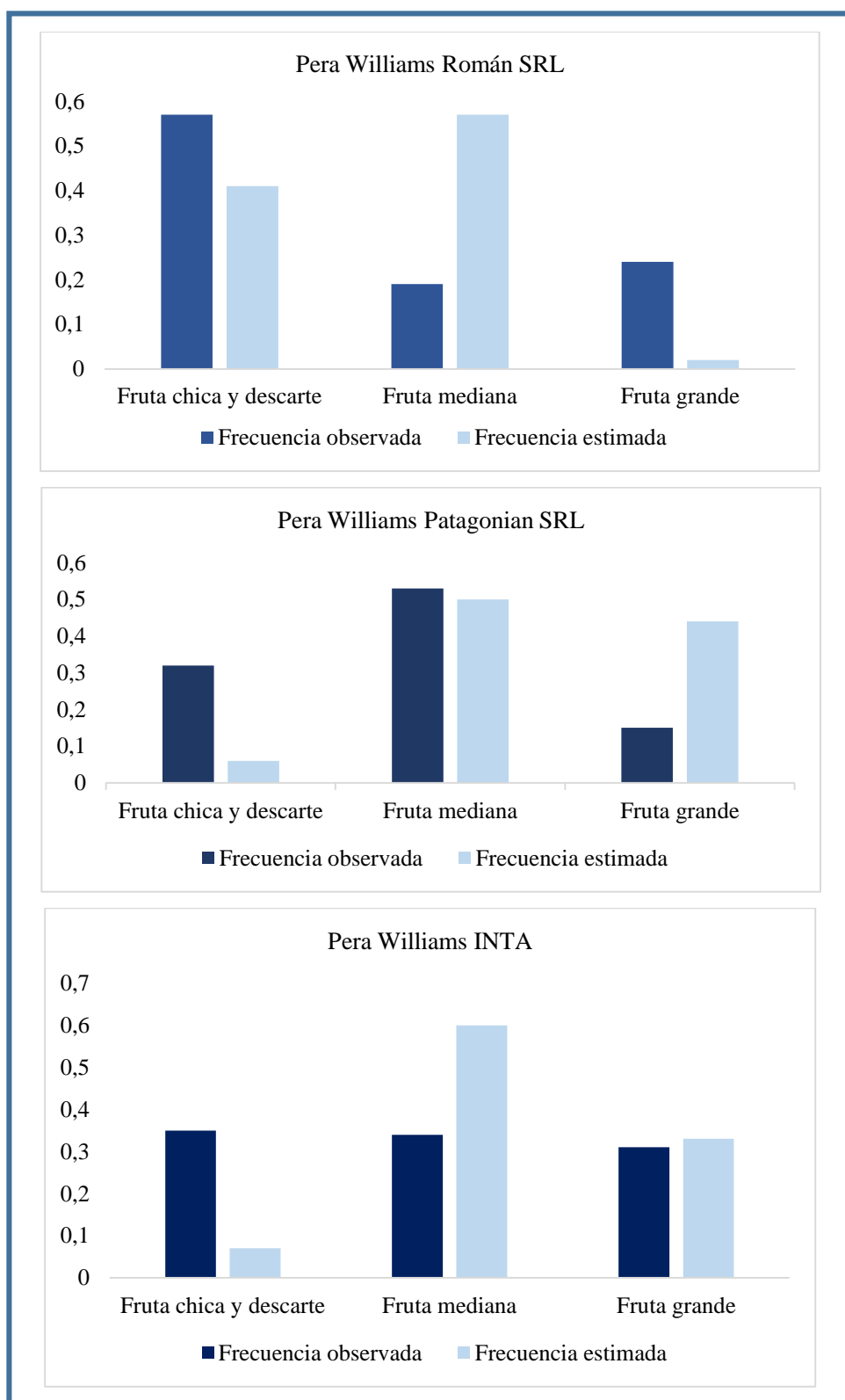


Figura 22: Histogramas de frecuencias relativas observadas y pronosticadas para cada clase de tamaño de frutos a cosecha en peras.

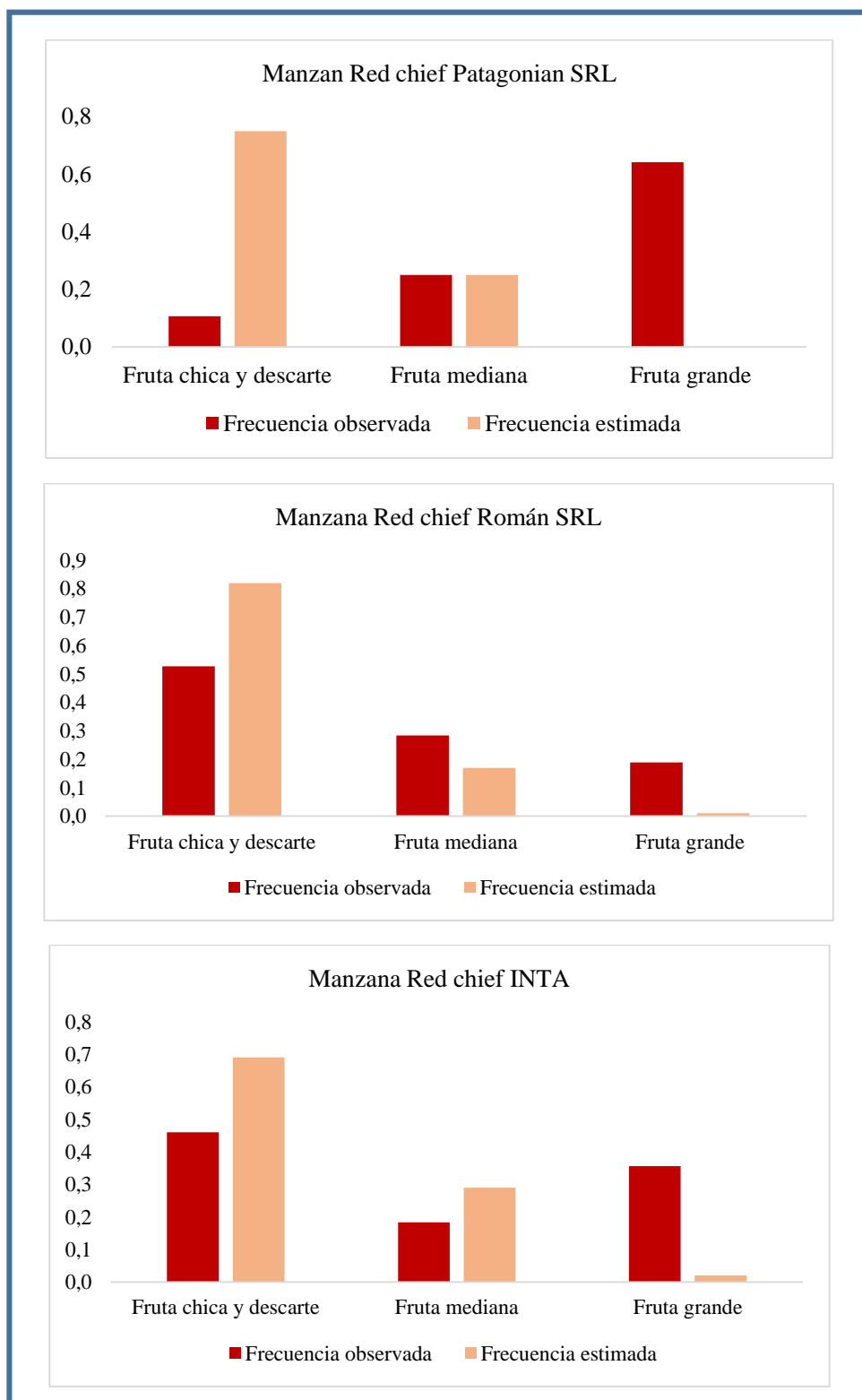


Figura 23: Histogramas de frecuencias relativas observadas y pronosticadas para cada clase de tamaño de frutos a cosecha en manzanas.

A partir de los resultados, se observó que la distribución de tamaños estimados o esperados al momento de cosecha no representó bien los tamaños finalmente obtenidos (Figura 22 y Figura 23). Las pruebas chi cuadrado arrojaron como resultado que las distribuciones de frecuencias de tamaños observadas no siguieron la distribución de frecuencias esperadas, con un nivel de significancia del 5% en todos los casos.

Otros autores han trabajado en la estimación de distribuciones de tamaños con resultados más alentadores. Bóbeda, (2020) realizó estimaciones de distribución de tamaños a cosecha de mandarinas y para ver cómo se comportaron los resultados construyó matrices de confusión y calculó la precisión global obteniendo que el 78% de la fruta fue clasificada correctamente en 8 categorías de tamaño, y las que no lo fueron, solo discreparon en una categoría. Por su parte, Giménez (2020) realizó también estimaciones de distribución de tamaños, matrices de confusión y calculó la precisión global, obteniendo valores aproximados de 0,6; y reduciendo la cantidad de categorías de tamaños logro mejorar la precisión global con valores de 0,77. En esta tesis, no se pudo realizar matrices de confusión porque los calibres para realizar la estimación de los tamaños a cosecha fueron tomados tempranamente al azar en frutos que no se marcaron, por lo cual no se obtuvo su tamaño final real a cosecha. Por ello, en esta tesis se comparó la distribución de la estimación de los tamaños a cosecha de 250 frutos medidos al azar en cada monte frutal versus la distribución de tamaños de los frutos cosechados de todas las plantas muestreadas (15 plantas en chacras de terceros, 60 plantas de pera en INTA y 72 plantas de manzana en INTA). Por este motivo también se decidió conveniente utilizar solo tres clases de tamaños, ya que los frutos tomados al azar para la predicción no se consideraron suficientes para establecer más clases. Similares conclusiones fueron presentadas por Marine et al., (2019), quienes midieron diámetros de manzanas a 60 DDPF en 93 árboles y estimaron su peso a cosecha para luego comparar estas distribuciones con las distribuciones reales, encontrando diferencias entre ambas distribuciones para todos los árboles estudiados. En su trabajo, miden entre el 7 y 30% de los frutos presentes en cada árbol, y concluyen que una estimación más precisa de la distribución del tamaño de la fruta requeriría medir aproximadamente el 50% de los frutos de cada árbol.

4.4. CONCLUSIONES

Los pronósticos realizados de forma tradicional presentaron errores y errores relativos aceptables tanto para peras como para manzanas (hasta 2,75 tn/ha y hasta 10%), por lo cual representan una herramienta de gran utilidad para estimar la producción a nivel de parcela.

Los pronósticos realizados a partir de imágenes no fueron tan precisos como los tradicionales, arrojando errores relativos desde tres a diez por ciento para manzanas y un poco mayores en peras. Esto se explicó por la estructura de los montes frutales utilizados y también por la dificultad del conteo de los frutos de un solo árbol en las imágenes. Sin embargo, en todos los huertos de manzana y en dos de los tres huertos de pera los errores relativos no superaron el 10%.

Estos resultados confirman la hipótesis dos planteada en esta tesis, es decir, es posible predecir anticipadamente la cantidad de frutos producidos en una parcela de frutales de pepita a partir del desarrollo de modelos de predicción que incluyan variables climáticas, de manejo del cultivo y tecnología de sensores.

Por su parte, el pronóstico tradicional arrojó errores menores al de imágenes, pero el costo de realización es bastante mayor respecto a este último. Considerando que, en cinco de los seis montes frutales estudiados, los errores fueron aceptables, y que es posible mejorarlo, el pronóstico por imágenes representa una alternativa muy interesante.

Este resultado responde a la hipótesis tres de esta tesis que plantea que las metodologías de conteo de frutos con sensores son precisas y menos costosas que el conteo manual.

Por último, las estimaciones de distribución de tamaños a cosecha no representaron adecuadamente las distribuciones obtenidas realmente, lo cual puede deberse a que el tamaño de muestra no fue representativo o lo suficientemente grande.

5. CONCLUSIONES GENERALES

El pronóstico de producción se basó en tres elementos: el número de frutos presentes en los árboles; la estimación del tamaño de los frutos a cosecha; y el número arboles del monte frutal. Para estimar el número de frutos, se utilizó, por un lado, la metodología de conteo manual y, por otro, se evaluaron metodologías de detección automática a partir de imágenes mediante algoritmos de visión artificial basados en redes neuronales. El número de frutos presentes en los árboles, tanto de peras como de manzanas, pudo ser estimado a partir de imágenes con una precisión aproximada de 0,85. Estos resultados pueden mejorarse si se aumenta la cantidad de datos de entrada a los modelos utilizados con más imágenes y etiquetas de frutas tomadas en la mayor diversidad de condiciones posibles. La posibilidad de estimar los frutos a partir de imágenes, reduce el tiempo de mano de obra destinada a dicha tarea, y esto permitiría aumentar el número de árboles muestreados para realizar los pronósticos.

Para estimar el tamaño de los frutos a cosecha, se ajustaron modelos mixtos no lineales que describen el crecimiento de los frutos de pera y manzanas para la región en estudio. Se contemplaron tanto efectos fijos como aleatorios a diferentes niveles como sitio, planta y fruto, y también se analizó la incorporación de covariables ambientales e índices verdes. El crecimiento fue descrito por la curva logística en su tercera parametrización. El uso de modelos mixtos multiniveles presentó la ventaja de poder discernir y cuantificar fuentes de variabilidad a diferentes niveles. El efecto aleatorio que mayor influencia tuvo en los parámetros de los modelos fue el fruto. La incorporación de variables ambientales e índices verdes mejoró el comportamiento de los modelos obtenidos, permitiendo contemplar la variabilidad asociada a cada temporada y sitio analizado. Si bien los modelos lograron describir bien el crecimiento de los frutos, a los fines del pronóstico no resultan prácticos si no están acompañados de un paquete tecnológico apropiado, por lo cual para hacer pronósticos, por el momento, se utilizarán los modelos ajustados sin los índices verdes. La capacidad predictiva de los modelos cuando se calculó con datos no contemplados en el proceso modelado fue razonable, con valores de sesgo medio cercanos a un mm en peras y menores para manzanas; y de sesgo medio relativo cercanos al dos y al uno por ciento para peras y manzanas respectivamente.

El número de árboles se obtuvo por registros del productor y con todos los elementos se realizaron dos pronósticos: i) basado en información recolectada por metodologías de conteo manual; ii) basado en información obtenida por imágenes. Ambos pronósticos resultaron buenos con errores relativos menores al 10 y 25% en manzanas y peras respectivamente.

Ambos pronósticos se compararon mediante un presupuesto parcial y como resultado se obtuvo que el pronóstico tradicional es tres veces más costoso que el realizado a partir de imágenes, pero en general, un poco más preciso.

A partir de los resultados surge también la necesidad de seguir trabajando a futuro en la mejora de la precisión de los algoritmos de detección de frutas a partir de imágenes y en el desarrollo de metodologías precisas de estimación de tamaños de frutos también a partir de imágenes. Además, sería interesante evaluar la posibilidad de construir un paquete tecnológico que permita medir (ya sea tamaños o número de frutos) y a la vez registrar las coordenadas geográficas de las respectivas mediciones para contar con datos de índices verdes de cada registro y poder incorporarlos a los pronósticos.

6. BIBLIOGRAFIA

1. ADARO, A., ALVAREZ, H., APCARIAN, A., ARAGÓN, J., ARUANI, M. C., BARNES, N., BEHEMER, S., BONDONI, M., CALVO, G., y OTROS, (2010). Pera Williams Manual para el productor y el empacador. Primera edición, Erregé y asociados. ISBN 978-987-25872-0-8
2. AGGELOPOULOU, K. D., WULFSOHN, D., FOUNTAS, S., GEMTOS, T. A., NANOS, G. D., BLACKMORE, S. (2010). Spatial variation in yield and quality in a small apple orchard, *Precision Agriculture*, 11:538-556. DOI: 10.1007/s11119-009-9146-9
3. ANNAMALAI, P. (2003). Citrus Yield Mapping System Using Machine Vision, ASAE Annual International Meeting, paper Number: 031002.
4. ARUANI, M. C., REEB, P. D., BARNES, N. E., (2014). Influence of soil properties on yield and fruit maturity at harvest of Williams pear. *Chilean Journal of Agricultural Research*, 74(4):460-467.
5. ASNER, G. P., (1998). Biophysical and Biochemical Sources of Variability in Canopy Reflectance. *Remote Sensing Environment*, 64:234–253.
6. ATAY, E., PIRLAK, L., ATAY, A.N. (2010). Determination of Fruit Growth in Some Apple Varieties Determination of Fruit Growth in Some Apple Varieties. *Journal of Agricultural Sciences*, 16(2010):1-8.
7. ÁVALOS-CERDAS, J. M., y VILLALOBOS-MONGE, A., (2018) Análisis económico: un estudio de caso en *Jatropha curcas* L. mediante la metodología de presupuestos parciales. *Agronomía Mesoamericana*, 29(1):95-104. DOI: 10.15517/ma.v29i1.27901.
8. AVANZA, M. M. (2010). 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 para el otorgamiento del grado académico de Doctor en el Área de Recursos Naturales, Universidad Nacional del Nordeste.
9. AVELLÁ, B., LANDRISCINI, S. G., PREISS, O., (2018). Complejo frutícola de Río Negro y Neuquén. Exportaciones, principales competidores y factores que condicionan la competitividad. *Revista Interdisciplinaria de Estudios Agrarios* N° 48.
10. BAGNATO, J. I., (2020). Aprende Machine Learning en Español, Teoría + Práctica Python. LeanPub.
11. BALZARINI, M., MACCHIAVELLI, R. E., CASANOVES, F. (2004). Aplicaciones de modelos mixtos en Agricultura y Forestería. *Aplicaciones de Modelos Mixtos*.

- Curso de capacitación Centro Agronómico Tropical de Investigación y Enseñanza. 189p. <https://www.researchgate.net/publication/28349108>
12. BARGOTI, S. y UNDERWOOD, J. 2017. Deep fruit detection in orchards. arXiv:1610.03677.v2[cs.RO]
 13. BENÍTEZ, C. (2001). Cosecha y poscosecha de peras y manzanas en los Valles irrigados de la Patagonia. Primera Edición, Estación experimental Patagonia Norte. ISBN: 987-521-033-1
 14. BERGH, O. (1982). Prediction of apple fruit size: a promising model. *Agroplanta*, 14(3):43-46.
 15. 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):11–18.
 16. BOBEDA, G.R.R. (2022). Modelo de predicción de volúmenes de producción y distribución de tamaño de frutos a cosecha en mandarino 'tangor Murcott'. Tesis para el otorgamiento del Grado Académico de: Doctor de la Universidad Nacional del Nordeste en Recursos Naturales- Doctor de la Universidad de Oviedo en Informática.
 17. BRAMARDI, S.J. (1995). Modelos de predicción de tamaños de fruta a la cosecha en base a mediciones sucesivas durante el período de crecimiento. Peras cvs.Wiliam'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.
 18. 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*, 475:283-293.
 19. BRAMARDI, S.J., TASSILE, V., REEB, P., DE BERNARDIN, F. (2005). Comparación de métodos para la estimación anticipada del peso del fruto medio a cosecha: Curvas de crecimiento vs. Modelos estocásticos. X Reunión Científica del Grupo Argentino de Biometría. Corrientes, Argentina.
 20. BRAMARDI, S. J., TASSILE, V., REEB, P. D., (2006). Crecimiento de frutos de pepita en el Alto Valle. Tablas de raleo. Editor: Área Estadística, Facultad De Ciencias Agrarias. Universidad Nacional Del Comahue.
 21. BRESILLA, K., DEMETRIO PERULLI, G., BOINI, A., MORANDI, B., CORELLI GRAPPADELLI, L, MANFRINI, L., (2019). Single-Shot Convolution Neural

- Networks for Real-Time Fruit Detection Within the Tree. *Front. Plant Sci.* DOI: 10.3389/fpls.2019.00611
22. BULANON, D. M., y KATAOKA, T. (2010). A Fruit Detection System and an End Effector for Robotic Harvesting of Fuji Apples, *Agricultural Engineering International: the CIGR Ejournal*. Manuscript 1285, Vol. XII:1–14.
 23. CALVO, G. (2008). Impacto del 1-Metilciclopropeno en cultivares de peras tempranas producidas en el Alto Valle de Río Negro. Tesis de Maestría en producción Vegetal, Facultad de Agronomía, Universidad de Buenos Aires.
 24. CALVO, G., GOMILA, T., MOLINA, G., (2013). Respuesta a la aplicación de 1MCP en pera Williams en función del estado de madurez caracterizado por el índice de diferencia de absorbancia (DA). *Revista Iberoamericana de Tecnología de Pos cosecha* ISSN: 1665-0204.
 25. CECCATO, P., FLASSE, S., TARANTOLA, S., JACQUEMOUD, S., GRÉGOIRE, J.M. 2021. Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sensing of Environment*, 77(1): 22-33, ISSN 0034-4257, DOI/10.1016/S0034-4257(01)00191-2
 26. CHOLLET, F. (2018). *Deep Learning with Python*. ISBN 9781617294433
 27. CHEN, S. W., SHIVAKUMAR, S. S., DCUNHA, S., DAS, J., OKON, E., QU, C., TAYLOR, C. J., KUMAR, V., (2016). Counting apples and oranges with deep learning: a data driven approach. *IEEE robotics and automation letters*. DOI:10.1109/LRA.2017.2651944
 28. COHEN, O., LINKER, R., NAOR, A. (2011). Estimation of the number of apples in color images recorded in orchards. *International Federation for Information Processing, Part I*, AICT 344:630-642.
 29. CPIARN (2015). Aportes para la Reconstrucción de una Fruticultura Sustentable. Technical report, Consejo Profesional de Ingeniería Agronómica Río Negro.
 30. ČRTOMIR, R., URŠKA, C., STANISLAV, T., DENIS, S., KARMEN, P., PAVLOVIČ, M. (2012). Application of Neural Networks and Image Visualization for Early Forecast of Apple Yield. *Erwerbs-Obstbau* 54:69–76. DOI: 10.1007/s10341-012-0162-y
 31. DE OLIVEIRA APARECIDO, L. E., FERREIRA, R. B., DE SOUZA ROLIM, G., DE SOUZA, B., DE SOUZA, P. S. (2016). Nonlinear agrometeorological models for estimating lychee fruit growth. *Revista Brasileira de Fruticultura*, 39(2):1-14. DOI 10.1590/0100-29452017 169

32. DE SILVA, H.N., TUSTIN, D.S., CASHMORE, W.M., STANLEY, C.J., LUPTON, G., McARTNEY. (1997). Fruit fresh mass-diameter relationship for 'Royal Gala' Apple across seasons and among fruit production regions of New Zealand. *Hortscience*, 32:1169-1173.
33. DORJ, U. O., LEE, M., YUN, S. (2017). An yield estimation in citrus orchards via fruit detection and counting using image processing. *Computers and Electronics in Agriculture*, 140, 103–112. DOI: 10.1016/j.compag.2017.05.019
34. ESCOLA, A. (2017). YOLO Object Detector for Onboard Driving Images. Tesis para acceder al otorgamiento del título de Computer Science, Universidad Autónoma de Barcelona.
35. ESCOLA, A., ARNO, J., SANZ, R., CAMP, F., MASIP, J., SOLANELLES, F., ROSELL, J. R., PLANAS, S. (2009) Sensing tree canopy parameters in real time for precision fruticulture/horticulture applications: Methodology set-up and first results. Conference paper.
36. FARJON, G., KRIKEB, O., BAR, A., AND VICTOR, H. (2019). Detection and counting of flowers on apple trees for better chemical thinning decisions. *Precision Agriculture*, 21:503-521. DOI: 10.1007/S11119-019-09679-1
37. 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
38. FERREYRA, P. (2018). Presupuesto parcial en un rodeo Aberdeen Angus con doble dosis de PGF2 α de diferentes marcas comerciales. Tesina de la Orientación Producción Animal, presentada como parte de los requisitos para optar al grado de Veterinario del estudiante. Facultad de Ciencias Veterinarias, Universidad Nacional del Centro de la Provincia de Buenos Aires.
39. FU, L., SUN, H., SHARMA, C., YUANCAI, L., ZHANG, H., y TANG, S. (2013). Nonlinear mixed-effects crown width models for individual trees of chinese fir (*cunninghamia lanceolata*) in south-central china. *Forest. Ecology and Management*, 302:210–220
40. GALAN, C., VAZQUEZ, L., GARCIA-MOZO, H., DOMINGUEZ, E., (2004). Forecasting olive (*Olea europaea*) crop yield based on pollen emission. *Field Crops Research*, 86: 43-51. DOI: 10.1016/S0378-4290(03)00170-9
41. GARRIZ, P. I. COLAVITA, G.M., ALVAREZ, H.L., SPERA, N.,BLACKHALL, V. (2006). A Model for Predicting Diameter of 'Red Sensation' Pears. *Proc. IXth IS*

- on Modelling in Fruit Research and Orchard Management, 147-152.
42. GEMTOS, T., FOUNTAS, S., TAGARAKIS, A., LIAKOS, V., (2013). Precision Agriculture Application in Fruit Crops: Experience in Handpicked Fruits. *Procedia Technology*, 8:324-332. DOI: 10.1016/j.protcy.2013.11.043
 43. GERON, A., (2005). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc., Canadá. ISBN: 978-1-492-03264-9
 44. GIMENEZ, G. N. (2012). Determinación de los patrones de crecimiento y abscisión de frutos para ajustar el momento oportuno de raleo en perales cv 'Williams'. Tesis para el otorgamiento del grado académico de Magister en Fruticultura de Clima Templado-Frío, Universidad Nacional del Comahue, Argentina.
 45. GIMENEZ, G.N. (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.
 46. GIMENEZ, G. N., y 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. XV Conferencia Española y V Encuentro Iberoamericano de Biometría. Bilbao
 47. GIRSHICK, R. 2015. FastR-CNN. In: *Proceedings of the IEEE International Conference on Computer Vision*. p.1440-1448. <https://doi.org/10.1109/ICCV.2015.169>.
 48. GIRSHICK,R., DONAHUE,J., DARRELL,T., MALIK,J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 580–587. <https://doi.org/10.1109/CVPR.2014.81>.
 49. GLENN, D. M., y TABB, A. (2019). Evaluation of Five Methods to Measure Normalized Difference Vegetation Index (NDVI) in Apple and Citrus. *International Journal of Fruit Science*, 19(2):191-210. DOI: 10.1080/15538362.2018.1502720
 50. GONGAL, A., AMATYA, S., KARKEE, M., ZHANG, Q., LEWIS, K. (2015). Sensors and systems for fruit detection and localization : A review. *Computers and Electronics in Agriculture*, 116:8-19. DOI:10.1016/j.compag.2015.05.021
 51. GOUDRIAAN, J., y MONTEITH, L. (1990). A mathematical function for crop growth based on light interception and leaf area expansion. *Annals of Botany*,

- 66(6):695–701.
52. HALL D.B., y BAILEY R.L. (2001). Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *Forest Science*, 47:311–321
 53. HARPER, J. K., CORNELISSE, S., KIME, L. F., HYDE, J.(2014). Alternativas agrícolas: Presupuestos para tomar decisiones agrícolas. PennState Cooperative Extension, College of Agricultural Sciences. The Pennsylvania State University, Code: EE0092S 05/14.
 54. HENTEN, E., HEMMING, J., TUIJL, B., KORNET, J. G., MEULEMAN, J., BOX, P. O. (2002). An Autonomous Robot for Harvesting Cucumbers in Greenhouses, *Autonomous robots*, 13:241–258. DOI: 10.1023/A:1020568125418
 55. HERNANDEZ REYES, M., (2001). Análisis económico de experimentos agrícolas con presupuestos parciales: Re-enseñando el uso de este enfoque. Boletín informativo del Centro de Información Agrosocioeconómica, N°1-2021, Universidad de San Carlos de Guatemala. https://www.researchgate.net/publication/334655730_Analisis_economico_de_experimentos_agricolas_con_presupuestos_parciales_Reensenando_el_uso_de_esto_enfoque
 56. HOČEVAR, M., ŠIROK, B., GODEŠA, T., STOPAR, M. (2014). Flowering estimation in apple orchards by image analysis. *Precision Agriculture*, 15(4):466–478. DOI: 10.1007/s11119-013-9341-6
 57. HUNT, E.R y ROCK, B.N. 1989. Detection of changes in leaf water content using Near- and Middle-Infrared reflectances. *Remote Sensing of Environment*, 30(1): 43-54. ISSN 0034-4257, [https://doi.org/10.1016/0034-4257\(89\)90046-1](https://doi.org/10.1016/0034-4257(89)90046-1).
 58. ITAKURA, K., NARITA, Y., NOAKI, S., HOSOI, F., (2021). Automatic pear and apple detection by videos using deep learning and a Kalman filter. *Osa Continuum*. Vol 4, N°5. DOI: 10.1364/OSAC.424583
 59. JIANG, L. y LI, Y. (2010) Application of Nonlinear Mixed-Effects Modeling Approach in Tree Height Prediction. *Journal of Computers*, 5(10):1575–1581. ISSN1796-203X. doi:10.4304/jcp.5.10.1575-1581.
 60. JIMENEZ, C.M. y ROYO DIAZ, J.B.R. 2003 . Statistical Model Estimates Potential Yields in Pear Cultivars ‘Blanquilla’ and ‘Conference’ before Bloom. *Amer. Soc. Hort. Sci.*, 128(4):452-457. <https://doi.org/10.21273/JASHS.128.4.0452>
 61. KAACK, K. y LINDHARD PEDERSEN, H. (2010). Prediction of Diameter, Weight

- and Quality of Apple Fruit (*Malus domestica* Borkh.) cv. ‘Elstar’ using Climatic Variables and their Interactions. *Europ. J. Hort. Sci.*, 75:(2),60-70. ISSN 1611-4426.
62. KANE, K. E., AND LEE, W. S. (2006). Spectral Sensing of Different Citrus Varieties for Precision Agriculture, ASABE Meeting Presentation, USA, paper number 061065.
 63. KARKEE, M., y ZHANG, Q. (2012). Mechanization and Automation Technologies in Specialty Crop Production, *Resource Magazine*. 19(5):16-17. DOI: 10.13031/2013.42314
 64. KOIRALA, A.; WALSH, K. B.; WANG, Z.; McCARTHY, C. 2019a. Deep learning for real time fruit detection and orchard fruit load estimation: benchmarking of ‘MangoYOLO. *Precision Agriculture*, 20:1107-1135. <https://doi.org/10.1007/s11119-019-09642-0>
 65. KOIRALA, A., WALSH, K. B., WANG, Z., McCARTHY, C., 2019b. Deep Learning – Method overview and review of use for fruit detection and yield estimation. *Computers and Electronics in Agriculture*. 162:219-234. DOI:10.1016/j.compag.2019.04.017
 66. KONISHI S. y KITAGAWA G. (2007). *Information Criteria and Statistical Modeling*. Springer, New York.
 67. KUMAR, A., SHARMA, P., AMBRAMMAL, S. K., (2014). Climatic Effects on Food Grain Productivity in India: A Crop Wise Analysis. *Journal of Studies in Dynamics and Change*, 1(1), 38-48. ISSN: 2348-7038
 68. KURTULMUS, F., LEE, W. S., VARDAR, A. (2011). Green citrus detection using ‘eigenfruit’, color and circular Gabor texture features under natural outdoor conditions. *Computers and Electronics in Agriculture*, 78(1):140-149. <https://doi.org/10.1016/j.compag.2011.07.001>
 69. KURTULMUS, F., LEE, W. S., VARDAR, A., (2014). Immature peach detection in colour images acquired in natural illumination conditions using statistical classifiers and neural network. *Precision Agric*, 15:57–79. DOI 10.1007/s11119-013-9323-8
 70. LAKSO, A. N., GRAPPADELLI, L. C., BARNARD, J., GOFFINET, M. C., GRAPPADELLI, L. C., BARNARD, J., AN, M. C. G. (1995). An expolinear model of the growth pattern of the apple fruit An expo linear model of the growth pattern of the apple fruit. *Journal of Horticultural Science*, 70(3):389-394. DOI: 10.1080/14620316.1995.11515308
 71. LIANG, Q., ZHU, W., LONG, J., WANG, Y., SUN, W., WU, W., (2018). A real-time detection framework for on-tree mango based on SDD Network. *Springer Nature*

- Switzerland, 423-436. DOI: 10.1007/978-3-319-97589-4_36
72. LINDSTROM, M. L. y BATES, D. M., (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46(3):673-687
 73. LINKER, R. (2016). A procedure for estimating the number of green mature apples in night-time orchard images using light distribution and its application to yield estimation. *Precision Agriculture*, 18:59-75. DOI:10.1007/s11119-016-9467-4
 74. LIU H., ZHENG Y., SHEN J., (2008). Goodness-of-fit measures or r^2 for repeated measures mixed effect models. *Applied Statistics*, 35(10):1081–1092.
 75. LIU, W., ANGUELOV D., ERHAN, D., SZEGEDY, C., REED, S., FU, C., BERG, A.C. 2016. Ssd: Singles hot multibox detector. In: *European Conference on Computer Vision*. Springer, 21–37. https://doi.org/10.1007/978-3-319-46448-0_2.
 76. MAGEIN, H. (1989). Growth and abscission dynamics of ‘Cox’s Orange Pippin’ and ‘Golden Delicious’ apple fruits. *Journal of Horticultural Science*, 64(3):265-273, DOI: 10.1080/14620316.1989.11515954
 77. MAHESWARI, P., RAJA, P., APOLO-APOLO, E., PÉREZ-RUIZ, M. (2021). Intelligent Fruit Yield Estimation for Orchards Using Deep Learning Based Semantic Segmentation Techniques-A Review. *Frontiers in plant Science*, v12:1-18. DOI: 10.3389/fpls.2021.684328
 78. MANN, K. K., SCHUMANN, A. W., OBREZA, T. A., (2011). Delineating productivity zones in a citrus grove using citrus production, tree growth and temporally stable soil data. *Precision Agriculture*, 12:457-472. DOI 10.1007/s11119-010-9189-y
 79. MAO, W., JIA, B., ZHANG, X., HUB, X. (2009). Detection and position method of apple tree image. *Computer and Computing Technologies in Agriculture II. IFIP Advances in Information and Communication Technology*, 294. DOI: 10.1007/978-1-4419-0211-5_29
 80. 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
 81. MARTÍNEZ-CASASNOVAS, J. A. y BORDES, X (2005). Viticultura de precisión: Predicción de cosecha a partir de variables del cultivo e índices de vegetación. *Revista de Teledetección*, 24:67-71

82. MAYORANO, F., RUBIALES, A., HERRERO, V., CLAUSSE, A. (2006). Modelo Computacional para el pronóstico de tamaño de frutas. *RIA*. 35 (2): 143-162.
83. MCARTNEY, S., PALMER, J. W., ADAMS, H. M. (1996). Crop loading studies with ‘ Royal Gala ’ and ‘ Braeburn ’ apples : Effect of time and level of hand thinning, 0671. DOI: 10.1080/01140671.1996.9513977
84. MIRANDA JIMÉNEZ, C. M. y ROYO DÍAZ, J. B., (2003). Statistical Model Estimates Potential Yields in Pear Cultivars ‘ Blanquilla ’ and ‘ Conference ’ before Bloom, 128(4):452–457.
85. MIRANDA JIMENEZ, C. M. y ROYO DIAZ, J. B. (2004). Statistical model Estimates potential yields in ‘Golden Delicious’ and ‘Royal Gala’ before bloom. *Journal of American society of Horticultural science* 129(1):20-25.
86. NAGY, A., FEHÉR, J., TAMÁS, J. (2018). Wheat and maize yield forecasting for the Tisza river catchment using MODIS NDVI time series and reported crop statistics. *Computers and Electronics in Agriculture*, 151:41-49. DOI: 10.1016/j.compag.2018.05.035
87. PARRISH, E. A., y GOKSEL, A. K. (1977). Pictorial Pattern Recognition Applied to Fruit Harvesting. *ASAE*, 20(5):0822-0827. DOI: 10.13031/2013.35657
88. PASQUA PAULINO, S. E., ALVES MOURÃO FILHO, F., NUNES MAIA, A., CANTUARIAS AVILÉS, T. E., NETO, D.D. (2007). Agrometeorological models for ‘Valencia’ and ‘Hamlin’ sweet oranges to estimate the number of fruits per plant. *Science Agriculture*, 64(1):1-11.
89. PERRY, E. M., DEZZANI, R. J., SEAVERT, C. F., PIERCE, E. F., (2010). Spatial variation in tree characteristics and yield in a pear orchard. *Precision Agriculture*, 11:42–60. DOI: 10.1007/s11119-009-9113-5
90. PINHEIRO, J. C. y BATES, D. M., (2000). *Mixed-Effects Models in S and S-Plus*. Springer, Nueva York. ISBN 0-387-98957-0
91. PORTIANSKY, E. L., (2013). *Análisis multidimensional de Imágenes digitales*. Primera Edición. Universidad Nacional de La Plata. ISBN: 978-950-34-1023-3
92. QUANG, A. V., DELBART, N., JAFFRAIN, G., PINET, C., MOIRET, A. (2022). Detection of degraded forests in Guinea, West Africa, based on Sentinel-2 time series by inclusion of moisture-related spectral indices and neighbourhood effect. *Remote Sensing of Environment*, 281:113230. <https://doi.org/10.1016/j.rse.2022.113230>
93. QURESHI, W. S., PAYNE, A., WALSH, K. B., LINKER, R., COHEN, O., DAILEY, M. N. (2016). Machine vision for counting fruit on mango tree canopies. *Precision*

- Agriculture. DOI: 10.1007/s11119-016-9458-5
94. R CORE TEAM R. 2018. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing. <http://www.R-project.org/>
 95. RATKOWSKY, D. A., (1983): Nonlinear Regression Modeling: a Unified Practical Approach, Marcel Dekker Inc., New York, USA.
 96. REDMON, J., DIVVALA, S., GIRSHICK, R., FARHADI, A. 2016. You only look once: unified, real time object detection. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 779–788. <https://doi.org/10.1109/CVPR.2016.91>.
 97. REDMON, J. y FARHADI, A., (2017). YOLO9000: Better, Faster, Stronger. arXiv:1612.08242v1
 98. REDMON, J. y FARHADI, A., (2018). YOLO v3: An incremental improvement. ArXiv preprint. Computer Vision and Pattern Recognition. <https://doi.org/10.48550/arXiv.1804.02767>
 99. REEB, P., BRAMARDI, S., ALVAREZ, O. (2003). Estudio de la variabilidad de la producción de manzanas Red Delicious en los montes frutales del Alto Valle de Rio Negro, Argentina. Agro sur, 31(1): 21-26. ISSN 0304-8802.
 100. REEB, P.D., BRAMARDI, S.J., TASSILE, V., GIMENEZ, G.N., CURETTI, M. ALVAREZ, H. (2018). Tabla de crecimiento de frutos. Departamento de Estadística, Universidad Nacional del Comahue. Proyecto de Extensión: Aporte a la planificación de tareas culturales en montes de peras y manzanas utilizando curvas de crecimiento de frutos en Rio Negro y Neuquen.
 101. REN, S., HE, K., GIRSHICK, R., SUN, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Advances in Neural Information Processing Systems, 91-99. DOI:10.48550/arXiv.1506.01497
 102. RODRIGUEZ, A. B. (2011). Influencia de la temperatura del aire y del raleo sobre el tamaño del fruto de pera (*Pyrus comunnis*) cv Williams durante sus distintas fases de crecimiento. Tesis para el otorgamiento del grado académico de Magíster en Ciencias Agropecuarias. Universidad Nacional de Córdoba.
 103. RODRIGUEZ A. B. y MUÑOZ A. R., (2005). Síntesis agrometeorológica para el período 1990-2004. EEA INTA Alto Valle; Boletín de divulgación técnica N° 53; INTA; Alto Valle del Río Negro; 38 pp.
 104. ROZADA RANEROS D.S. (2021). Estudio de la arquitectura YOLO para la detección

- de objetos mediante deep learning. Tesis para el otorgamiento del grado académico de Máster en Ingeniería de Telecomunicación, Universidad de Valladolid.
- 105.SA, I., GE, Z., DAYOUB, F., UPCROFT, B., PEREZ, T., MCCOOL, C., (2016). DeepFruits: A Fruit Detection System Using Deep Neural Networks. *Sensors*: 16: 1222. DOI:10.3390/s16081222
 - 106.SALVO, S., MUÑOZ, C., AVILA, J., BUSTOS, J., RAMIREZ-VALDIVIA, M., SILVA, C., VIVALLO, G. (2012). An estimate of potential blueberry yield using regression models that relate the number of fruits to the number of flower buds and to climatic variables. *Scienza Horticulturae*, 133:56-63. DOI:10.1016/j.scienta.2011.10.020
 - 107.SANCHEZ E. E. y VILLAREAL P. (2011); Programa Nacional Frutales; Estación Experimental Agropecuaria Alto Valle; Instituto Nacional de Tecnología Agropecuaria; Río Negro, Argentina.
 - 108.SERVICIO NACIONAL DE SANIDAD Y CALIDAD AGROALIMENTARIA (SENASA), (2021). Anuario Estadístico 2021, Centro Regional Patagonia Norte, Argentina 2022.
 - 109.SCHNEIDER, C., RASBAND, W. y ELICEIRI, K. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9, 671–675. <https://doi.org/10.1038/nmeth.2089>
 - 110.SOBER E. (2002). Instrumentalism, parsimony and the akaike framework. *Philosophy Science*, 69:112–123,
 - 111.SOZZI G. O. (2007). Árboles frutales: ecofisiología, cultivo y aprovechamiento. Editorial Facultad de Agronomía, Universidad de Buenos Aires. ISBN: 950-29-0974-7.
 112. STAJNKO, D., y ÈMELIK, Z. (2005). Modelling of Apple Fruit Growth by Application of Image Analysis. *Agriculturae Conspectus Scientificus*, Vol. 70(2): 59-64.
 113. STAJNKO, D, RAKUN, J., BLANKE, M. (2009). Modelling Apple Fruit Yield Using Image Analysis for Fruit Colour, Shape and Texture. *European Journal Horticulture Science*, 74(6):260–267.
 114. 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.

115. STANLEY, C., TUSTIN, D., LUPTON, G., MCARTNEY, S., CASHMORE, W., DE SILVA, H. (2000). Towards understanding the role of temperature in apple fruit growth responses in three geographical regions within New Zealand. *J. of Hort. Sci. and Biotechnology* 74 (4), 413-422.
116. STORTI M., (2008). Tesis de grado: Determinación de Parámetros hidrofísicosquímicos de los suelos relacionados con los procesos de degradación salina y drenaje restringido; Escuela Superior de Salud y Ambiente; Universidad Nacional del Comahue, Neuquén, Argentina.
117. TASSILE, V., (2020). Aplicación de modelos mixtos no lineales en el crecimiento de frutos de peral (*Pyrus communis*) cv Williams. Tesis para el otorgamiento del grado académico de Magister en Estadística Aplicada, Universidad Nacional de Córdoba, Argentina.
118. TASSILE V., BRAMARDI S., REEB P., DE BERNARDIN F. (2005). Modelos de crecimiento de frutos de pepita: mejora en la precisión de las estimaciones a través de la inclusión de covariables ambientales y componentes aleatorias. Presentado en X Reunión Científica del Grupo Argentino de Biometría, Universidad Nacional del Nordeste, Corrientes.
119. 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.
120. TIAN, Y., YANG, G., WANG, Z., WANG, H., LI, E., LIANG, Z., (2019). Apple detection during different growth stages in orchards using the improved YOLO-V3 model. *Computers and electronic in agricultura.* 157: 417-426. DOI: 10.1016/j.compag.2019.01.012
121. VARGAS BAEZA, V. (2020). Sistema de visión artificial para el control de calidad en piezas cromadas. Tesis para el otorgamiento del grado académico de maestro en ingeniería de manufactura. Instituto politécnico nacional de Mexico.
122. VELEZ, J.P., SCARAMUZZA, F., VILLARROEL, D. (2016) Teledetección: qué tecnología se adapta mejor para lograr un buen diagnóstico del cultivo. Recopilación de presentaciones técnicas del curso internacional de agricultura y ganadería de precisión con agregado de valor en origen. 1ra Edición, Ediciones INTA, Buenos

Aires.

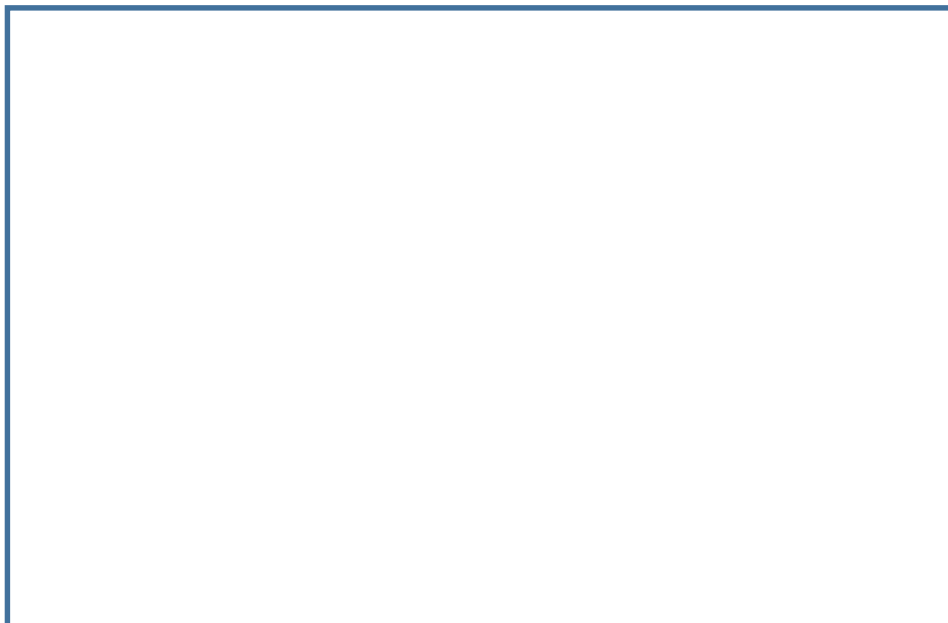
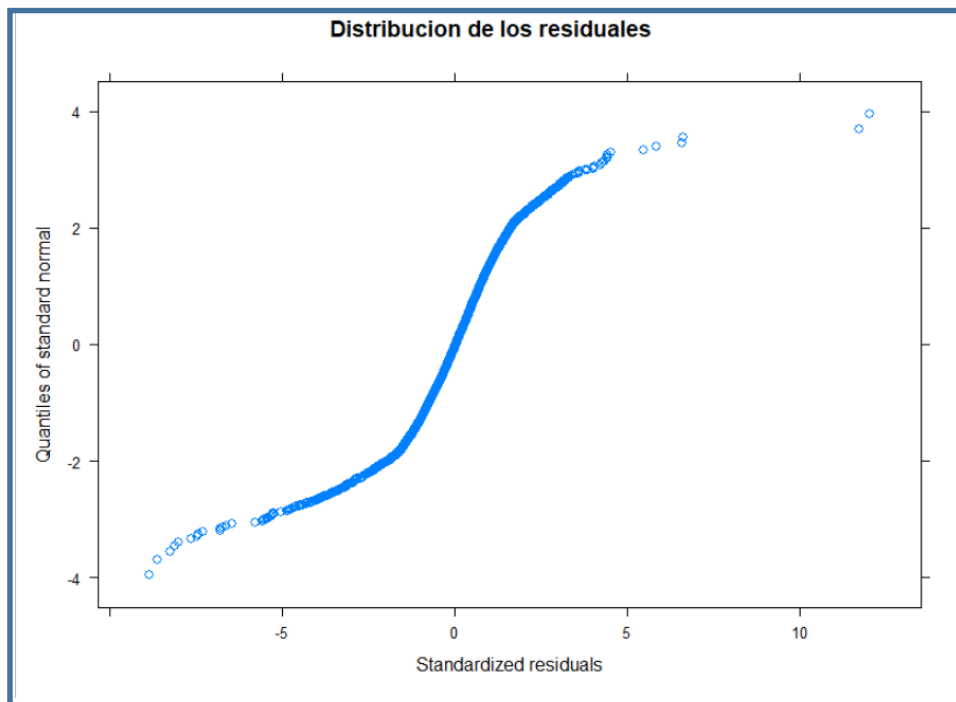
123. VIAU, A. A., JANG, J., PAYAN, V., DEVOST, A., (2005). The Use of Airborne LIDAR and Multispectral Sensors for Orchard Trees Inventory and Characterization. *Information and Technology for Sustainable Fruit and Vegetable Production*, FRUTIC 05:12-16.
124. VILLATORO, B. (2020). Uso de índices de vegetación NDWI y MSI obtenidos de imágenes satelitales sentinel-2 y landsat 8 para monitoreo de humedad, como indicador de momento óptimo de cosecha en caña de azúcar. CENGICANÑA: Memoria Presentación de resultados de investigación Zafra.
125. VONESH E.F. (1996). A note on the use of laplace's approximation for nonlinear mixed-effects models. *Biometrika*, (83):447–452.
126. WACHS, J. P., STERN, H. I., BURKS, T., ALCHANATIS, V. (2010). Low and high-level visual feature-based apple detection from multi-modal images, *Precision Agriculture*, 11:717–735. DOI: 10.1007/s11119-010-9198-x
127. WANG, C. Y., BOCHKOVSKIY, A., LIAO, H. Y. M (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv:2207.02696v1. DOI:10.48550/arXiv.2207.02696
128. WANG, D. y HE, D., (2021). Channel Pruned YOLOv5s based-Deep learning approach for rapid and accurate apple fruitlet detection before fruit thinning. *Biosystem engineering*, 210: 271-281. DOI:1016/j.biosystemseng.2021.08.015
129. WANG, Q., NUSKE, S., BERGERMAN, M., SINGH, S. (2012). Automated Crop Yield Estimation for Apple Orchards, *Experimental Robotics*, Springer Tracts in Advanced Robotics, 88. 1DOI:10.1007/978-3-319-00065-7_50
130. WULFSOHN, D., ZAMORA, F. A., TÉLLEZ, C. P., LAGOS, I. Z., GARCÍA-FIÑANA, M. (2012). Multilevel systematic sampling to estimate total fruit number for yield forecasts. *Precision Agriculture*, 13(2), 256–275. DOI:10.1007/s11119-011-9245-2
131. YANG, Y. y HUANG, S. (2013). On the statistical and biological behaviors of nonlinear mixed forest models. *European Journal Forest*, 132:727–736.
132. YE, X., SAKAI, K., ASADA, S., SASAO, A., (2008). Application of narrow-band TBVI in estimating fruit yield in citrus *Biosystems Engineering*, 99(2):179-189 doi:10.1016/j.biosystemseng.2007.09.016
133. ZADRAVEC, P. (2015). Fruit Growth Patterns of Four Apple Cultivars Using Nonlinear Growth Models *European Journal Horticulture Science*, 79(2):52–59, 2014,

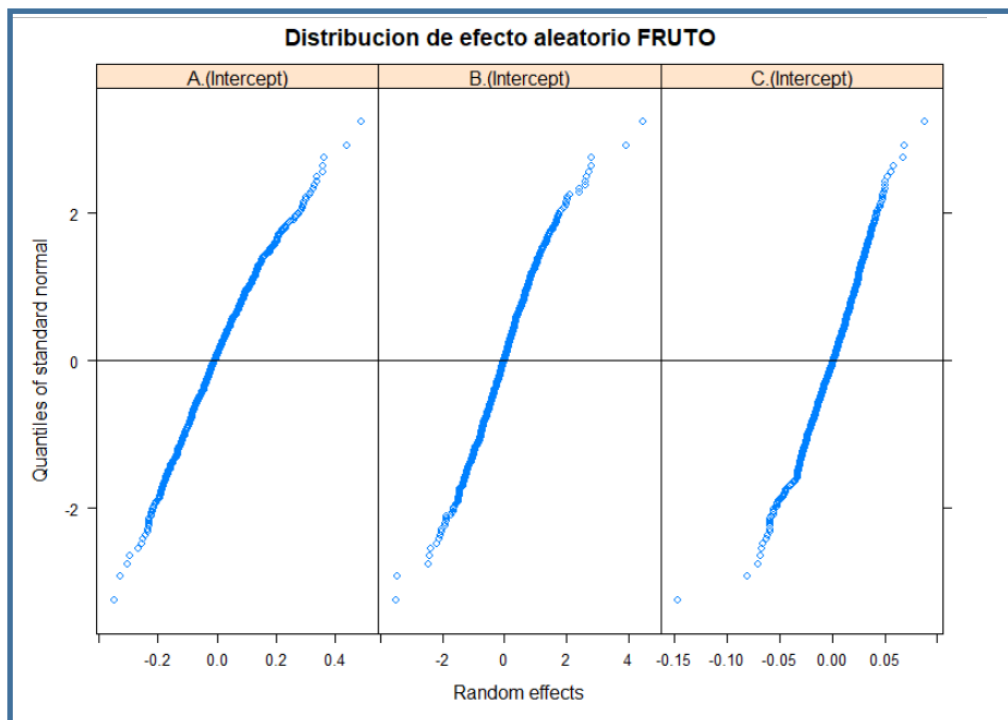
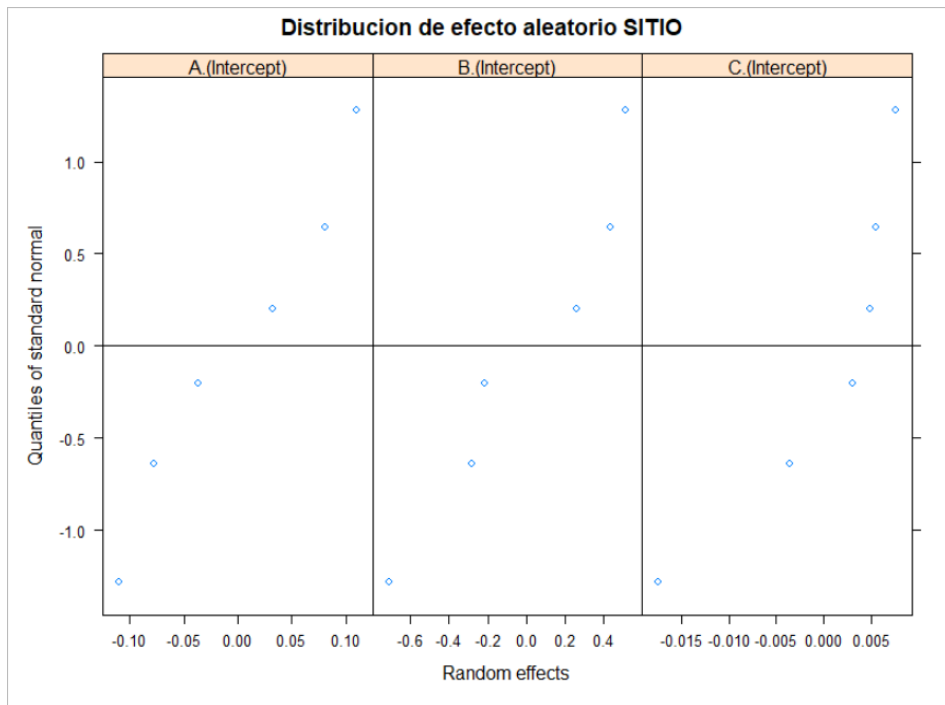
ISSN 1611-4426.

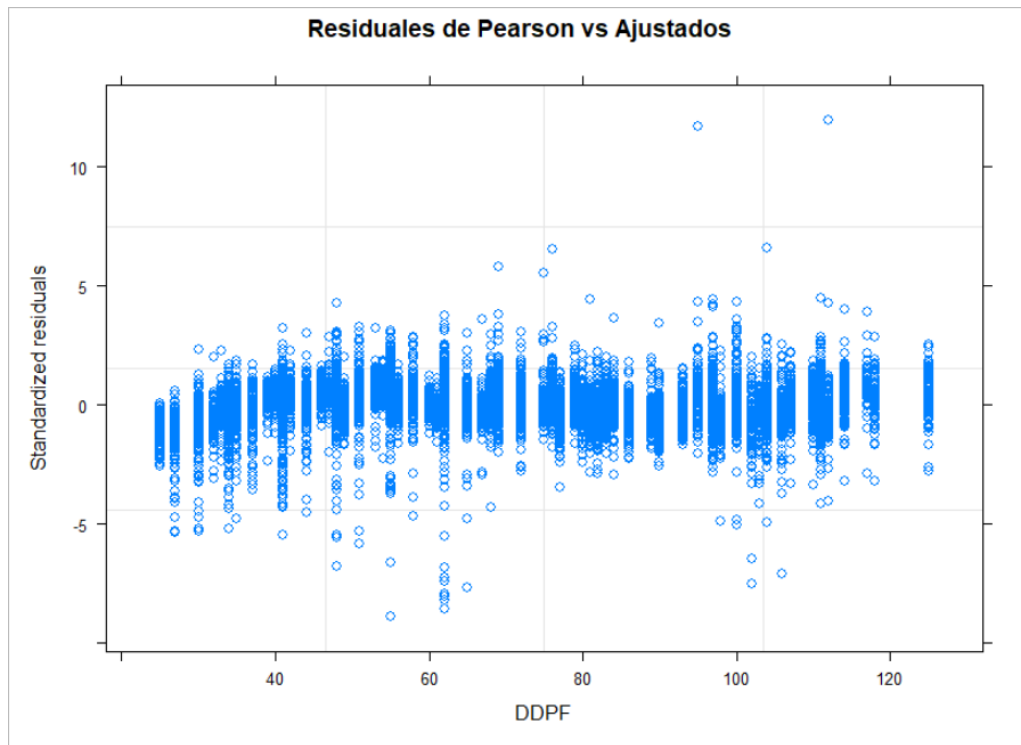
134. ZUDE, M. (2003). Comparison of índices and multivariate models to non-destructively predict the fruit chlorophyll by means of visible spectrometry in apple fruit. *Analytica Chimica Acta*, 481:119-126. doi.org/10.1016/S0003-2670(03)00070-9
135. ZHOU, R., DAMEROW, L., SUN, Y., BLANKE, M. M. (2012). Using colour features of cv. “Gala” apple fruits in an orchard in image processing to predict yield. *Precision Agriculture*, 13(5), 568–580. DOI:10.1007/s11119-012-9269-2

7. ANEXO

7.1. Gráficos de residuales para el modelo seleccionado en peras







7.2. Gráficos de residuales para el modelo seleccionado de manzanas:

