



Scientia Agropecuaria

Web page: <http://revistas.unitru.edu.pe/index.php/scientiaagrop>

Facultad de Ciencias
Agropecuarias

Universidad Nacional de
Trujillo

RESEARCH ARTICLE



Biomass estimation of a high Andean plant community with multispectral images acquired using UAV remote sensing and Multiple Linear Regression, Support Vector Machine and Random Forests models

Estimación de la biomasa de una comunidad vegetal altoandina utilizando imágenes multiespectrales adquiridas con sensores remotos UAV y modelos de Regresión Lineal Múltiple, Máquina de Vectores Soporte y Bosques Aleatorios

Andrés C. Estrada Zúñiga^{1*} ; Jim Cárdenas Rodríguez¹ ; Juan Víctor Bejar Saya² ; Javier Ñaupari Vásquez² 

¹ Universidad Nacional de San Antonio Abad del Cusco, Cusco, Av. De la Cultura 733, Cusco. Perú.

² Universidad Nacional Agraria La Molina, Lima Av. La Molina 15024. Lima. Perú.

* Corresponding author: andres.estrada@unsac.edu.pe (A. C. Estrada Zúñiga).

Received: 25 May 2022. Accepted: 30 September 2022. Published: 10 October 2022.

Abstract

Remote sensing with large-scale satellite images for precision studies in grasslands has spatial and spectral resolution limitations. Against this, using spectral signs and vegetation indices obtained with microsensors transported by unmanned aerial vehicles (UAV) constitutes a more accurate alternative for biomass estimation. In the fieldwork, images were acquired with microsensors, and fixed transects of 100 m were used where vegetation samples were collected. The photographs acquired with the UAV were processed in Pix 4D, Arc Gis, and algorithms elaborated in R programming language. The biomass estimation was carried out with Multiple Linear Regression, Vector Support Machine, and Random (Forest Random) models. The Random model showed a Kappa coefficient of 0.94 in the training set and 0.901 in the test set ($R^2 = 0.482$). The Random Forest model predicted 3 g/pixel of MV for Puna grass in the rainy season and 2 g/pixel for the dry season; the predicted biomass for the Tola bush was 15 g/pixel of MV for both seasons of the year. The estimation of biomass/hectare for the tolar plant community with its tolar shrub and Puna grass components was 6,535.88 kg/ha for the rainy season and 6,588.81 kg/ha for the dry season. The difference between the biomass estimated in the field and the biomass estimated with Random Forest was 5.48% for the rainy season and 9.63% for the dry season.

Keywords: Aerial biomass; machine learning; Unmanned Aerial Vehicle (UAV); Tolar; Support Vector Machine; Random Forest.

Resumen

La teledetección con imágenes satelitales de gran escala para estudios de precisión en pastizales presenta limitaciones en su resolución espacial y espectral; frente a ello el uso de signos espectrales e índices de vegetación obtenidos con microsensores transportados por vehículo aéreo no tripulado (VANT) constituyen una alternativa de mayor precisión para la estimación de biomasa. En el trabajo de campo, además de adquirir las imágenes con los microsensores, se utilizaron transectas fijas de 100 m donde se recolectaron muestras de vegetación. Las fotografías adquiridas con el VANT se procesaron en Pix 4D, Arc Gis y algoritmos elaborados en el lenguaje de programación R. La estimación de biomasa se realizó con los modelos de Regresión Lineal Múltiple, Máquina de Soporte Vectorial y Random (Bosques Aleatorios). El modelo Random mostró un coeficiente Kappa de 0,94 en el set entrenamiento y de 0,901 en el set de prueba ($R^2 = 0,482$). El modelo Random Forest predijo 3 g/píxel de MV para césped de puna en la época de lluvia y 2 g/píxel para la época seca; la biomasa predicha para el arbusto de Tola fue de 15 g/píxel de MV para ambas épocas del año. La estimación de biomasa/hectárea para la comunidad vegetal tolar con sus componentes arbusto de tolar y césped de puna fue de 6535,88 kg/ha para la época de lluvia y de 6588,81 kg/ha para la época seca. La diferencia entre las estimaciones de biomasa en campo y la estimación con Random Forest fue de 5,48% para época de lluvia y de 9,63% para época de estiaje.

Palabras clave: Biomasa aérea; aprendizaje automático; Vehículo Aéreo No Tripulado (VANT); tolar; Máquina de Vectores Soporte (MVS); Random Forest.

DOI: <https://dx.doi.org/10.17268/sci.agropecu.2022.027>

Cite this article:

Estrada Zúñiga, A. C., Cárdenas Rodríguez, J., Bejar Saya, J. V., & Ñaupari Vásquez, J. (2022). Estimación de la biomasa de una comunidad vegetal altoandina utilizando imágenes multiespectrales adquiridas con sensores remotos UAV y modelos de Regresión Lineal Múltiple, Máquina de Vectores Soporte y Bosques Aleatorios. *Scientia Agropecuaria*, 13(3), 301-310.

1. Introducción

En los Andes del sur del Perú nace la franja de la puna seca y se extiende desde los 15° latitud sur y comprende desde las tierras altas del departamento de Arequipa (Estrada et al., 2018). Hasta Tucumán en Argentina a 30° latitud sur (Tapia & Flores, 1984). El territorio peruano es considerado andino amazónico y la presencia de la cordillera de los Andes le otorgan características ecológicas muy particulares con diversidad de zonas de vida (Pulgar, 1996; Holdrige, 1987).

En la zona de producción homogénea de puna seca se encuentran comunidades vegetales de tola (*Parastrephia* sp), pajonales y césped de puna (Estrada & Ñaupari, 2021), que proveen diversos servicios ecosistémicos, entre ellos el de proveer alimentos, combustible, regulación del ciclo hidrológico y representa el 65% de los aproximadamente 2 millones de hectáreas de pastos naturales de la Reserva Nacional de Salinas y Aguada Blanca de la Región Arequipa (Flores et al., 2014).

Los SIG y la teledetección son utilizados en el estudio de coberturas vegetales y aplicados también en la gestión de los recursos naturales, (Acebedo, 2011; Berrio et al., 2015; Sharma et al., 2022). Sin embargo, las imágenes satelitales de gran escala presentan limitaciones en sus resoluciones espacial espectral y temporal (Buitrón & Calisaya, 2012; Azcoiti, 2016). Por ello, la incorporación de los sistemas aéreos no tripulados (UAS = *Unmanned Aerial System*) y los microsensores (Pilho et al., 2019; Ezenne et al., 2019) se posicionan como una nueva alternativa para obtener imágenes de alta resolución, superando limitaciones presentadas por las imágenes satelitales (D'Oleira et al., 2012; Norasma et al., 2021; Elshikha et al., 2022).

Los sensores multispectrales pasivos transportados por Drone pueden captar la radiación fotosintéticamente activa a partir del grado de verdor de la vegetación (Flores et al., 2021). Sin embargo, la capacidad de estimar la productividad neta y la producción de biomasa aún es limitada (Sharma et al., 2022) por falta desarrollo de modelos adecuados (Librán et al., 2020). Considerando el

potencial de los microsensores en la agricultura, el estudio planteo como objetivo estimar la producción de biomasa aérea de la comunidad vegetal tola de puna seca a partir de signos espectrales e índices de vegetación con imágenes multispectrales adquiridas con sensores transportados por sistemas aéreos no tripulados (UAS) (Kharuf et al. 2018; Zhu et al., 2018; Sharma et al., 2022).

2. Materiales y métodos

Ubicación de ámbito de estudio

El estudio se realizó en el fundo Quena del distrito de Callalli, Provincia de Caylloma Departamento Arequipa, a una altitud desde los 3800 y corresponde a la franja altiplánica de la provincia de Caylloma, región Arequipa, Esta zona semiárida se clasificada como zona homogénea de producción de puna seca, que presenta una precipitación pluvial media de 350 mm/año y una temperatura de 12,2 °C media máxima y (-10,08 °C) media mínima.

Proceso metodológico del estudio

Para la recopilación de muestras en campo y fotografías con los microsensores se colocaron puntos de control en tierra, que son referentes para realizar las correcciones geométricas en el proceso de elaboración de las ortofotografías (Sharma et al., 2022; D'Oleira, 2012). En las zonas de vuelo se colocaron cuatro puntos de control en tierra. Estos puntos eran cuadrantes de triplay de 80 x 60 cm, pintados de color azul, rojo y blanco, fácilmente identificados por el sensor y reconocidos durante el tratamiento de las imágenes (Figura 1). Cada uno de los puntos de control cuentan con datos georreferenciados que fueron tomados con GPS diferencial Spectra Precision Promark 120.

Se tomaron 100 muestras de tola (*Parastrephia* sp.) a lo largo de las transectas fijas para la estimación de biomasa y adyacente a estas muestras con un cuadrante de 0,25 x 0,25 cm se tomaron 100 muestras de césped a lo largo de la transecta (Figura 1).

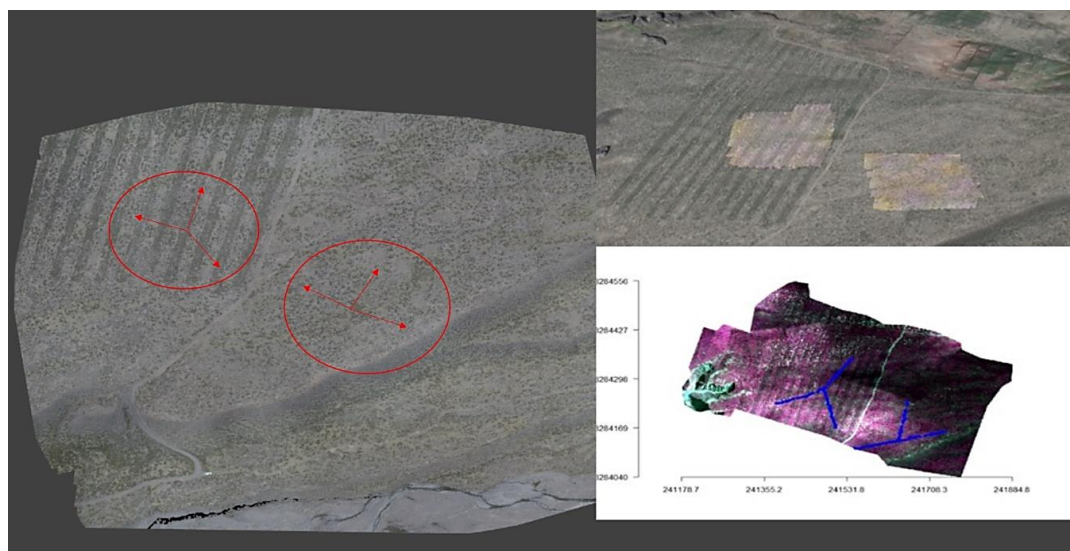


Figura 1. Puntos de control en tierra y transectas para muestreo de biomasa.

Para la recopilación de fotografías multispectrales se utilizaron microsensores transportados por UAV el que estuvo compuesto por un Drone Spyder DJI Multirrotor, que transportaba una cámara multispectral de cinco bandas, azul, verde, rojo, borde rojo e infrarrojo cercano (NIR) de marca Micasense con resolución a 100 m de altura de 12 Mpx con una resolución geométrica de 6,35 cm/pixel y con autonomía de vuelo de 20 minutos por batería a una altura de 4500 msnm. El segundo Drone fue el eBee de ala fija: con una cámara Parrot Sequoia de 16 Mpx. con una resolución de 13 cm/pixel a 120 m de altura de vuelo, y una autonomía de vuelo de 40 minutos por batería a 4500 msnm.

Etapas de gabinete, procesamiento de datos y estimación de biomasa

En la etapa de gabinete se elaboró una base de datos con información de: peso materia verde y peso en materia seca de las muestras de tola y césped recolectadas en campo (Chávez & Bolaños, 2021; Elshikha et al., 2022; Robb et al., 2020). Al mismo tiempo se procesaron las fotografías adquiridas con los UAV en la plataforma de fotogrametría Pix-4D, obteniendo como producto las ortofotografías multispectrales.

Para la estimación de biomasa a partir de imágenes multispectrales se elaboraron algoritmos en el lenguaje de programación R, utilizando las librerías "caret" (Max et al., 2021), "Performance Analytics" (Peterson et al., 2020), "corrplot" (Wei & Simko, 2017), "Random Forest" (Liaw & Wiener, 2022), "rgdal" (Bivand et al., 2017), "raster" (Hijmans et al., 2020), "sp" (Pebesma & Bivand, 2013), que

utilizaron como insumo ortomosaicos multispectrales. El algoritmo elaborado utilizando ortofotografías multispectrales calcula diferentes índices de vegetación (NDVI, SAVI, EVI) y genera mapas de cobertura (Elshikha et al., 2022) utilizando combinaciones de bandas espectrales e índices de vegetación a través del algoritmo Random Forest (Sharma et al., 2022).

Previo a la estimación de biomasa se evaluaron individualmente los modelos: Modelo de Regresión Lineal Múltiple (Amat, 2016), Model Support Vector Machines (SVM) (Gregory & Mira, 2016); Modelo Random Forest – RF (Bosques aleatorios) para predicción de biomasa, estimación de biomasa y generación de mapas de cobertura (Gregory & Mira, 2016). La evaluación se realizó a través de un set de prueba, tomando en cuenta los parámetros de regresión R^2 , RMSE y MAE

3. Resultados y discusión

3.1. Índices de vegetación estimados para la comunidad vegetal tolar de puna seca

Los valores de NDVI en la época de lluvia para la comunidad vegetal tolar fueron 0,0 a 0,8 y mientras que los valores de 0,0 a 1,00 alcanzaron los índices GNDVI, SAVI, EVI y MCARI. Los valores del NDVI de esta comunidad vegetal en la época seca fue de -0,4 a 0,4. Los índices GNDVI, NDRE, SAVI y EVI alcanzaron rangos de valor más amplios que van de -0,6 a 0,6. Los índices MCARI y WDRVI muestran rangos negativos desde -1,0 a 0,0, estos valores negativos encontrados corresponden a las coberturas de agua y suelo desnudo (Figura 2 a y b).

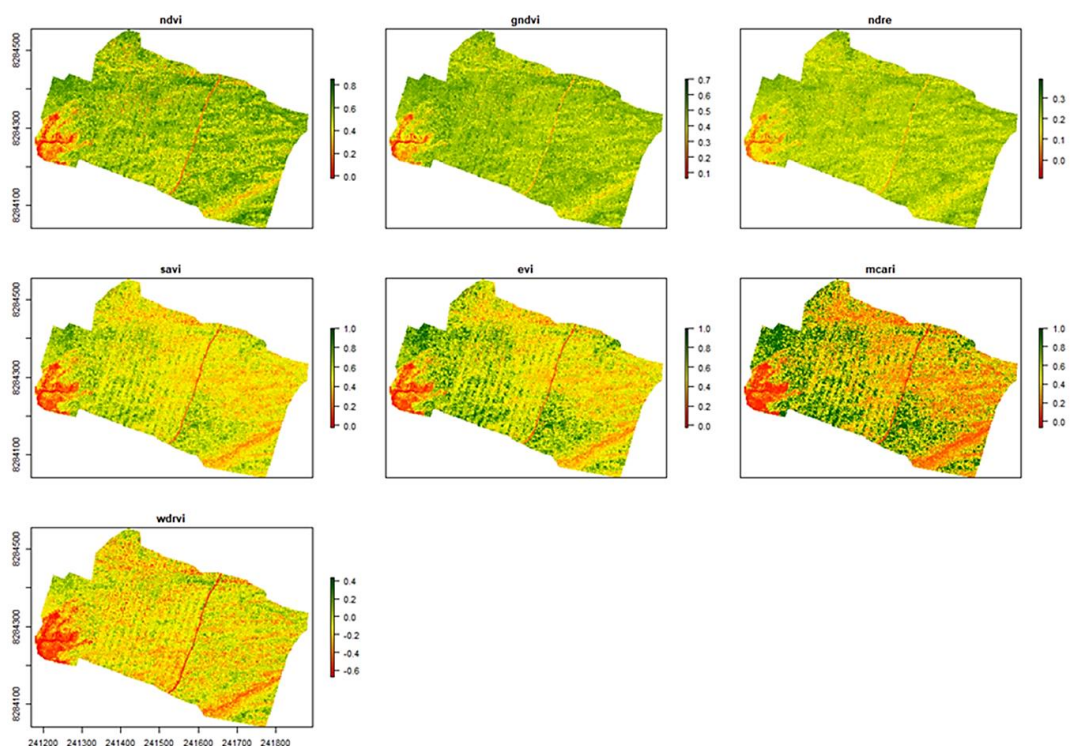


Figura 2.a Mapas de índices de vegetación en época seca y época de lluvia estimados con el "Modelo Césped de Luna".

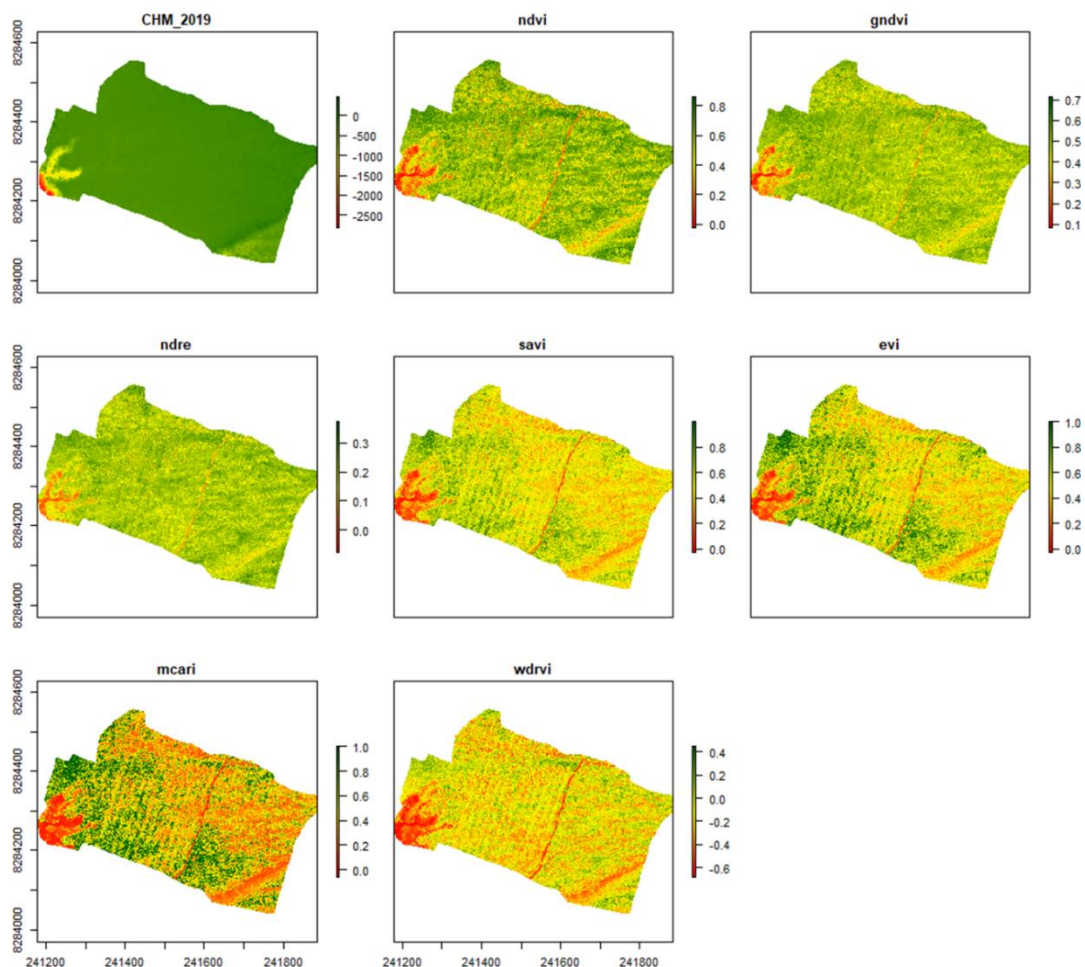


Figura 2.b Mapas de índices de vegetación en época seca y época de lluvia estimados con el "Modelo Tola".

Tabla 1

Evaluación de modelos para predicción de biomasa

Modelo	Época de Lluvia			Época de Seca		
	Precisión		Kappa Test	Precisión		Kappa Test
	Entrenamiento	Prueba		Entrenamiento	Prueba	
Elastic Net	0,9111	0,9196	0,8946	0,7942	0,8067	0,7680
Machine Learning para clasificación (SVM)	0,9001	0,9093	0,8809	0,7797	0,7888	0,7465
Random Forest.	0,9450	0,9409	0,9071	0,8413	0,8333	0,791

3.2. Evaluación de modelos para realizar estimación de biomasa y el coeficiente de correlación

Realizada la evaluación de los modelos *Random forest*, *Machine Learning (SVM)* y *Elastic net* para estimación de biomasa. Random Forest presentó mejor precisión siendo el de entrenamiento de 0,9450 y de verificación o prueba de 0,9409 y el coeficiente de Kappa fue de 0,9071 para muestras adquiridas en la época de lluvia. Para los resultados obtenidos se utilizaron Random forest como estimador de biomasa (Tabla 1)

Para el final de la época seca, *Random Forest* presenta valores de precisión de entrenamiento (0,8413) valor de precisión de prueba (0,8333), e índice de Kappa (0,791 superiores a los modelos *Elastic net* y *SVM*. Considerando estos resultados presentados para estimar la producción

de biomasa en la comunidad vegetal tolar se utilizó *Random Forest*.

La media decreciente en *Gini* (MDG = *Mean Decrease Gini*) para la comunidad vegetal tolar de puna seca, establece que índices de vegetación son de alta importancia en la estimación de biomasa. La priorización de variables determinadas a través de la MDG, establece que, para la estimación de biomasa, debe usar diferentes índices de vegetación según la época del año, quedando establecido que para estimar biomasa en imágenes adquiridas al final de la época de lluvia es mejor usar los índices de vegetación WDRVI, NDVI y GNDVI, y se puede usar con restricciones los índices de vegetación SAVI, EVI, mientras que para la época seca los índices que muestran mejor performance para estimación de biomasa son EVI, RVI y WRDVI.

Las variables priorizadas en el párrafo anterior se sustentan en los resultados del MDG en muestras adquiridas al final de la época de lluvia. Quedando establecido que los índices de vegetación más importantes en la estimación de biomasa son WDRVI, NDVI y GNDVI, y se encuentran en más de 2000 nodos de árboles de decisión aleatoria. Los índices SAVI y EVI son índices de menor importancia para la estimación de biomasa y se encuentran entre 1000 a 1500 nodos. Para muestras adquiridas al final de la época de lluvia a través de la MDG, se determinó como la variable de mayor prioridad el índice de vegetación EVI, el que se encuentra en más de 4500 nodos, seguido por WRDVI y RVI, que están entre 3000 a 4000 nodos de árboles de decisión aleatoria.

El estudio de **Hernández et al. (2019)** para estimar el rendimiento de trigo con imágenes adquiridas con Drone usando índices de vegetación DRE y NDVI, recomiendan el uso de imágenes o mapas de los índices NDRE y NDVI como referencia para diferenciar zonas de mayores rendimientos de biomasa aérea, considerando que el uso de imágenes NDVI muestran mejor performance en la estimación de biomasa en zonas con menor cobertura vegetal. El NDRE genera un mapa de mayor precisión en zonas bajo riego y en coberturas vegetales con mayor índice de área foliar. El estudio realizado en la comunidad vegetal tolar utilizando imágenes multispectrales que los índices de vegetación de GNDVI, NDVI, WRDVI tienen buena performance en la estimación de biomasa aérea, y un segundo grupo que son el SAVI o índice corregido por suelo y EVI o índice corregido por vegetación (EVI). Para predecir producción de biomasa aérea para la época

seca, se identificó como variables de alta importancia los índices MCARI, EVI y SAVI y las bandas RED y GREEN, todo ello apoyados por el modelo Random Forest, que predice la producción de biomasa más cercana a la realidad (**Ishida et al., 2017**).

El coeficiente de correlación de Pearson, aplicado a muestras de MV adquiridas en la época de lluvia, define una asociación del 3% ($p = 0,03$) entre la biomasa estimada con los datos de campo y la biomasa estimada a partir de los índices de vegetación NDRE, NDVI, GNDI, MACARI, EVI y SAVI. El índice WDRVI muestra una asociación negativa muy débil ($p=0,001$). Para el análisis a partir de MS, se encuentra una asociación positiva muy débil del 3 al 6%, con valores de ($p= 0,03$) a ($p=0,05$). El mismo coeficiente aplicado a muestras de MV tomadas al final de la época seca, estima una asociación negativa débil para los índices de vegetación GNDVI ($p=-0,23$), NDVI ($p=-0,22$), y NDRE, RVI, EVI, WRDVI, ($p = -0,22$ a $p=-0,12$), siendo el índice MCARI ($p=0,07$). el único caso de correlación positiva muy débil.

3.3. Predicción de biomasa utilizando el Modelo Random Forest

El modelo *Random Forest* para época de lluvia en la cobertura de césped de la comunidad vegetal tolar predice una producción de biomasa en un rango de 1 a 8 g/píxel de MV, siendo los píxeles más frecuentes los que predicen una producción de 1 y 2 g de biomasa/píxel. La predicción de biomasa en base a MS fue de 0,1 a 4 g/píxel, siendo los píxeles que predicen producen 0,5 a 1,5 g de biomasa los más frecuentes (**Figura 3**).

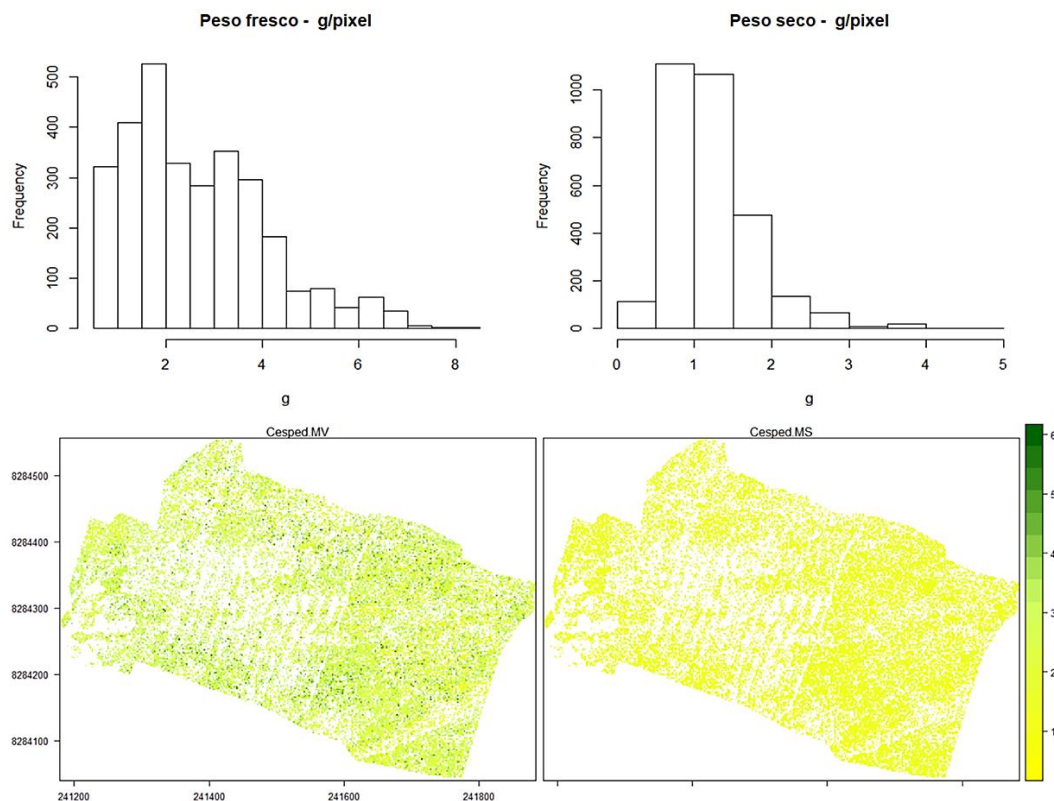


Figura 3. Estimación de biomasa con modelo RF para cobertura césped de puna.

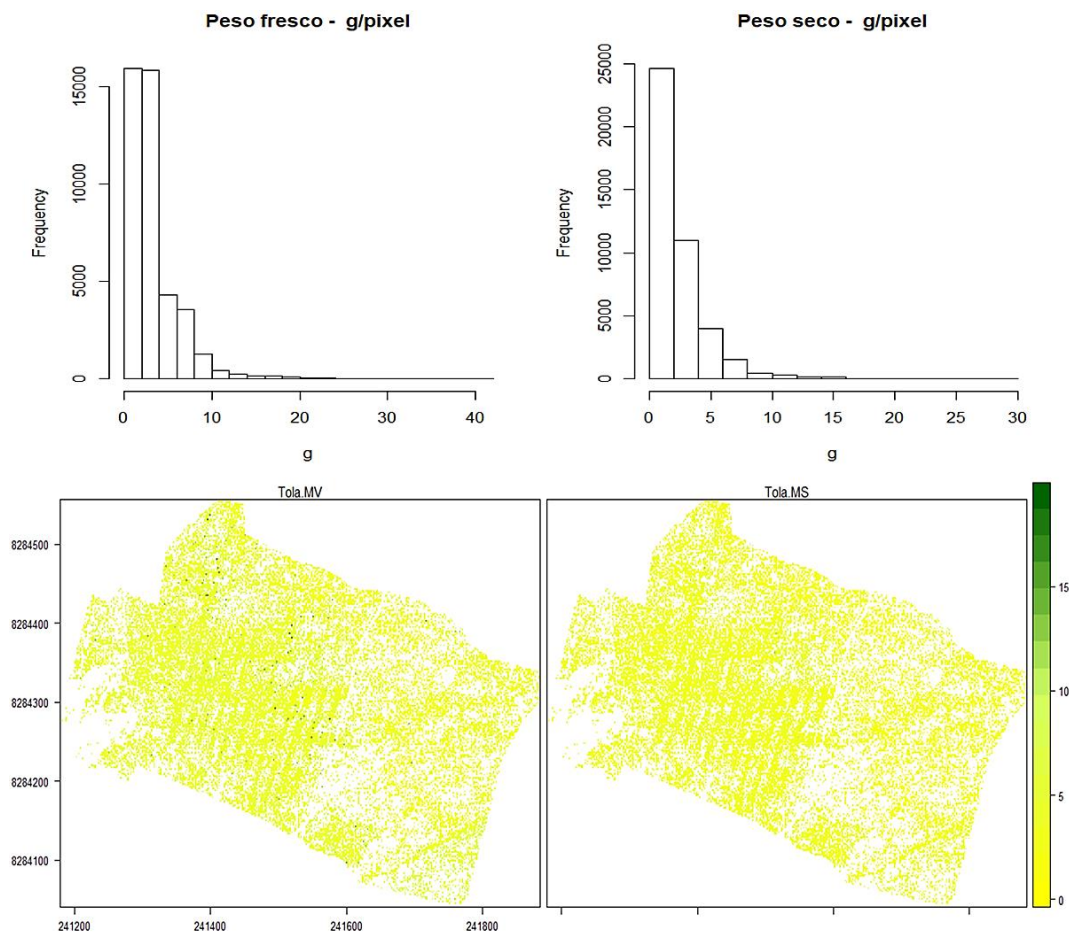


Figura 4. Estimación de biomasa con modelo RF para cobertura tolar.

El modelo aplicado con datos recopilados en época seca estimó una producción de 5 a 30 g/pixel de MV, siendo los pixeles con valores de producción de 10 y 11 g de biomasa los más frecuentes. Los resultados con datos de MS en la misma época predicen una producción de 1 a 8 g/pixel de biomasa, siendo los frecuentes los pixeles que estiman 2 a 4 g de biomasa. Los resultados encontrados durante el proceso de estimación de biomasa, para el césped del Fundo Quena, muestran poca variación entre época de lluvia y época seca (Figura 3).

Para las muestras adquiridas al final de la época de lluvia, el modelo *Random Forest* para cobertura tolar de la comunidad vegetal, predijo una producción de biomasa de 1 a 40 g/pixel, siendo los pixeles que estiman 2 a 3 g de biomasa los más frecuentes. El modelo para las muestras de la época de lluvia predijo de 1 a 60 g/pixel de producciones de biomasa siendo los pixeles que producen 5 g/pixel de biomasa en materia verde los más frecuentes. La predicción de biomasa en base a materia seca para ambas épocas presentó mayor frecuencia de pixeles que predicen 4 a 5 g/pixel de biomasa (Figura 4). La evaluación de los modelos como predictores de biomasa a partir de índices de vegetación para la época de lluvia y época seca, muestran un coeficiente de determinación ($R^2 = 0,484$) para el modelo *Random Forest*, siendo este modelo el que tiene la mayor posibilidad de predecir la producción biomasa más

cercana a la realidad: El modelo de Regresión Lineal ($R^2 = 0,0427$ y $R^2 = 0,146$) y el modelo SVM ($R^2 = 0,0016$ y $R^2 = 0,174$) presentan los más bajos coeficientes de determinación, siendo su posibilidad de estimación de biomasa muy imprecisa. Ante la evidencia, se concluye que dichos modelos no estarían en condiciones de estimar la producción de biomasa en ninguna época. Estos valores son muy cercanos a los coeficientes de determinación ($R^2 = 0,31$ y $R^2 = 0,22$) reportados por Paredes (2018) en el caso de los pajonales de la sierra central y los reportes de Kharuf et al. (2018) que obtuvo coeficientes de determinación de ($R^2 = 0,21$ a $R^2 = 0,39$). Sharma et al. (2022) para arbustos determino coeficientes de determinación de $R^2 = 0,20 - 0,25$ que son bajo y parecidos a los encontrados en el estudio.

Los coeficientes de determinación son bajos, por lo que su posibilidad para predecir biomasa de la comunidad vegetal tolar no será el óptimo. Considerando que el estudio no puede controlar variables de la naturaleza, se puede explicar que el modelo *Random Forest* ha estimado predicciones de producción de biomasa más cercanas a lo observado en campo para las dos épocas de estudio. Los bajos coeficientes de determinación y correlación encontrados en el presente estudio y los reportados por Bethany et al. (2019), Botello et al. (2019) y Gruner et al. (2018) atribuyen a: las diferencias propias de los cultivos y las comunidades vegetales para cada una de las zonas de

estudio, a la característica heterogénea de la vegetación contenida en un pixel, al número de observaciones por pixel para el método de estimación de biomasa utilizado y al ruido de la imagen según la época adquirida.

En el gráfico de dispersión (Figura 5) Random forest muestra distribución adecuada y predice mejor la producción de biomasa para la cobertura de césped en base a MS. Los valores predichos y observados están concentrados entre 1 y 3 g/pixel de biomasa en base a MV para el final de la época de lluvia, y de 5 a 15 g/pixel de biomasa para el final de la época seca. Para la cobertura tola, los gráficos muestran 3 a 10 g MV/pixel de biomasa.

3.4. Estimación de producción de biomasa de la comunidad vegetal tolar del área de estudio

El modelo *Random Forest*, al clasificar la cobertura estimó el área de la comunidad vegetal tolar del fundo Quena en 20 ha, así como determino tres clases de cobertura: tola, césped y suelo. La cobertura tola ocupa el 44,61% (9,190 ha.) para la época de lluvia y 49,78% (9,956 ha.) para la época seca, siendo la especie dominante de la comunidad. La clase césped de puna tiene un comportamiento variable: en la época de lluvia ocupa el

47,82% (9,85 ha.) y en la época de seca ocupa el 27,86% (5,57 ha) de la comunidad, mostrando una diferencia de 19,6% (4,28 ha) entre épocas. La clase suelo desnudo también presenta un comportamiento variable entre épocas de muestreo: en la época de lluvia ocupa el 7,57% del área y en la época seca el 22,6%.

El modelo *Random Forest* a partir de imágenes multispectrales e índices de vegetación para la clase césped de puna seca, estimó una producción de biomasa de 3282,50 kg MV/ha para la época de lluvia y de 2630,11 kg MV/ha de biomasa para la época seca y la predicción de biomasa en MS fue de 864,86 kg/ha para la época de lluvia y 1245,73 kg/ha para la época seca (Tabla 4). Mientras que la producción estimada con datos de muestras colectadas en campo fue de 1326,00 kg MV/ha para época de lluvia y 975 kg MV /ha para época seca. La biomasa en MS se estimó en 450 kg/ha para época de lluvia y 351,00 kg/ha para época seca (Tabla 4). Las diferencias de las estimaciones de producción de biomasa para el componente césped de puna en base a MS y MV son muy altas.

Tabla 2

Evaluación de los modelos para predicción de biomasa

Modelo	Época Lluvia			Época Seca		
	RMSE	RSq	MAE	RMSE	RSq	MAE
Linear Regression	2,768	0,043	1,859	3,490	0,146	2,470
Support Vector Machine (SVM)	1,432	0,002	1,131	3,325	0,174	2,904
Random Forest	1,994	0,484	1,148	3,623	0,418	2,817

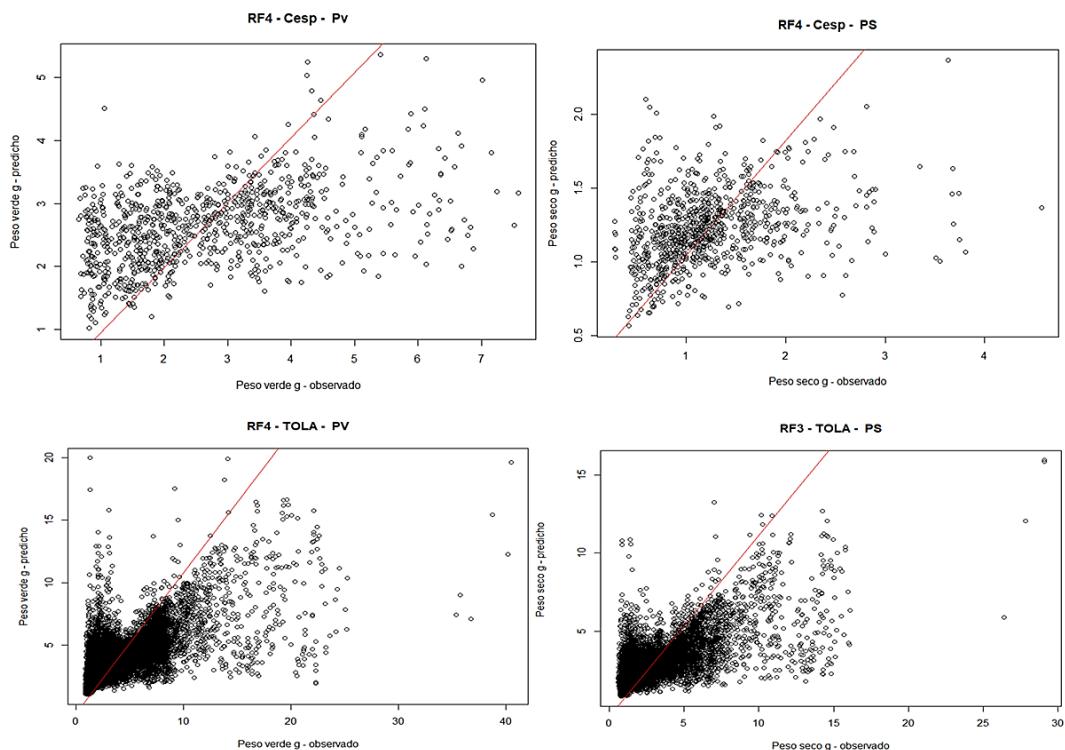


Figura 5. Comparación entre predicción RF y observado en campo.

La producción de biomasa de la clase tola estimada con datos de campo para época de lluvia fue de 15875,00 kg/ha., mientras que *Random Forest* para la misma época estimó una producción de biomasa 20914,42 kg MV/ha. Para la época seca, *Random Forest* estimó la producción de 13535 kg MV/ha, y la estimación con datos de campo fue de 12850 kg MV/ha, mostrando una diferencia del 5,33%. Las diferencias de estimación en base a materia seca fueron menores para muestras de ambas épocas. Estos resultados muestran que el modelo *Random Forest* estima resultados más cercanos a la producción de biomasa en campo.

Las predicciones de biomasa en pastizales están influenciadas de manera directa con la época de muestreo. **Santa et al. (2013)**, a partir del NDVI, estimó 189,6 g/m² de biomasa MV para el inicio de lluvias mes de diciembre y para el mes de marzo a la producción de biomasa se incrementó a 105,74 y 115,22 g/m². **Paredes (2018)**, en la sierra central del Perú, obtuvo valores de biomasa disponible entre 1 558 y 3 559 kg de MS/ha. Para la comunidad vegetal tolar, utilizando el algoritmo *Random Forest*, con índices de vegetación NDVI se estimó una producción de biomasa de 10 889,82 kg/ha, que resulta mayor a la estimación realizada con datos de campo, que fue de 7 716,3 kg/ha de materia verde para la época de lluvia. Los resultados encontrados en la estimación de biomasa de la comunidad vegetal tolar de puna seca son diferentes a los de **Santa et al. (2013)** y superiores a los de **Paredes (2018)**, situación que se explica por la diferencia de la composición florística de cada comunidad estudiada.

El moldeo *Random Forest* utilizando imágenes multiespectrales de la comunidad vegetal tolar de puna seca en época de lluvia (coberturas de césped, tola y suelo desnudo) estimó una producción de biomasa de 10889,82 kg MV /ha y la estimación de producción de biomasa con datos de campo fue de 7716,3 kg/ha, encontrando una diferencia entre los dos métodos de estimación del 41,13% en MV y 5,48% en MS para la misma época. La estimación de biomasa para la época seca fue de 7470,57 kg MV/ha, y 6668,37 kg MS/ha de biomasa para las muestras de campo, encontrando una diferencia de 12,03% para MV entre métodos de estimación de biomasa y para MS 9,63% de diferencia entre los dos métodos.

Utilizando el algoritmo *Random Forest*, con imágenes de alta resolución espacial y espectral, finas, se puede realizar la clasificación de cobertura para cada componente de la comunidad vegetal. El modelo también permite elaborar

coberturas por especies de pastos. Este resultado coincide con los reportes de **Bethany et al. (2019)**, quienes realizaron estudios de pastizales con sensores hiperspectrales de alta resolución en plataformas UAS, concluyendo que el uso de sensores de alta resolución y el modelo de predicción *Random forest* muy útil para la identificación y clasificación de pastizales, así como para la estimación de la biomasa.

La percepción remota es una herramienta valiosa de apoyo para las estimaciones de biomasa.

La correlación entre los valores de NDVI estimados y el estado de salud de las plantas es alta, encontrándose valores del NDVI más altos (superior 0,49) en las plantas más saludables y los valores bajos de 0,21 a 0,39, se encuentran en las parcelas más afectadas por el estrés nutricional de las plantas. Finalmente, los valores casi nulos de NDVI, por debajo de 0,21, se hallan en los sectores de vegetación pobre o suelo.

La clasificación comunidades vegetales y la estimación de biomasa puede realizarse con ortofotografías NDVI obtenidas con sensores de alta resolución y que estén en el rango espectral de 600 a 875 nm que cubran la banda roja e infrarrojo cercano del espectro. Sin embargo, la tarea de clasificación de coberturas resulta potenciada con el uso de imágenes adquiridas en el espectro visible (**Bethany et al., 2019; Elshikha et al., 2022**). Estos resultados muestran concordancia con los resultados del presente estudio, toda vez que las imágenes adquiridas con sensores multiespectrales de alta resolución y procesadas en el modelo *Random forest*, facilitaron la clasificación de coberturas y predijeron la producción de biomasa por pixel cuyos valores estuvieron cercanos a los estimados en el trabajo de campo.

En el Perú, un estudio de pastizales altoandinos a partir de signos espectrales e índices de vegetación con un sensor MODIS TERRA (**Paredes, 2018**), muestra que las relaciones entre las variables NDVI-biomasa y EVI-biomasa, tuvieron coeficientes de determinación de $R^2 = 0,31$ y $R^2 = 0,22$ respectivamente. En Córdoba, Argentina, un estudio que utilizó imágenes LANSAT 5 TM con datos de biomasa en campo y NDVI obtenidos por sensores remotos, reportó el coeficiente de determinación más alto para estudios de biomasa de pastizal, que es de ($R^2 = 0,50$) (**Santa et al., 2013; Zorogastua, 2012**). Estos estudios recomendaron el uso de imágenes satelitales, los coeficientes de determinación encontrados en el estudio están en el mismo rango de los encontrados ($R^2: 0,3 - 0,9$) por **Elshikha et al. (2022)** para arbustos.

Tabla 3

Comparación de producción de biomasa con *Random Forest* y datos de campo

Clase		Época de Lluvia				Época Seca			
		Estimación Biomasa		Diferencia (RF CAMPO)		Estimación Biomasa		Diferencia (RF CAMPO)	
		Campo (kg/ha)	Random Forest (kg/ha)	(kg/ha)	%	Campo (kg/ha)	Random Forest (kg/ha)	kg/ha	%
Materia Verde	Césped	1326,00	3282,50	1956,50	147,55	975,00	2630,11	1655,11	169,75
	Tola	15875,00	20914,42	5039,42	31,74	12850,00	13535,20	685,20	5,33
	Comunidad Vegetal	7716,13	10889,82	3173,68	41,13	6668,37	7470,57	802,20	12,03
Materia Seca	Césped	450,84	864,86	414,02	91,83	351,00	1245,73	894,73	254,91
	Tola	13406,25	13723,63	317,38	2,37	11877,25	12538,68	661,43	5,57
	Comunidad Vegetal	6196,32	6535,88	339,55	5,48	6010,28	6588,81	578,53	9,63

4. Conclusiones

Es factible predecir producción biomasa utilizando imágenes multiespectrales adquiridas con microsensores, Las estimaciones pueden realizarse con diferencias que van del 5,48% para época de lluvia y de 9,63% para época de estiaje.

Los mejores meses de muestreo para estimar la producción de biomasa usando índices de vegetación en tolares son los meses de marzo y octubre. por que permiten realizar predicciones y estimaciones de producción de biomasa más cercanas a las estimaciones de producción en campo.

Random forest es el algoritmo que mejor desempeño muestra en la estimación de biomasa para matorrales y la comunidad vegetal tolar.

Agradecimientos

A la UNSAAC y Laboratorios de ciencia animal y cambio climático de la escuela profesional de Zootecnia de la Facultad de Ciencias Agrarias; al laboratorio de evaluación y utilización de pastizales de la Facultad de Zootecnia de la UNALM, al equipo docente del Doctorado en Ciencia animal.

ORCID

A. C. Estrada Zúñiga  <https://orcid.org/0000-0002-1588-6399>

J. Cárdenas Rodríguez  <https://orcid.org/0000-0002-8775-9014>

J. V. Bejar Saya  <https://orcid.org/0000-0002-0350-3557>

J. A. Ñaupari Vásquez  <https://orcid.org/0000-0002-1554-2633>

Referencias bibliográficas

- Acebedo, R. (2011). Sistemas de teledetección activos y pasivos embarcados en sistemas aéreos no tripulados para la monitorización de la tierra Barcelona España. 15pp.
- Amat, J. (2016). Introducción a la Regresión Múltiple. https://www.cienciadatos.net/documentos/25_regresion_li_neal_multiple
- Azcoiti, J. (2016). Evaluación de aplicaciones para el tratamiento de imágenes mediante correlación Automática. Tesis de Maestría Universidad Pública de Navarra.
- Berrio, M., Mosquera, T., & Álzate, D. (2015). Uso de drones para el análisis de imágenes multiespectrales en agricultura de pre-utilización de imágenes multiespectrales obtenidas por vehículos aéreos no tripulados para la determinación de zonas intra-lote con mayor rendimiento y calidad en el trigo (*Triticum aestivum*), Universidad Nacional de Villamaría Argentina.
- Bethany, M., Lucier, A., & Aryal, J. (2019). Classification of Lowland Native Grassland Communities Using Hyperspectral Unmanned Aircraft System (UAS) Imagery in the Tasmanian Midlands. *Drones*, 3(1), 5.
- Botello, A., Reyes, F., Chavez, L., Galvis, A., & Hidalgo, R. (2019). Colegio de Postgraduados, Campus Montecillo. Posgrado en Hidro ciencias México- Estado de México, México.
- Bivand, R., Keitt, T., & Rowlingson, B. (2017). *rgdal: Bindings for the "Geospatial" Data Abstraction Library* (1.2-16). <https://r-forge.r-project.org/projects/rgdal/>
- Buitrón, C., & Callisaya, J. (2012). Estudio espacial multitemporal de variaciones en superficie observada a través de imágenes satelitales LANDSAT en una región del Parque Nacional Sajama Bolivia, DIACONIA. La Paz Bolivia.
- Chávez, E., Paz, F., & Bolaños, M. (2017). Estimación de biomasa y cobertura aérea usando radiometría e imágenes digitales a nivel de campo en pastizales y matorrales. *Terra Latinoamericana*, 35(3), 247-257.
- D'Oleira, S., Marzoff, I., Klaus, D., & Ries, J. (2012). Unmanned aerial Veicle (UAV) for Monitoring soil erosion in Morocco. *Remote Sens*, 4(11), 3390-3416.
- Elshikha, D., Hunsake, D., Waller, P., Thorp, K., Dierig, D., Wang, G., Cruz, M., Katterman, K., Bronson, G., Wall, W., & Thompson, A. (2022). Estimation of direct seeded guayule cover, crop coefficient, and yield using UAS based multispectral and RGB data. *Agricultural Water Management*, 265, 107540.
- Estrada, A., Cárdenas, J., Ñaupari, J., & Zapana, J. (2018). Capacidad de carga de pastos de puna húmeda en un contexto de cambio climático. *Revista de investigaciones altoandinas*, 20(3), 361-368.
- Estrada, A., & Ñaupari, J. (2021). Detección e identificación de comunidades vegetales altoandinas, Bofedal y Tolar de Puna Seca mediante ortofotografías RGB y NDVI en drones "Sistemas Aéreos no Tripulados". *Scientia Agropecuaria*, 12(3), 291-301.
- Ezenne, G., Jupp, L., Mantel, S., & Tanner, J. L. (2019). Current and potential capabilities of UAS for crop water productivity in precision agriculture. *Agricultural Water Management*, 218, 158-164.
- Pilho, F., Heldens, W., Kong, Z., & Lange, E. (2019). Drones: Innovate technology for use in precision pest management. *Journal of Economic Entomology*, 113(1), 1-25.
- Flores, D., Gonzáles, H., Lozano, R., Vazquez, J., & Hernandez, J. (2021). Automated Agave detection and counting using a Convolutional Neural Network and Unmanned Aerial Systems. *Drones*, 5(1), 4.
- Flores, E. R., Ñaupari, J. A., & Tacuna, R.E. (2014). La economía del cambio climático en el Perú: Ganadería altoandina en la economía del cambio climático en el Perú: Desarrollo, CEPAL.
- Gregory, D., & Mira, J. (2016). Estudios de simulación estocástica con Random Forest frente a árboles individuales, Universidad Politécnica de Madrid España.
- Grüner, E., Astor, T., & Wachendorf, M. (2019). Biomass Prediction of Heterogeneous Temperate Grasslands Using an SfM Approach Based on UAV Imaging. *Agronomy*, 9(2), 54.
- Hernández, L., Rudge, M., Bartolo, R., & Erskine, P. (2019). Identifying species and monitoring understorey from UAS derived data, a literature review and future directions. *Drones*, 3(1), 9.
- Hijmans, R. J., Etten, J. Van, Sumner, M., Cheng, J., Baston, D., Bevan, A., Bivand, R., Bussetto, L., Canty, M., Fasoli, B., Forrester, D., Gray, J., Greenberg, J. A., Hiemstra, P., Karney, C., & Mattiuzzi, M. (2020). *raster: Geographic Data Analysis and Modeling* (Version 3.4-5, p. 249). CRAN. <https://cran.r-project.org/package=raster>
- Holdridge, L. R. (1987). Ecología basada en zonas de vida, servicio editorial IICA, tercera edición, San José de Costa Rica. ISBN 92 9039 1316 pp. 2-26.
- Ishida, T., Kurihara, J., Viray, F., & Takashi, J. (2017). A novel approach for vegetation classification using UAV-based hyperspectral imaging. *Computers and Electronics in Agriculture*, 144, 80-85.
- Kharuf, S., Hernandez, L., Orosco, R., & Delgado, I. (2018). Análisis de imágenes multiespectrales adquiridas con vehículos aéreos no tripulados. *Ingeniería electrónica, automática y comunicaciones*, 39(2), 79-91.
- Liaw, A., & Wiener, M. (2022). *Breiman and Cutler's Random Forests for Classification and Regression* (Version 4.7-1.1, Issue 3, pp. 239-248).
- Librán, F., Klaus, F., Tscharntke, T., & Grass, I. (2020). Unmanned aerial Vehicles for biodiversity friendly agricultural landscape a systematic review. *Science of the total environment*, 732, 139204
- Max, K., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Ziem, A., Scrucca, L., & Hunt, T. (2021). caret: Classification and Regression Training. In *R package version 6.0-86* (Version 6.0-90). <https://cran.r-project.org/package=caret>
- Norasma, C., Fadzilah, M., Roslin, N., Zanariah, Z., Tarmidi, Z., & Candra, S. (2019). Unmanned aerial vehicle applications in

- agriculture, IOP Publishing South conference on engineering and technology. *IOP Conf. Ser.: Mater. Sci. Eng.*, 506, 012063.
- Paredes, M. (2018). Uso de índices de vegetación del sensor MODIS TERRA en la estimación de biomasa aérea en pajonales altoandinos, Universidad Nacional Agraria la Molina Lima Perú.
- Pebesma, E., & Bivand, R. (2013). *Classes and Methods for Spatial Data* (Version 1.5-0). <https://edzer.github.io/sp/>
- Peterson, B. G., Carl, P., Boudt, K., Bennett, R., Ulrich, J., Zivot, E., Cornilly, D., Hung, E., Lestel, M., Balkissoon, K., Wuertz, D., Christidis, A. A., Martin, R. D., Zhou, Z., & Shea, J. M. (2020). *PerformanceAnalytics: Econometric Tools for Performance and Risk Analysis* (Version 2.0.4., p. 240). <https://github.com/braverock/PerformanceAnalytics>
- Pulgar, J. (1996). Geografía del Perú, las ocho regiones naturales, Promoción editorial Inca S.A. décima edición, Lima Perú, ISBN 9972 40 027 3pp 15 al 23.
- Robb, C., Harday, A., Doonan, J., & Brook, J. (2020) Semi automated field plot segmentation from UAS imagery for experimental agriculture. *Frontiers in plant Science*, 11, 591886.
- Santa, V., Rosa, M., Mónaco, N., & Heguiabehere, A. (2013). Determination of relationship between data obtained from biomass derived from field and remote sensing NDVI along the Arroyo Chucul (PACIA Córdoba). *Revista de la Facultad de Agronomía UNLPam*, 22(S2), 157-162.
- Sharma, P., Leigh, L., Chang, J., Maimaitjiang, M., Caffé, M. Above (2022). Ground Biomass Estimation in Oats Using UAV Remote Sensing and Machine Learning. *Sensors*, 22, 601.
- Tapia, M., & Flores, J. (1984). Ecosistema de los andes del sur del Perú y su relación con los pastizales, Capítulo 2, in "Pastoreo y pastizales en los andes del sur del Perú", Programa colaborativo de apoyo a la investigación en rumiantes menores.
- Wei, T., & Simko, V. (2017). Corrplot: Visualización of a Correlation Matrix (Versión 0.84, p. 18). <https://cran.r-project.org/web/packages/corrplot/corrplot.pdf>.
- Zhu, F., Jiewei, L., Honghui, X., & Goodman, E. (2018). Automatic Tobacco plant detection in UAV images via deep Neural Networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(3), 876-887.