

MÁSTER UNIVERSITARIO EN INGENIERÍA GEOMÁTICA Y GEOINFORMACIÓN

TRABAJO DE FIN DE MASTER

ESTIMACIÓN DE LA HUMEDAD DEL COMBUSTIBLE VIVO USANDO IMÁGENES SENTINEL-2 EN PARCELAS FORESTALES DE LA COMUNIDAD VALENCIANA.

Autor: Karla Gabriela Freire Quintanilla (karfrqui@topo.upv.es)

Tutor: Ángel Antonio Balaguer Beser (abalague@mat.upv.es)

Cotutor: José Luis Soriano Sancho (jlsoriano@vaersa.org)

Centro: E.T.S.I. Geodésica, Cartográfica y Topográfica Universitat Politècnica de València

Febrero del 2019

AGRADECIMIENTOS

A Ángel Antonio Balaguer Beser mi tutor, por haberme brindado la oportunidad de recurrir a su capacidad y conocimiento en el proceso de este proyecto, pero no solo por eso, sino por la paciencia, apoyo y enseñanza permanentes a lo largo del desarrollo de este estudio, un gracias infinito, ya que su constancia y tiempo brindado ha hecho que este trabajo finalice obteniendo buenos resultados.

A José Luis Soriano Sancho mi cotutor, que fue nuestra fuente de información y conocimiento para el desarrollo del estudio, haciendo que nunca se detenga el proyecto y entregando todo su apoyo, buenas energías y consejos para que se vaya encaminando de la mejor manera el trabajo.

Al Jorge Recio por dedicarme parte de su tiempo y conocimiento en el ámbito de teledetección en cuanto a índices de vegetación extraídos de imágenes satelitales.

A Rosa Ana Quintanilla y Carlos Freire mis padres, por el apoyo incondicional, por ser los principales promotores de mis sueños, por ser mis pilares y darme con su ejemplo las ganas de superarme y seguir creciendo como profesional, porque a pesar de la distancia siempre han estado a mi lado al igual que mi hermana Anita.

A Sergio Andrade, gracias por estar presente desde el primer momento en que tome este camino, por ser un apoyo incondicional a lo largo del máster compartiéndome tus conocimientos, por luchar juntos para conseguir este nuestro sueño, gracias por tu infinito amor.

A Paola Hernández, gracias amiga querida porque hiciste que este proyecto, el master, y la estadía aquí sea más llevadera. ¡Lo logramos!

Dedicatoria

A mis padres, por su interés y apoyo incondicional en mis estudios entregándome así la mejor herencia, por enseñarme que todo es posible sin importar el tiempo y el espacio, por demostrarme que nunca hay que bajar los brazos ni dejarse vencer, por conservar una familia unida y luchadora saliendo siempre adelante buscando el bienestar de mi hermana y el mío, por su amor incondicional y sobre todo por ser los mejores. Madre querida a ti por tu esfuerzo impresionante y tu amor invaluable.

COMPROMISO

“El presente documento ha sido realizado completamente por el firmante, no ha sido entregado como otro trabajo académico previo y todo el material tomado de otras fuentes ha sido convenientemente entrecuadrado y citado su origen en el texto, así como referenciado en la bibliografía”

“Los datos de humedad de los combustibles vivo forestales utilizados en este trabajo han sido proporcionados por la empresa VAERSA con la autorización de la Generalitat Valenciana, gracias a la colaboración mantenida para este fin entre los tutores de este trabajo y la Dirección General de Prevención de Incendios Forestales de la Generalitat Valenciana.”

RESUMEN

El contenido de agua que tiene la vegetación es un parámetro crítico que afecta la inflamabilidad de la vegetación y el comportamiento del fuego. Dicho parámetro se conoce como humedad del combustible vivo (HCV) y su valor depende de la especie forestal que se esté analizando. Se necesitan estimaciones actualizadas de HCV en las distintas especies forestales para la evaluación del riesgo de incendios forestales. Además, se tiene que estimar la proporción que las especies forestales guardan en cada parcela de muestreo. Sin embargo, el monitoreo detallado y constante de HCV en el campo es costoso y consume mucho tiempo.

En el presente trabajo se ha construido un nuevo método de estimación temporal de los valores de HCV en diferentes parcelas forestales de la Comunidad Valenciana. Para ello se utilizaron técnicas de regresión múltiple, usando como predictores diferentes índices de vegetación obtenidos desde enero de 2016 a través de imágenes del satélite Sentinel-2, en un conjunto de píxeles tomados en cada parcela forestal. Dichas parcelas están repartidas por las tres provincias de la Comunidad Valenciana, en las cuales se cuenta con datos mensuales de HCV para las diferentes especies forestales, teniendo en cuenta la proporción que guardan las mismas en la parcela forestal. Los datos de HCV han sido proporcionados por la Dirección de Prevención de Incendios Forestales de la Generalitat Valenciana a través de la empresa pública VAERSA.

Para mejorar la precisión de los modelos de regresión de HCV en cada parcela, calculados a partir de los índices de vegetación, se incluyeron variables meteorológicas, extraídas para diferentes periodos de tiempo. Se calculó la estimación del valor de HCV en cada parcela en otros valores de tiempo no usados en la muestra de entrenamiento del modelo de regresión, analizando la capacidad predictiva de los modelos en cada parcela forestal.

Palabras clave: Humedad del combustible vivo; regresión múltiple; Sentinel-2, parcelas forestales; Comunidad Valenciana.

ABSTRACT

The water content in the vegetation is a critical parameter that affects the inflammability of the vegetation as well as fire behaviour. This parameter is known as live fuel moisture (LFM) and its value depends on the forest species being analyzed. To evaluate the risk of forest fires, updated estimations of LFM are required. Moreover, it is necessary to estimate the proportion that forest species have in each sampling plot. However, detailed and constant monitoring are time and cost intensive.

This research generated a new method to estimate temporal values of LFM in various plots located in the Valencian Community. Accordingly, multiple regression techniques will be used. The predictors comprise various vegetation indexes generated through the analysis of satellite imagery of each plot taken by Sentinel-2 satellite since 2016. These parcels are distributed across three provinces of the Valencian Community, in which there is monthly data of LFM of the forest species, considering their proportion in the forest parcel. The LFM data has been obtained from the “Dirección de Prevención de Incendios Forestales de la Generalitat Valenciana” through the public enterprise VAERSA.

To improve the precision of LFM regression models in each parcel, calculated through vegetation indexes, the use of meteorological variables obtained at different time periods was tested. Also, the analysis of the predictive capacity of the models in each forest parcel was achieved by generating the LFM values for each parcel at times different from those employed to build the models.

Keywords: Live fuel moisture; multiple regression, Sentinel-2, forest parcels, Valencian Community.

INDICE DE IMÁGENES

Imagen 1. Ecosistemas forestales de la Comunidad Valenciana.	14
Imagen 2. Localización de los puntos de muestreo.	15
Imagen 3. Comportamiento temporal de la HCV en el punto de muestreo Fredes.	19
Imagen 4. Porción de FCC de las especies forestales en los puntos de muestreo.	20
Imagen 5. Formato físico de los archivos Sentinel-2.	22
Imagen 6. Resultado de la corrección atmosférica de las imágenes satelitales en Sen2cor.	23
Imagen 7. Mosaico de imágenes satelitales Sentinel-2 para la Comunidad Valenciana.	24
Imagen 8. Superficie para el cálculo de índices de vegetación.	27
Imagen 9. Tendencia de índices de vegetación en el periodo enero 2016-julio2018 en el punto de muestreo Buñol.	28
Imagen 10. Tendencia de índices de vegetación interpolados a la fecha de toma de muestra en el punto de muestreo Buñol.	28
Imagen 11. Relación entre el índice de vegetación NDVI y la HCV en el punto de muestreo Castielfabid.	29
Imagen 12. NDVI obtenidos para el año 2017 en la Comunidad Valenciana. Composiciones de máximo valor para los meses de diciembre (invierno) y agosto (verano)	30
Imagen 13. NDVI obtenidos para el año 2017 en la Comunidad Valenciana. Composiciones de máximo valor para los meses de marzo (primavera) y octubre (otoño)	31
Imagen 14. Estaciones meteorológicas.	33
Imagen 15. Esquema de metodología de trabajo.	36
Imagen 16. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Castielfabid.	39
Imagen 17. Gráfica de R-Cuadrada Ajustada para HCV_P en Castielfabid.	41
Imagen 18. Gráfica de Cp de Mallows para factorR.	43
Imagen 19. Histograma para residuos de la distribución normal.	45
Imagen 20. Gráfica de residuos estudentizado vs predicho.	46
Imagen 21. Gráfica de residuos estudentizado vs número de fila.	46
Imagen 22. Matriz de correlación.	47
Imagen 23. Matriz de correlación sin variables polinomiales.	50

Imagen 24. Gráfica comparativa de modelos calculados por Regresión de Componentes Principales (PCR).	55
Imagen 25. Gráfica HCV_P vs HCV_P filtrado de Castielfabid.	57
Imagen 26. Gráfica ajuste del modelo para índices de vegetación suavizados en el punto de muestro Castielfabid.....	59
Imagen 27. Gráfica ajuste del modelo para la Comunidad Valenciana en el punto de muestro Altea.....	62
Imagen 28. Gráfica ajuste del modelo para la Comunidad Valenciana en el punto de muestro Castielfabid.	63
Imagen 29. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Castielfabid.	65
Imagen 30. Histograma para residuos de la distribución normal.	66
Imagen 31. Gráfica de residuos estudentizado vs predicho.	66
Imagen 32. Gráfica de residuos estudentizado vs número de fila.	67
Imagen 33. Matriz de correlación.	67
Imagen 34. Gráfica comparativa de modelos calculados por Regresión de Componentes Principales (PCR).	73
Imagen 35. Gráfica de resultados de R-cuadrada ajustada para los modelos relacionados entre HCV e IVeg.	75
Imagen 36. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Castielfabid.....	76
Imagen 37. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Gilet.....	77
Imagen 38. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Cortes de Pallás.....	77
Imagen 39. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Gandia.	78
Imagen 40. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Biar.....	79
Imagen 41. Gráfica de resultados de R-cuadrada ajustada para los modelos relacionados entre HCV, IV y variables meteorológicas.	83
Imagen 42. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Puebla Tornesa.	84

Imagen 43. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Chelva.	85
Imagen 44. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Gilet.	85
Imagen 45. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Villargordo del Cabriel.	86
Imagen 46. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Buñol.	87
Imagen 47. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Altea.	88
Imagen 48. Representación de la HCV de Fredes en el rango de intervalo de confianza.	90

INDICE DE TABLAS

Tabla 1. Puntos de muestreo y ecosistemas forestales.	16
Tabla 2. Especies muestreadas en 18 parcelas de la Comunidad Valenciana.	17
Tabla 3. Análisis descriptivo estadístico de los datos de HCV de las especies forestales en Fredes.	18
Tabla 4. Características de bandas Sentinel-2.	21
Tabla 5. Escenas de imágenes satelitales.	21
Tabla 6. Estaciones meteorológicas.	32
Tabla 7. Variables meteorológicas analizadas.	35
Tabla 8. Regresión múltiple HCV_P.	38
Tabla 9. Modelos con Mayor R-Cuadrada Ajustada	40
Tabla 10. Modelos con Menor Cp	42
Tabla 11. Modelos con el mejor criterio de información.	44
Tabla 12. Pruebas de Normalidad para Residuos	45
Tabla 13. Matriz de covarianzas	48
Tabla 14. Comparación de modelos datos de entrenamiento y test.	54
Tabla 15. Regresión múltiple HCV_P y las 12 variables de índices de vegetación aplicadas el filtro Savitzki Golay.	58
Tabla 16. Modelos de regresión múltiple para los 18 puntos de muestreo con datos suavizados por el filtro Savitsky Golay.	59
Tabla 17. Regresión múltiple HCV_P y las variables de índices de vegetación y meteorológicas para la Comunidad Valenciana.	61
Tabla 18. Regresión múltiple HCV_P y las variables de índices de vegetación y meteorológicas.	64
Tabla 19. Pruebas de Normalidad para Residuos	65
Tabla 20. Comparación de modelos datos de entrenamiento y test.	72
Tabla 21. Modelos de regresión múltiple para los 18 puntos de muestreo entre la media ponderada de HCV y las variables de índices de vegetación.	74
Tabla 22. Modelos de regresión múltiple para los 18 puntos de muestreo entre la media ponderada de HCV y las variables de índices de vegetación más meteorológicas.	80
Tabla 23. Estimación de la HCV en el punto de muestreo Fredes.	89
Tabla 24. Fechas de toma de muestreo que no ingresan en el intervalo de confianza.	91
Tabla 25. Presupuesto TFM.	99

INDICE

AGRADECIMIENTOS	2
COMPROMISO	3
RESUMEN	4
ABSTRACT	5
1. INTRODUCCIÓN	12
2. OBJETIVOS	13
2.1. Objetivo General.....	13
2.2. Objetivos Específicos.....	13
3. LOCALIZACIÓN DEL ESTUDIO	14
4. DATOS.....	18
4.1. Humedad de combustible vivo para las especies forestales.....	18
4.2. Índices de vegetación.....	20
4.2.1. Adquisición de imágenes satelitales.....	21
4.2.2. Corrección atmosférica de imágenes satelitales con Sen2cor.....	23
4.2.3. Generación de índices de vegetación.....	24
4.3. Datos meteorológicos.....	31
4.3.1. Cálculo de variables meteorológicas.....	33
5. METODOLOGÍA	36
5.1. Análisis de modelos de regresión múltiple entre HCV e índices de vegetación.	37
5.1.1. Método de regresión paso a paso hacia atrás usando la técnica del P-valor.	37
5.1.2. Comparación de modelos de regresión con diferente número de variables.	39

5.1.3.	Análisis del cumplimiento de las hipótesis del modelo de regresión: normalidad, linealidad, homocedasticidad e independencia.....	45
5.1.4.	Análisis de Multicolinealidad.....	47
5.1.5.	Modelos con una muestra de entrenamiento y una de test.....	51
5.1.6.	Análisis de validación cruzada.....	55
5.1.7.	Filtro de suavizado Savitsky-Golay.....	57
5.2.	Análisis de modelos de regresión múltiple entre HCV e índices de vegetación más variables meteorológicas.....	61
5.2.1.	Modelo de regresión múltiple general para la Comunidad Valenciana. ...	61
5.2.2.	Modelo de regresión múltiple para los 18 puntos de muestreo.....	63
5.2.3.	Análisis del cumplimiento de las hipótesis del modelo de regresión: normalidad, linealidad, homocedasticidad e independencia.....	65
5.2.4.	Análisis de Multicolinealidad.....	67
5.2.5.	Modelos con una muestra de entrenamiento y una de test.....	69
6.	RESULTADOS.....	73
6.1.	Modelos de regresión lineal múltiple.....	73
6.1.1.	Análisis entre la media ponderada de la humedad de combustible vivo y los índices de vegetación.....	73
6.1.2.	Análisis entre la media ponderada de la humedad de combustible vivo y los índices de vegetación más las variables meteorológicas.....	79
6.2.	Estimación del valor de la media ponderada de HCV con los modelos de regresión utilizando el intervalo de confianza.....	88
7.	CONCLUSIONES.....	92
8.	BIBLIOGRAFIA.....	95
9.	PRESUPUESTO.....	98
10.	CARTOGRAFÍA.....	100
11.	ANEXOS.....	107

1. INTRODUCCIÓN

La humedad del combustible vivo es una medida del contenido de agua de la vegetación viva, es un fuerte determinante de la ignición del fuego, la velocidad de propagación y la intensidad (Peterson *et al.*, 2008). Si bien el contenido de humedad de la vegetación muerta puede obtenerse con bastante facilidad a partir de modelos dependientes del clima, ya que los combustibles muertos dependen de la variabilidad atmosférica, estimar el contenido de humedad de la vegetación viva es más complicado porque depende de las propiedades fisiológicas que pueden variar entre las diferentes especies de plantas. Para cuantificar el contenido de humedad de la vegetación viva, la humedad del combustible vivo se define como la relación porcentual de la diferencia entre el peso húmedo y el peso seco de una muestra de vegetación (Boksoon *et al.*, 2018).

Estudios previos han determinado metodología para la obtención de la humedad de combustible vivo, por ejemplo (Soriano Sancho, 2016), obteniendo el peligro de incendios forestales según la clasificación de niveles de HCV, bajo: mayor que 120%, moderado: entre 80% y 120%, alto: entre 60% y 80%, y crítico; menor del 60%. La mayor parte de especies en la Comunidad Valenciana presentan valores de humedad que se establecen entre 75-100% un umbral que está por debajo de los incendios más virulentos, aunque este valor depende de la época del año que se considere y también puede variar de un año a otro.

En la actualidad, la cobertura espacial y el muestreo temporal de los datos de HCV están muy limitados ya que el trabajo de campo para las mediciones de HCV requiere mucho esfuerzo. La humedad de combustible vivo se mide de manera manual semanalmente, quincenalmente o mensualmente en un número limitado de sitios de muestreo. Por ejemplo, la Dirección de Prevención de Incendios Forestales de la Generalitat Valenciana a través de la empresa pública VAERSA, generalmente toma muestras de HCV en 18 sitios de la Comunidad Valenciana una vez al mes y en diferentes especies forestales, dejando vacíos de datos en áreas donde el clima y las variaciones geofísicas pueden afectar sustancialmente la HCV.

Un posible enfoque para superar las limitaciones espaciales y temporales de las mediciones manuales de HCV es utilizar índices de vegetación (IVeg) derivados de datos satelitales. Estudios previos demuestran la capacidad de los datos de sensores remotos para la estimación de HCV (Boksoon *et al.*, 2018). Según el tipo de sensor utilizado, se encuentran disponibles diferentes resoluciones espectrales, espaciales y temporales. MODIS es uno de los sensores ópticos más utilizados debido a su alta resolución temporal, proporcionando información espectral en una resolución espacial aproximada lo que limita su uso en áreas de muestreo a pequeña escala o con vegetación muy heterogénea. Otros autores evaluaron el potencial de los sensores remotos por microondas como una alternativa para la estimación de HCV (Marino *et al.*, 2018). Aun se necesita más investigación para proporcionar modelos operacionales y herramientas para la predicción precisa de HCV.

La nueva generación de sensores SENTINEL-2 proporciona información espectral con una resolución espacial alta, lo que ofrece la oportunidad de mejorar las estimaciones de HCV para fines operativos, pero solo se ha probado su aplicación en unos pocos trabajos (Marino *et al.*, 2018).

Físicamente, los HCV dependen de índices meteorológicos, la humedad del suelo, la evapotranspiración y la fisiología de las plantas (Soriano Sancho, 2016). Sin embargo, los IVeg recuperados de las mediciones de sensores remotos satelitales están relacionadas con el verdor de la superficie y la biomasa de la vegetación representada por el índice de área de la hoja verde, que se ven afectadas y por lo tanto correlacionan con el HCV. Los IVeg y HCV son variables interdependientes con tendencias interanuales similares, lo que insinúa una posible estimación de HCV a partir de IVeg (Argañaraz *et al.*, 2016).

No obstante, también existen diferencias entre ellos, el crecimiento de las plantas requiere no solo de la humedad, sino también de la temperatura y precipitación y estos índices meteorológicos también pueden variar durante el desarrollo de las diferentes especies de plantas. Por tanto, para recuperar la HCV de los datos de IVeg, es necesario realizar una investigación cuidadosa de las características de IVeg utilizando en conjunto los datos de mediciones in situ y satelitales.

El objetivo de este trabajo es valorar la capacidad del satélite Sentinel-2, para la estimación de HCV mediante estudios de modelos de regresión. La investigación se centra en diferentes especies forestales representativas del matorral y arbóreo que se encuentran comúnmente en áreas mediterráneas propensas a incendios, y han sido identificadas como especies indicadoras para la evaluación del riesgo de incendios forestales. Cada especie forestal cuenta con su valor de HCV por lo que es necesario calcular la media ponderada de humedad de combustible vivo de las especies que se encuentran un área específica por medio de diferentes criterios que se tomaran en cuenta a lo largo del presente trabajo.

2. OBJETIVOS

2.1. Objetivo General.

Estimar la humedad de combustible vivo usando índices de vegetación extraídos de imágenes obtenidas con el satélite Sentinel-2 juntamente con otras variables de tipo meteorológico, en 18 parcelas de la Comunidad Valenciana teniendo en cuenta la proporción que guarda cada especie forestal.

2.2. Objetivos Específicos.

- Cuantificar el porcentaje de la media ponderada del contenido de humedad de combustible vivo de las especies forestales muestreadas en las 18 parcelas distribuidas en las 3 provincias de la Comunidad Valenciana a partir de la fracción de cabida cubierta.

- Descargar y procesar imágenes satelitales del sensor Sentinel-2 que cubran los 18 puntos de muestreo en el periodo desde enero del 2016 a julio del 2018.
- Valorar la capacidad del satélite Sentinel-2 por medio del cálculo de índices de vegetación para obtener la estimación de HCV mediante la utilización de modelos de regresión múltiples utilizándolos como predictores.
- Mejorar la precisión de los modelos de regresión calculados a partir de índices de vegetación introduciendo variables meteorológicas, extraídas para la misma temporalidad de muestreo.

3. LOCALIZACIÓN DEL ESTUDIO

El presente estudio se realiza en la Comunidad Valenciana, situada en el este de la península Ibérica, cuenta con una superficie de 2.325.500 ha, según el Instituto Cartográfico Valenciano, esta zona cuenta con una superficie forestal de 1.319.517 ha, lo que supone el 57% del territorio total. Esta superficie cuenta con una clasificación de cobertura vegetal definidos por el Plan de Acción Territorial Forestal de la Comunidad Valenciana (PATFOR, 2011), son 23 los ecosistemas forestales como se muestra en la Imagen 1, 10 de los ecosistemas pertenecen a arbolado, 12 a no arbolados y uno a mixto. Se observa que de los ecosistemas destaca los pinares de *Pinus halepensis* perteneciente al arbolado y la garriga que es una especie arbustiva cuya especie es *Quercus coccifera* perteneciente al no arbolado.

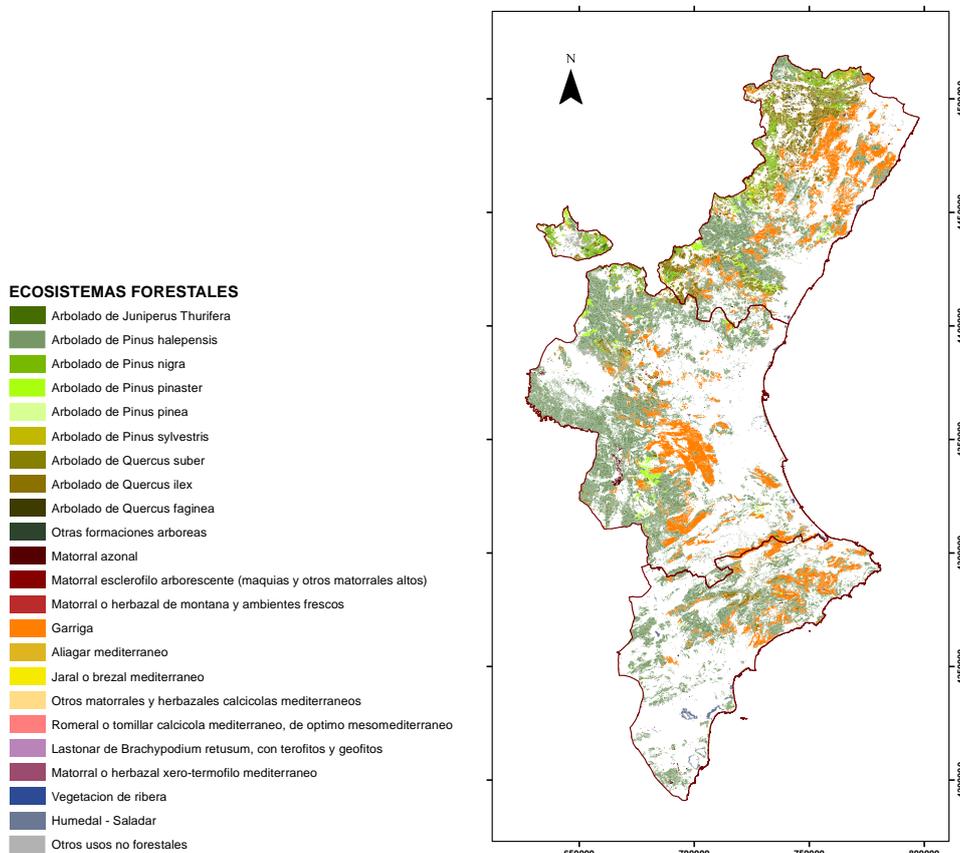


Imagen 1. Ecosistemas forestales de la Comunidad Valenciana.

Fuente: (PATFOR, 2011)

Las especies forestales con más representación en la Comunidad Valenciana aparte de las dos ya nombradas son *Quercus ilex*, *Pinus nigra* y *Pinus pinaster*, del estrato arbóreo, y a *Rosmarinus officinalis*, *Ulex parviflorus*, *Pistacea lentiscus* y *Juniperus oxycedrus* en el estrato arbustivo.

A partir de lo antes mencionado la empresa Valenciana de Aprovechamiento Energético de Residuos, SA (VAERSA) perteneciente a la Conselleria de Agricultura, Medio Ambiente, Cambio Climático y Desarrollo Rural, localiza una serie de puntos de muestreo donde se toman muestras combustibles de las especies forestales más representativas de la zona, con el fin de obtener una elevada representatividad respecto al territorio.

Se localiza un total de 18 puntos de muestreo como se muestra en la Imagen 2. Estos puntos de muestreo son representativos de la zona en cuanto a especies, estratos de vegetación y estructura del combustible. También se consideró los aspectos climáticos, topográficos y de accesibilidad.

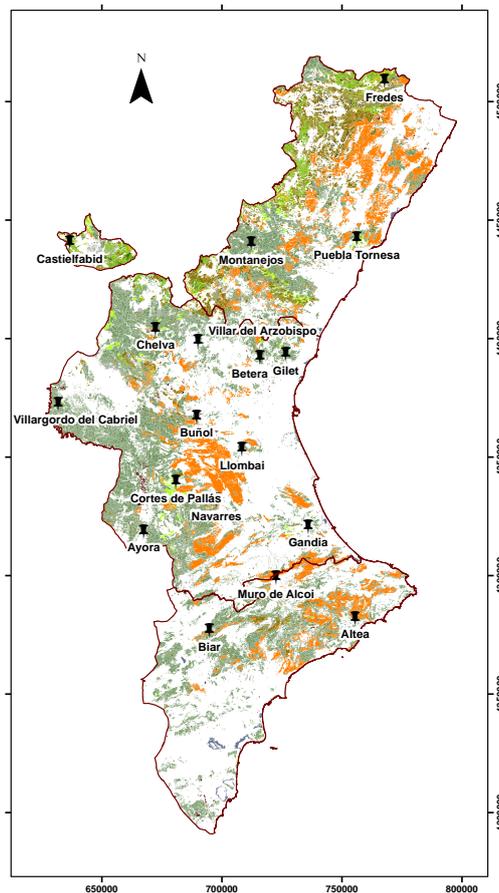


Imagen 2. Localización de los puntos de muestreo.

Fuente: VAERSA

En la Tabla 1 se detalla las características de cada punto de muestreo, sus coordenadas y el ecosistema forestal al que pertenece según la clasificación de PATFOR.

Tabla 1. Puntos de muestreo y ecosistemas forestales.

Fuente: VAERSA. (PATFOR, 2011).

Nº	PUNTO DE MUESTREO	Coordenadas (UTM)			ARBOLADO/ DESARBOLADO	ECOSISTEMA
		X	Y	Z		
1	Fredes	767823	4508581	1021	Arbolado	Arbolado de <i>Pinus nigra</i>
2	Puebla Tornesa	756294	4442164	371	Arbolado	Arbolado de <i>Pinus pinaster</i>
3	Montanejos	712282	4440006	635	Arbolado	Arbolado de <i>Pinus halepensis</i>
4	Castielfabid	636667	4440355	1315	Arbolado	Arbolado de <i>Juniperus Thurifera</i>
5	Chelva	672464	4404094	721	Desarbolado	Garriga
6	Villar del Arzobispo	690194	4398656	480	Desarbolado	Matorral esclerófilo arborescente (maquias y otros matorrales altos)
7	Betera	715904	4391951	207	Arbolado	Arbolado de <i>Pinus halepensis</i>
8	Gilet	726619	4393272	302	Arbolado	Arbolado de <i>Pinus halepensis</i>
9	Villargordo del Cabriel	631749	4372505	857	Arbolado	Arbolado de <i>Pinus halepensis</i>
10	Buñol	689507	4366847	467	Desarbolado	Lastonar de <i>Brachypodium retusum</i> , con terófitos y geófitos
11	Llombai	708449	4353298	139	Desarbolado	Garriga
12	Cortes de Pallás	680786	4339390	868	Arbolado	Arbolado de <i>Pinus pinaster</i>
13	Navarrés	697857	4331838	220	Desarbolado	Garriga
14	Ayora	667506	4318494	692	Arbolado	Arbolado de <i>Pinus halepensis</i>
15	Gandia	736028	4320541	396	Desarbolado	Matorral o herbazal xero-termófilo mediterráneo
16	Muro de Alcoi	722610	4298929	477	Desarbolado	Garriga
17	Altea	755562	4281788	301	Arbolado	Arbolado de <i>Pinus halepensis</i>
18	Biar	694871	4276564	829	Arbolado	Arbolado de <i>Pinus halepensis</i>

Hay 11 parcelas de muestreo que se encuentran en el estrato arbolado y 7 en desarbolado; de las parcelas que se encuentran en arbolado 7 son *Pinus Halepensis* siendo el estrato arbustivo más grande de la Comunidad Valenciana típica de los montes mediterráneos; el resto de arbolados corresponden a *Pinus Pinaster*, *Pinus Nigra* y *Juniperus Thurifera*. Las parcelas con vegetación desarbolada, en su mayor parte pertenece a *Garriga* que representa a los matorrales seguida por los *Matorrales o Herbazales xero-termófilo mediterráneo* y *Lastonar de Brachypodium retusum*, con terófitos y geófitos.

La Tabla 2 muestra las especies que se toma en cada punto de muestreo por VAERSA de enero del 2016 a julio del 2018, la elección de las especies sigue la premisa de representatividad en los ecosistemas forestales, se toma muestras de un total de 19 especies distintas típicas de los montes mediterráneos; las especies que se repiten en los puntos de muestreo son *Rosmarinus officinalis* y *Pinus Halepensis*. Se recoge un total de 125 muestras de vegetación en cada toma de muestra repartidas en los 18 puntos de muestreo, hay que tomar en cuenta que VAERSA recoge la muestrea de las especies forestales una vez al mes y en ciertos puntos de muestreo en los meses de verano lo hace quincenalmente.

Tabla 2. Especies muestreadas en 18 parcelas de la Comunidad Valenciana.
Fuente: VAERSA.

Nº	Punto de Muestreo	Ph	Qc	Qi	Ro	Up	Em	Ca	Pp	Pn	Ps	Jo	Jc	Jt	Jph	Pl	Cs	St	Au	Ra	Total
1	Fredes			x	x	x				x	x	x			x						7
2	Puebla Tornesa	x	x		x	x	x		x			x				x					8
3	Montanejos	x	x		x	x						x			x	x				x	8
4	Castielfabid			x	x					x			x	x							5
5	Chelva	x	x	x	x	x												x			6
6	Villar del Arzobispo	x	x		x		x					x			x						6
7	Betera	x	x		x	x	x					x				x		x			8
8	Gilet	x	x		x		x					x				x					6
9	Villargordo del Cabriel	x	x		x							x			x			x	x		7
10	Buñol	x	x		x	x	x	x				x					x				8
11	Llombai	x	x		x	x	x					x				x	x				8
12	Cortes de Pallás	x	x	x	x	x		x	x			x									8
13	Navarrés	x	x		x	x	x					x									6
14	Ayora	x	x	x	x	x												x			6
15	Gandia	x	x	x	x	x	x		x							x					8
16	Muro de Alcoi	x	x	x	x	x		x				x				x					8
17	Altea	x	x		x	x	x									x	x				7
18	Biar	x			x	x						x			x						5
	Total	16	15	7	18	14	9	3	3	2	1	13	1	1	5	8	3	4	1	1	125

Nota: correspondencia entre abreviaturas y nombre de las especies: **Ph**(*Pinus halepensis*); **Qc** (*Quercus coccifera*); **Qi**(*Quercus ilex*); **Ro**(*Rosmarinus officinalis*); **Up**(*Ulex parviflorus*); **Em**(*Erica multiflora*); **Ca**(*Cistus albidus*); **Pp**(*Pinus pinaster*); **Pn**(*Pinus nigra*); **Ps**(*Pinus sylvestris*); **Jo**(*Juniperus oxycedrus*); **Jc**(*Juniperus communis*); **Jt**(*Juniperus thurifera*); **Jph**(*Juniperus phoenicea*); **Pl**(*Pistacea lentiscus*); **Cs**(*Ceratonia siliqua*); **St**(*Stipa tenacissima*); **Au**(*Arbutus unedo*); **Ra** (*Rhamnus alaternus*)

4. DATOS

4.1. Humedad de combustible vivo para las especies forestales.

El análisis de humedad de combustible vivo que contiene cada especie vegetal detallada en la Tabla 2 es primordial, ya que forma parte de los ecosistemas forestales endémicos de la Comunidad Valenciana, mismos que al ser afectados por factores externos como su estado fenológico, variables meteorológicas a más del relieve y el suelo, presentan una mayor o menor deshidratación a lo largo del tiempo, existiendo la probabilidad de que dichas especies presenten características de un componente combustible, considerándose como un parámetro del comportamiento de los incendios forestales.

Por tal razón el presente estudio se basa en los datos de humedad de combustible vivo (HCV) calculados por VAERSA para la Generalitat Valenciana, para las especies anteriormente detalladas de los 18 puntos de muestreo en el periodo de enero del 2016 a julio del 2018. Los datos de HCV han sido recopilados en intervalos de aproximadamente un mes. La Tabla 3 muestra como ejemplo el análisis descriptivo estadístico de los datos de HCV para el punto de muestreo Fredes de las diferentes especies analizadas en el sector.

Tabla 3. Análisis descriptivo estadístico de los datos de HCV de las especies forestales en Fredes.

Fuente: VAERSA.

	Humedad de combustible vivo (%)						
	Ro	Jo	Up	Qi	Jph	Pn	Ps
Mínimo	58.46	63.59	63.11	60.57	52.37	79.41	89.77
Primer cuartil	100.38	85.09	82.22	74.56	82.31	95.87	109.88
Media	120.718	100.527	96.5359	80.6714	86.3559	99.8617	115.431
Mediana	119.59	89.66	90.43	76.69	87.8	100.15	112.45
Tercer cuartil	138.91	117.94	110.3	84.05	91.88	103.23	125.55
Máximo	244.75	175.32	175.9	150.88	105.03	119.25	138.54

La humedad de combustible vivo varía según la especie forestal y el tiempo en que se realice el muestreo, en la Imagen 3 se observa el comportamiento temporal de las siete especies forestales recogidas en el punto de muestreo Fredes, todas las especies forestales presentan valores bajos en los meses de verano siendo julio, agosto y septiembre; tienen un comportamiento en toda la temporalidad muy parecido a pesar de que las especies con valores más altos de HCV son *Rosmarinus officinalis* y *Pinus sylvestris*.

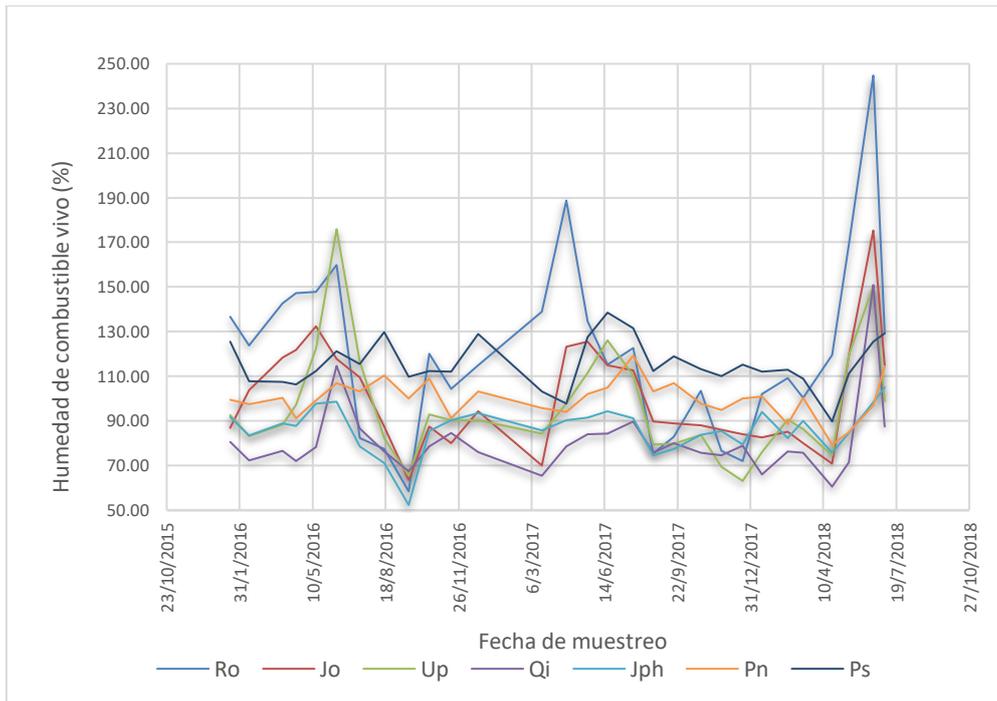


Imagen 3. Comportamiento temporal de la HCV en el punto de muestreo Fredes.

El cálculo de la media ponderada de la humedad de combustible vivo para cada punto de muestreo se realiza a partir de la fracción de cabida cubierta (FCC) es decir la parte de suelo ocupado por la proyección de las copas de la vegetación, siendo diferente para cada punto de muestreo, VAERSA nos ha proporcionado dichos datos para el cálculo de la media ponderada de HCV. La mayoría de las especies forestales tienen una espesura incompleta al ser arbolados y matorrales es decir que no existirá tangencia entre sus copas y la distancia entre ellas será superior al diámetro medio de sus copas.

La Imagen 4 muestra la proporción de fracción de cabida cubierta que hay en cada punto de muestreo, se observa que hay un gran porcentaje de **Ro** (*Rosmarinus officinalis*), **Ph** (*Pinus Halepensis*) y **Qc** (*Quercus coccifera*) siendo las más predominantes.

La obtención de la media ponderada de HCV en cada punto de muestreo y para cada fecha se obtiene a partir de la suma del producto de HCV de cada especie y su FCC esto dividido para la suma de FCC.

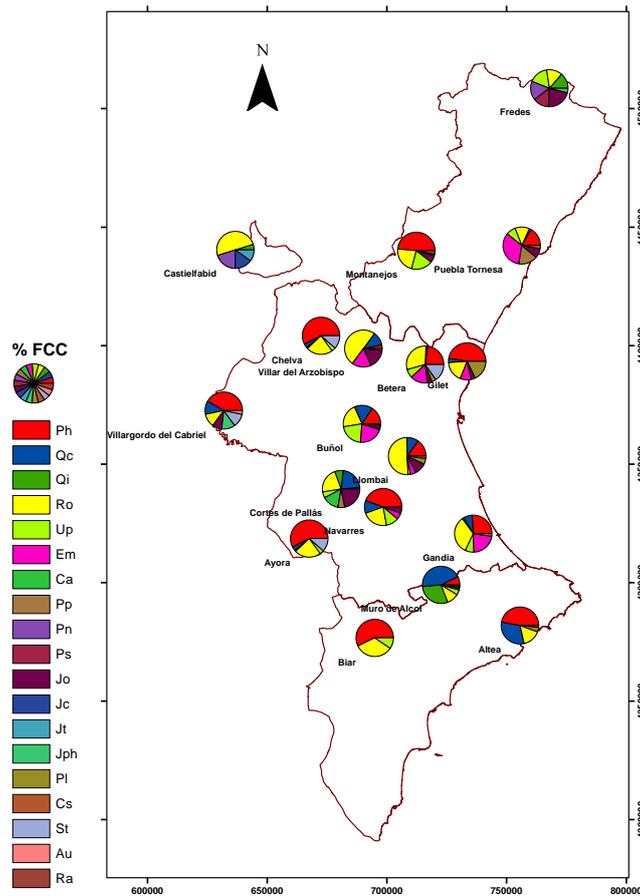


Imagen 4. Porción de FCC de las especies forestales en los puntos de muestreo.
Fuente: VAERSA.

4.2. Índices de vegetación.

Los índices de vegetación (IVeg) son desarrollados con la finalidad de obtener información sobre las características y estado de la vegetación, estos son calculados partiendo de datos espectrales adquiridos desde imágenes satelitales. Los IVeg aprovechan el comportamiento específico radiométrico de la vegetación sana en las diferentes bandas espectrales, especialmente entre el visible y el infrarrojo cercano; el visible perteneciente a la banda roja, muestra los pigmentos de la hoja que absorben la mayor parte de la energía que reciben, estas sustancias afectan en menor fracción al infrarrojo cercano. Estos contrastes son los que aprovecha los IVeg para establecer diferencias en la vegetación, sana y vigorosa, cuando los contrastes son mayores, y vegetación enferma o de baja densidad, cuando los contrastes son menores (Carvacho *et al.*, 2010).

El número de sistemas satelitales de observación de la tierra ha ido aumentando y por tal razón existe mayor información para uso de los usuarios. El satélite Sentinel-2 ofrece imágenes de libre disposición, obteniendo series de datos temporales con resoluciones espaciales buenas y densas ampliando su uso en varios campos.

El satélite Sentinel-2 forma parte de la misión de observación de la Tierra desarrollada por la Agencia Espacial Europea (ESA) y se lanzó el 23 de julio de 2015. Este satélite transporta un solo instrumento multispectral (MSI) con 13 canales espectrales. La resolución varía entre 10 y 60 m. Las bandas visible y de infrarrojo cercano (NIR) tienen 10, las bandas de borde rojo de vegetación y las bandas de infrarrojo de onda corta (SWIR) tienen 20, y el resto de las bandas tienen una resolución de 60 m (Agency, 2015), como se muestra en la Tabla 4.

Tabla 4. Características de bandas Sentinel-2.
Fuente: (SENTINEL, 2018)

Bandas		Longitud de onda (nm)	Resolución espacial (m)
Banda 2	Azul	490	10
Banda 3	Verde	560	
Banda 4	Rojo	665	
Banda 8	Infrarrojo cercano - NIR	842	
Banda 5		705	
Banda 6		740	
Banda 7		783	
Banda 8A		865	
Banda 11		Infrarrojo lejano - SWIR	1610
Banda 12	2190		
Banda 1	Aerosol	443	60
Banda 9	Vapor de agua	940	
Banda 10	Cirrus	1375	

Con estos antecedentes, se realiza la recopilación de imágenes satelitales pertenecientes a la Comunidad Valenciana en el periodo de enero del 2016 a julio del 2018 con el fin de obtener el cálculo de diferentes índices de vegetación en dicho periodo y relacionarlos con la media ponderada de humedad de combustible vivo en 18 puntos de muestreo.

4.2.1. Adquisición de imágenes satelitales.

La descarga de imágenes Sentinel-2 se realizó desde la página web de Copernicus (COPERNICUS, 2018), en donde fue necesario registrarse creando una cuenta para la posterior descarga de imágenes. Al estar los 18 puntos de muestreo distribuidos por toda la Comunidad Valenciana se descargó siete escenas de imágenes satelitales que cubren todos los puntos de muestreo y de cada escena una por cada mes en el periodo de tiempo antes mencionado. En la Tabla 5 se observa las escenas y puntos de muestreo que cubren las imágenes satelitales descargadas.

Tabla 5. Escenas de imágenes satelitales.

SXH	SXJ	SYH	SYJ	TBF	TXK	TYK
Biar	Ayora Navarrés Cortes de Pallás Buñol Villargordo de Cabriel	Altea	Muro del Alcoi Gandia Llombai Betera Gilet	Fredes	Castielfabid Villar del Arzobispo Chelva	Puebla Tornesa Montanejos

Las imágenes descargadas se encuentran en un fichero ZIP, el cual contiene algunas carpetas con una colección de información del producto SENTINEL-2 entre ellas incluye:

- Un archivo *manifest.safe* que contiene la información general del producto en XML.
- Una imagen de vista previa en formato JPEG2000.
- Subcarpetas para conjuntos de datos de medición, incluidos datos de imagen (gránulos / mosaicos) en formato GML-JPEG2000.
- Subcarpetas para información de nivel de datos
- Una subcarpeta con datos auxiliares (por ejemplo, boletín de Sistemas de referencia y rotación de la Tierra (IERS)).
- Vistas previas de HTML.

En la Imagen 5, la Agencia Espacial Europea (ESA) describe el formato físico de archivos Sentinel-2 antes mencionados.

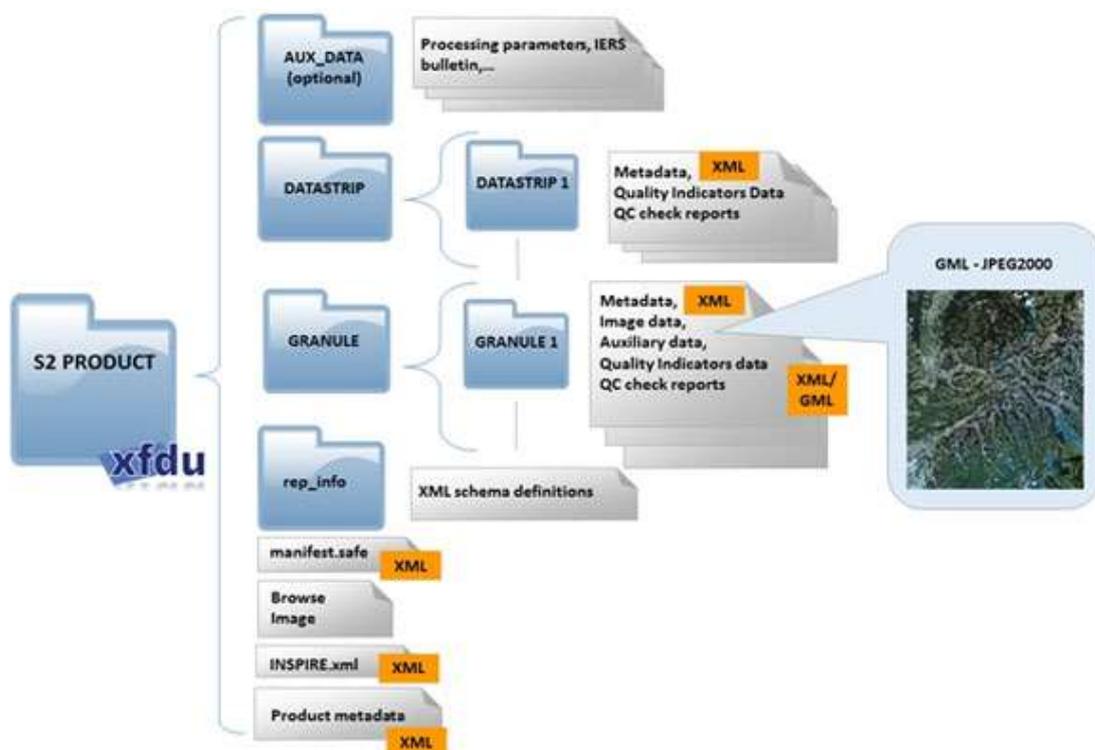


Imagen 5. Formato físico de los archivos Sentinel-2.

Fuente: (ESA, 2018)

4.2.2. Corrección atmosférica de imágenes satelitales con Sen2cor.

Las imágenes satelitales descargadas pasaran por el software SNAP mediante su extensión SEN2COR el cual pasará a las imágenes del nivel de producto L1C a L2A realizando la corrección atmosférica, del terreno y del cirro de los datos de entrada. Crea imágenes de reflectancia corregidas en el terreno y en el cirro; Grosos ópticos de aerosol adicionales, vapor de agua, mapas de clasificación de escena e indicadores de calidad para las probabilidades de nubes y nieve. El formato del producto de salida es equivalente al Producto de Usuario de Nivel 1C: imágenes JPEG 2000, y se obtiene tres resoluciones diferentes, 60, 20 y 10 m. (European Space Agency, 2018).

Como resultado se obtiene unas imágenes más nítidas y con valores de píxel corregidos, en la Imagen 6 a la izquierda se observa como ejemplo la imagen satelital en nivel L1C del 1 de mayo del 2016 perteneciente a la escena SXJ, después de haber pasado por el procesamiento en SNAP se observa que las nubosidades bajan y son hasta omitidas a la derecha.

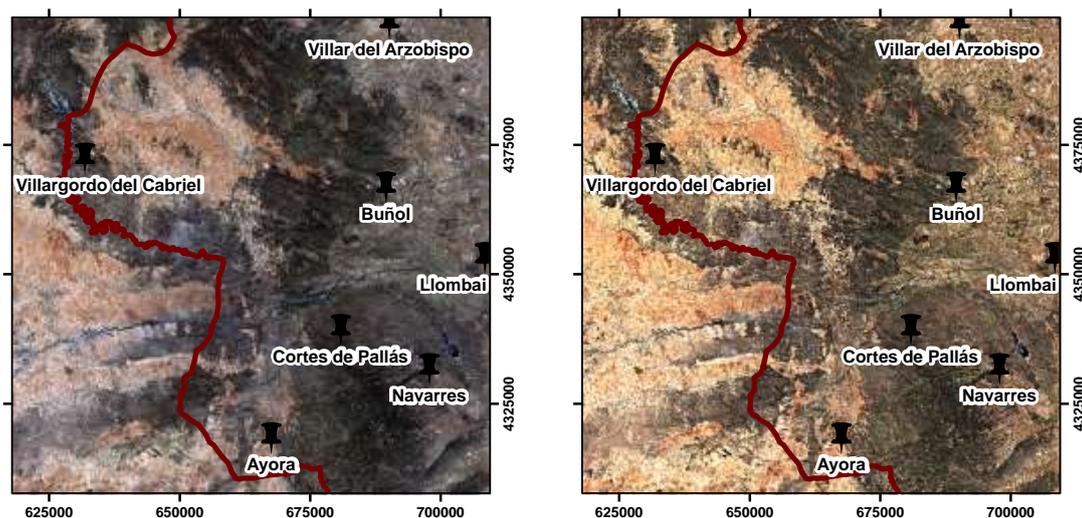


Imagen 6. Resultado de la corrección atmosférica de las imágenes satelitales en Sen2cor.

Esta corrección se realizó en todas las imágenes satelitales descargadas desde enero del 2016 hasta julio del 2018. En la Imagen 7 se observa como las escenas descargadas y procesadas cubren en su totalidad la superficie de la Comunidad Valenciana y abarca todos los puntos de muestreo.



Imagen 7. Mosaico de imágenes satelitales Sentinel-2 para la Comunidad Valenciana.

4.2.3. Generación de índices de vegetación.

Una vez procesadas las imágenes satelitales se procede a generar 6 índices de vegetación comúnmente utilizados para estimar HCV por varios autores: índice de vegetación normalizado (NDVI), índice de vegetación mejorado (EVI), índice de vegetación ajustado al suelo (SAVI), índice de diferencia de agua normalizada (NDWI), índice de resistencia atmosférica visible (VARI), índice de vegetación verde (VIgreen), la mayor parte de estos índices miden las variaciones de verdor, mismas que pueden estar relacionadas indirectamente con el contenido de agua de la hoja ya que las variaciones de humedad de la vegetación afectan la actividad de la clorofila, la estructura interna de la hoja y el índice de área foliar (LAI) de la vegetación.

NDVI (Índice de vegetación normalizada)

Este índice se utiliza para estimar la cantidad, calidad y desarrollo de la vegetación con base a la medición de la intensidad de la radiación de ciertas bandas del espectro electromagnético que la vegetación emite o refleja.

Para el cálculo de dicho índice se utiliza la información de las bandas roja e infrarroja, mediante la siguiente fórmula, (Rouse Jr., 1973):

$$NDVI = \frac{NIR-ROJO}{NIR+ROJO} \quad (1)$$

Mediante la diferencia entre la reflectancia de las bandas 8 (infrarrojo cercano) y 4 (visible – rojo) dividido por la suma de estas dos bandas de reflectancia.

Un aspecto de gran interés del NDVI es que varía entre unos márgenes conocidos (-1 y +1), lo que facilita notablemente su interpretación. Por ejemplo, podemos señalar como umbral crítico para cubiertas vegetales un valor de NDVI en torno a 0,1 y para la vegetación densa entre 0,5 y 0,7 (Holben, 1986)

EVI (Índice de vegetación mejorado)

El índice de vegetación mejorado (EVI), propuesto por (Huete A. R., 1997) siendo una alternativa más sólida de índices tradicionales, por ser más robusto frente a la aportación de suelo y de las influencias atmosféricas. Se define como:

$$EVI = 2.5 \frac{NIR-ROJO}{NIR+6*ROJO-7.5*AZUL+1} \quad (2)$$

Consigue la respuesta de las variaciones estructurales del dosel vegetal incluyendo el LAI tipo y arquitectura del dosel y fisonomía de las plantas. Fue desarrollado para optimizar la señal de la vegetación con sensibilidad mejorada para altas densidades de biomasa, separando la señal proveniente de la vegetación y de la influencia atmosférica.

SAVI (Índice de vegetación ajustado del suelo)

Entre los factores que modifican notablemente el comportamiento del NDVI figura la proporción de vegetación/suelo observada por el sensor. Los mismos valores de NDVI pueden corresponder a cubiertas vigorosas, pero poco densas, o a cubiertas densas con poca vitalidad. Para incluir explícitamente el factor suelo, clave cuando se trabaja en zonas áridas. Huete y colaboradores propusieron incluir en la fórmula del NDVI un parámetro (L) que indica un posible valor en 0.5, que ajuste el índice a una reflectividad promedio del fondo (Huete *et al.*, 1992) de esta forma se calcula el índice de vegetación ajustado al suelo (SAVI).

Para el cálculo de SAVI se utiliza la información de las bandas roja e infrarroja, mediante la siguiente fórmula, (Huete A. , 1988):

$$SAVI = \frac{1.5*(NIR-ROJO)}{NIR+ROJO+0.5} \quad (3)$$

Este índice pretende minimizar el efecto del suelo sobre los otros índices de vegetación, especialmente manifiesto en superficies parcialmente recubiertas.

NDWI (Índice de diferencia de agua normalizada)

El índice de diferencia de agua normalizado (NDWI, por sus siglas en inglés) se utiliza para la detección remota de agua líquida de la vegetación desde el espacio. La ecuación

para determinar el NDWI utiliza las bandas verdes y NIR (McFeeters, 1996) desde que fue probado como óptimo para la extracción de agua (Kaplan & Avdan, 2017); las bandas están ubicadas en la meseta de alta reflectancia de las copas de vegetación. Perciben profundidades similares a través de los doseles de vegetación. NDWI es menos sensible a los efectos atmosféricos que el NDVI.

$$NDWI = \frac{Verde - NIR}{Verde + NIR} \quad (4)$$

VARI (Índice de resistencia atmosférica visible)

La resistencia del VARI a los efectos atmosféricos (en comparación con el NDVI) se logra mediante un proceso de autocorrección del efecto atmosférico en el canal rojo, utilizando la diferencia en la radiancia entre los canales azul y rojo para corregir la radiancia. Las simulaciones que utilizan cálculos de transferencia radiactiva en espectros aritméticos y de superficie natural, para diversas condiciones atmosféricas, muestran que VARI tiene un rango dinámico similar al NDVI, pero es, en promedio, cuatro veces menos sensible a los efectos atmosféricos que el NDVI (Kaufman *et al.*, 1992). Se emplea para evitar problemas de saturación a valores altos de índice que presentan SAVI y NDVI. La ecuación para calcular VARI es la siguiente (Tello *et al.*, 2017):

$$NDWI = \frac{Verde - Rojo}{Verde + Rojo - Azul} \quad (5)$$

VIgreen (Índice de vegetación verde)

VIgreen índice de verdor es sensible a cambios en la absorción de clorofila de la vegetación y al LAI, que coinciden con los cambios en el contenido de agua (Hardy y Burgan 1999)

Se calcula a partir de la ecuación (Tucker *et al.*, 1978):

$$VIgreen = \frac{Verde - Rojo}{Verde + Rojo} \quad (6)$$

donde Verde se refiere a la banda visible del verde y Rojo la banda del rojo.

En términos del equilibrio entre la reflectancia verde (longitudes de onda de 560nm) y la reflectancia roja (longitudes de onda de alrededor de 665 nm), se puede encontrar tres grupos de patrones de reflectancia espectral para los componentes principales de la cobertura del suelo (Takeshi *et al.*, 2010):

Vegetación verde (coníferas, árboles de hoja caduca y pasto): la reflectancia de la banda verde es más alta que la reflectancia de la banda roja.

Suelos (arena marrón, limo y arcilla seca): la reflectancia de la banda verde es más baja que la reflectancia de la banda roja.

Agua / nieve: la reflectancia de la banda verde y la reflectancia de la banda roja son en su mayoría iguales.

Con estos antecedentes se procede con la obtención del valor concreto de los 6 índices de vegetación para los 18 puntos de muestreo en análisis, al tener como dato la coordenada de donde se realizó el muestreo de las especies forestales se procede a sacar la media de cada índice de vegetación en un radio de 30 metros es decir se tomarán los nueve pixeles que se encuentren rodeando al punto de muestreo, este proceso se lo realizó mediante el desarrollo de un script en Python (ANEXO 1), como ejemplo se muestra en la Imagen 8 la superficie tomada para el cálculo de la media de los índices de vegetación en el punto de muestreo Buñol, de esta manera se obtendrá un valor promedio del ecosistema forestal de cada punto de muestreo.

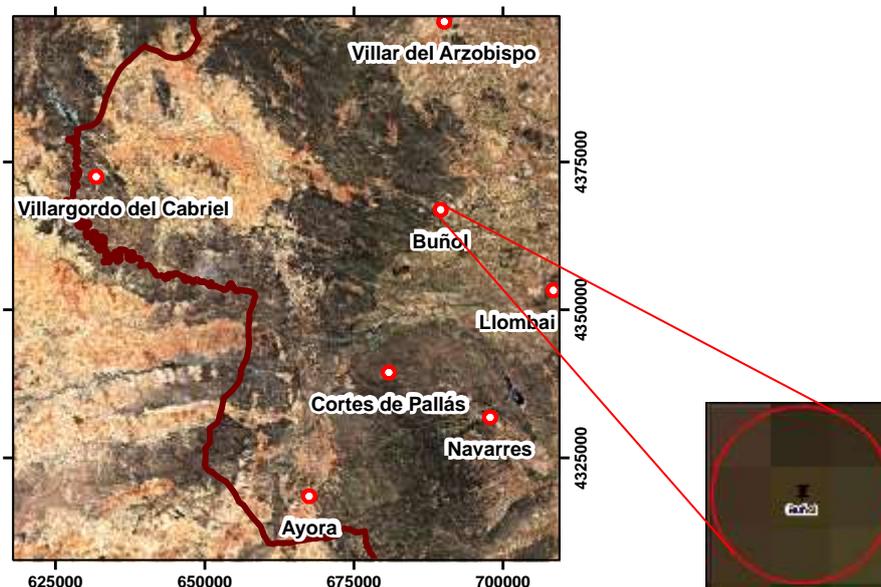


Imagen 8. Superficie para el cálculo de índices de vegetación.

De esta manera se obtiene los índices de vegetación para cada imagen satelital en cada punto de muestreo, como ejemplo en la Imagen 9 se observa la tenencia de los valores obtenidos para el periodo en estudio en el punto de muestreo Buñol. Los índices de vegetación demuestran un comportamiento homogéneo según las características de cada índice, hay que tener en cuenta que el rango de SAVI no tiene cambios muy evidentes ya que la conducta del mismo en todas las fechas de muestreo actúa de la misma manera y es similar por el tipo de vegetación que se encuentra en los pixeles estudiados. El rango de este índice de vegetación se encuentra entre 1.49 y 1.50. Así también se puede observar que el NDVI es el más homogéneo al ser un índice normalizado.

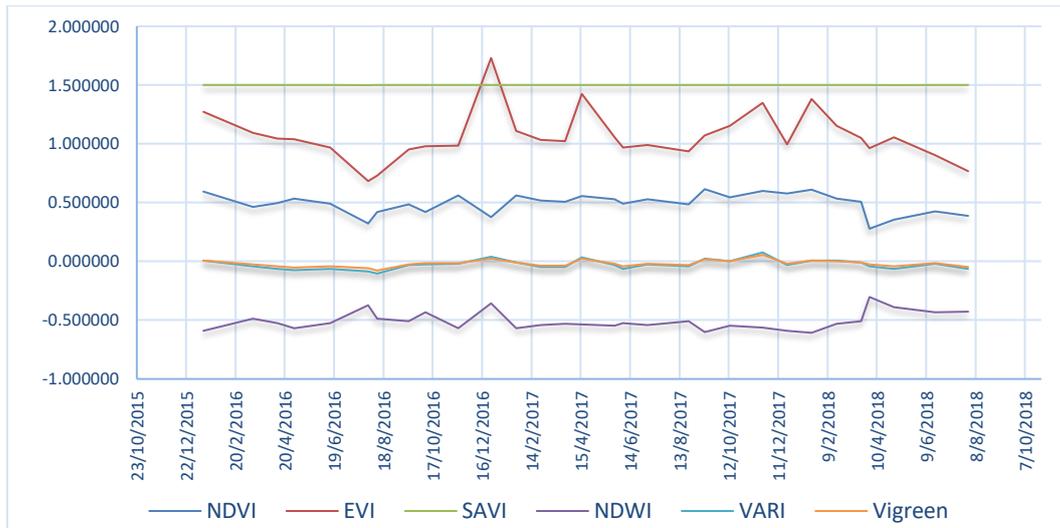


Imagen 9. Tendencia de índices de vegetación en el periodo enero 2016-julio 2018 en el punto de muestreo Buñol.

Los datos calculados de IVeg pertenecen a la fecha de toma del satélite de la imagen, por tal razón, se realiza una interpolación de los datos para obtener los mismos en la fecha de muestreo en campo por VAERSA de las especies forestales pertenecientes a cada punto de muestreo. Se realizó una interpolación lineal obteniendo los datos como se muestra en la Imagen 10 se observa que la gráfica no varía mucho en comparación con la Imagen 9 ya que la fecha de toma de imagen satelital no está muy lejana a la fecha de la toma de muestra en campo. Los datos resultantes de la interpolación son los que se van a relacionar con la HCV en el punto de muestreo Buñol.

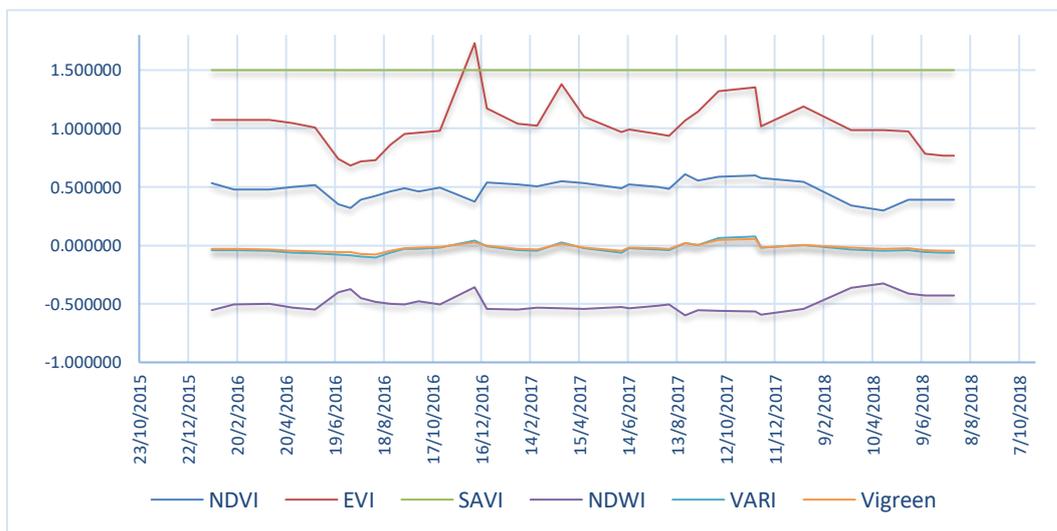


Imagen 10. Tendencia de índices de vegetación interpolados a la fecha de toma de muestra en el punto de muestreo Buñol.

Los índices de vegetación antes mencionados están relacionados con la media ponderada de humedad de combustible vivo, ya que han sido desarrollados con la finalidad de obtener información sobre el estado de la vegetación y las características de la mismas, entregando así la estimación de la calidad y desarrollo de la planta. En la Imagen 11 se puede observar la relación existente entre uno de los índices de vegetación este caso NDVI y la HCV en el punto de muestreo Castielfabid, el rango de los valores en que se encuentra el IVeg es de 0.22 a 0.46 indicando que la vegetación en esta temporalidad del muestreo es alta, se observa como el comportamiento de la misma coincide con la HCV, el NDVI aumenta en los meses de primavera y los picos de HCV siguen el mismo rumbo mientras que en los meses secos de verano el índice tiene el comportamiento contrario.

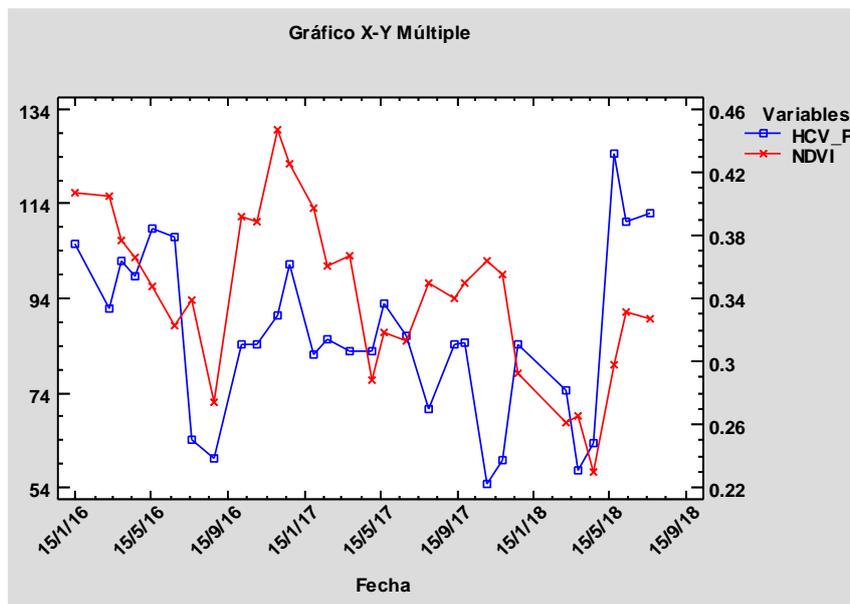


Imagen 11. Relación entre el índice de vegetación NDVI y la HCV en el punto de muestreo Castielfabid.

En el presente trabajo los seis índices de vegetación se obtienen para las 18 parcelas de estudio, sin embargo, estos se pueden calcular para toda la Comunidad Valenciana. En la Imagen 12 y 13 se puede observar como ejemplo el NDVI en el año 2017 para toda la Comunidad en sus cuatro estaciones del año. El NDVI es el índice que muestra un comportamiento más homogéneo con respecto a los otros cinco IVeg, al ser un índice normalizado de -1 a 1 se pretende mostrar que los valores cercanos a -1 son sectores que no cuentan con cubierta vegetal (agua, suelo desnudo), los valores cercanos a 0 indican ausencia de vegetación (barbechos), mientras que los valores cercanos a 1 proyectan la vigorosidad de la vegetación.

En la Imagen 12 a la izquierda se puede observar el comportamiento del NDVI en invierno específicamente en el mes de diciembre, se tiene tonalidades más verdes, es decir que en este temporal la vegetación es un poco más densa que en verano como se puede observar a la derecha de misma imagen, el NDVI en verano específicamente en el mes de agosto cuenta con una tonalidad más naranja indicando que la cubierta de la vegetación es menor.

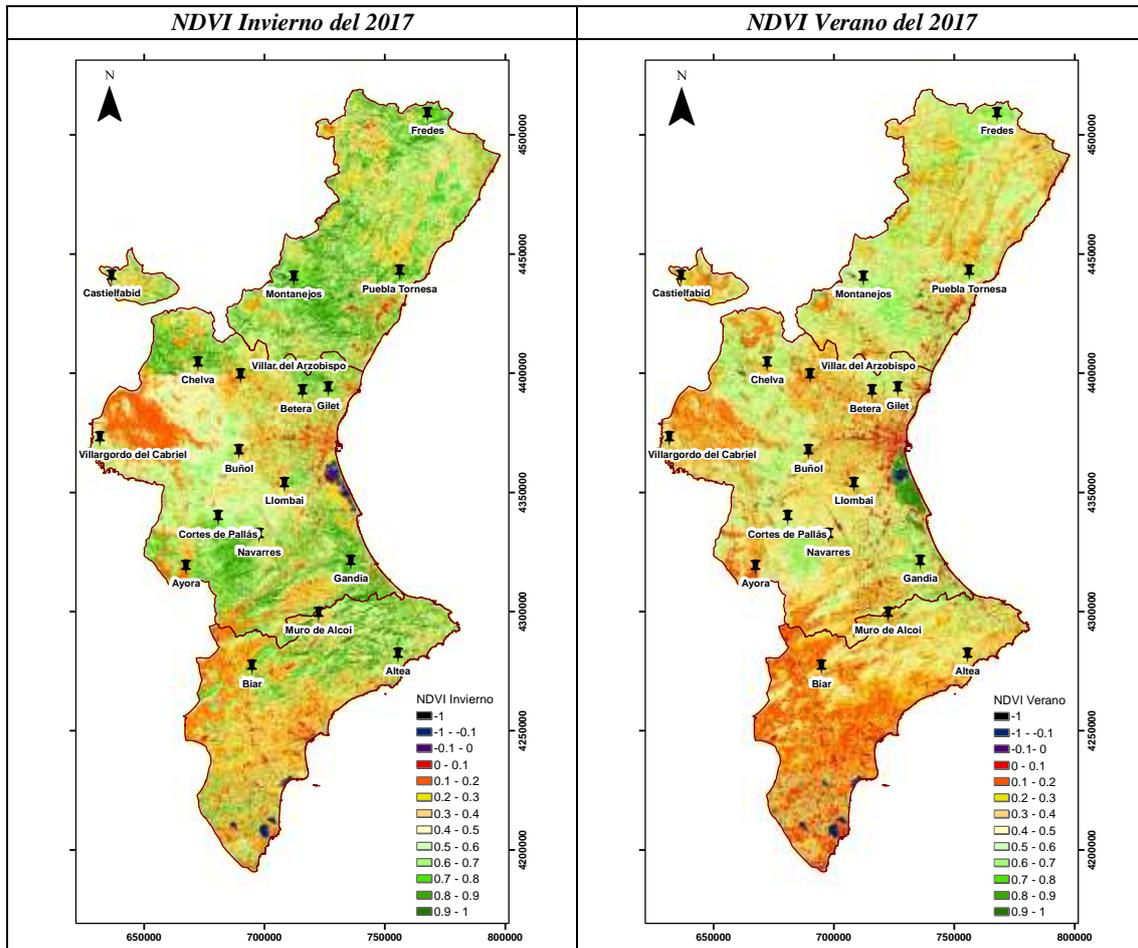


Imagen 12. NDVI obtenidos para el año 2017 en la Comunidad Valenciana. Composiciones de máximo valor para los meses de diciembre (invierno) y agosto (verano)

En la Imagen 13 a la izquierda se puede observar el comportamiento del NDVI en primavera específicamente en el mes de marzo, se tiene tonalidades más verdes, es decir que en este temporal la vegetación es un poco más densa que en otoño como se puede observar a la derecha de misma imagen, el NDVI en otoño específicamente en el mes de octubre cuenta con una tonalidad más naranja indicando que la cubierta de la vegetación es menor.

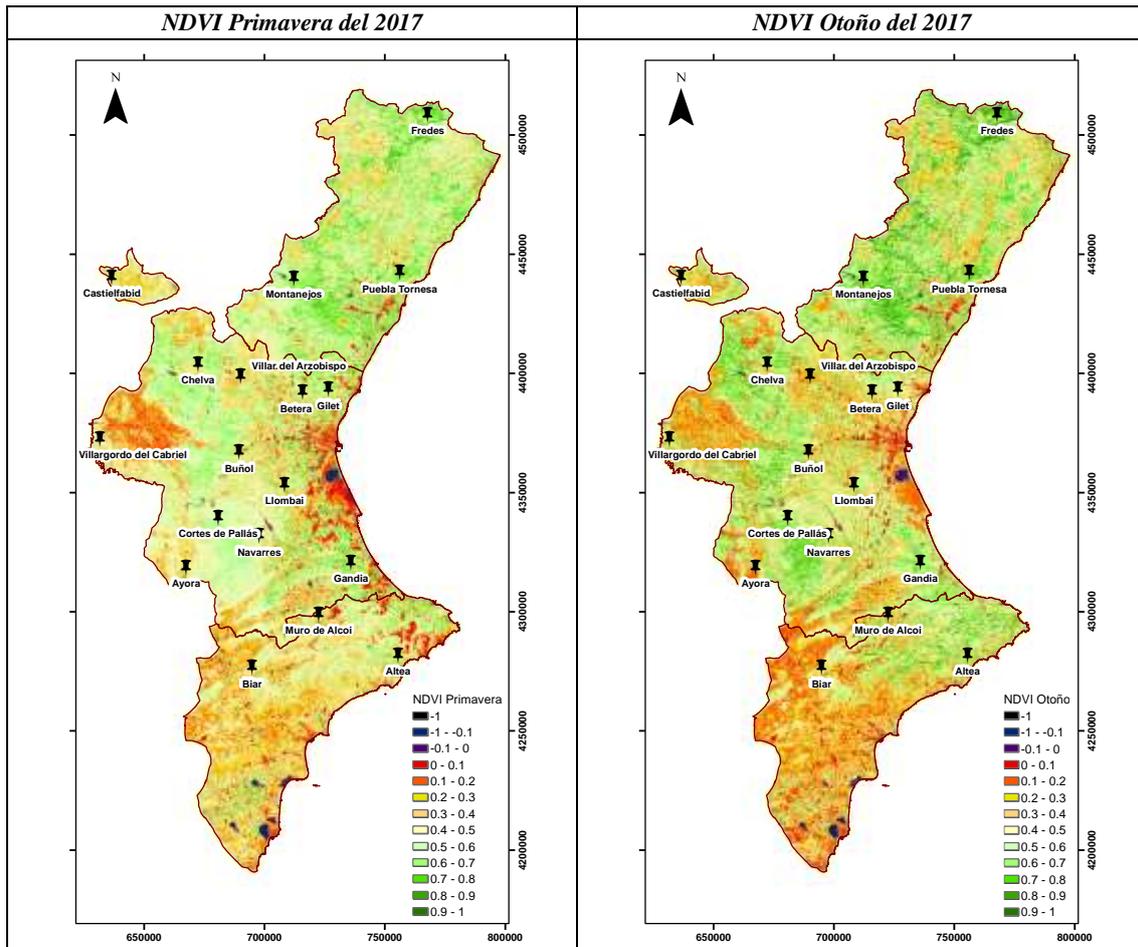


Imagen 13. NDVI obtenidos para el año 2017 en la Comunidad Valenciana. Composiciones de máximo valor para los meses de marzo (primavera) y octubre (otoño)

4.3. Datos meteorológicos.

La fenología de las especies forestales depende del estado meteorológico del ecosistema, obteniendo así el desarrollo de las plantas relacionado a los parámetros ambientales. Por tal motivo se realiza la recopilación de datos de variables meteorológicas, con el fin de relacionarlos con la humedad de combustible vivo y los índices de vegetación.

Para ellos se hace uso de los datos meteorológicos de las estaciones cercanas a los 18 puntos de muestreo, dichos datos se recopilan de las estaciones pertenecientes a entidades gestoras: Agencia Estatal de Meteorología (AEMET) y el Instituto Valenciano de Investigaciones Agrarias (IVIA). En la Tabla 6 se muestra las estaciones de donde se obtuvo la información para el presente estudio. Los datos de las estaciones fueron facilitados por VAERSA. En la Imagen 14 se observa la distribución de las estaciones meteorológicas en la Comunidad Valenciana y su cercanía a los puntos de muestreo.

Tabla 6. Estaciones meteorológicas.

PUNTO DE MUESTREO			ESTACION METEOROLÓGICA				
Nº	NOMBRE	Z	NOMBRE	ENTIDAD	X	Y	Z
1	Fredes	1021	La Pobra de Benifass	AEMET	767442	4509489	1168
2	Puebla Tornesa	371	Atzeneta del Maestrat	AEMET	740794	4456745	420
3	Montanejos	635	Arenos-Pantano	AEMET	709126	4440309	601
4	Castielfabid	1315	Ademuz	AEMET	646378	4435814	705
5	Chelva	721	Chelva	AEMET	671925	4402232	515
6	Villar del Arzobispo	480	Lliria	AEMET	701243	4393010	198
7	Betera	207	Betera	IVIA	717496	4385473	180
			Lliria	AEMET	701243	4393010	198
8	Gilet	302	Sagunto	AEMET	735386	4398825	179
9	Villargordo del Cabriel	857	Utiel	AEMET	650752	4382016	761
10	Buñol	467	Buñol	AEMET	684388	4368927	585
11	Llombai	139	Carlet CE Coop	IVIA	714049	4344831	180
			Bicorp	AEMET	691008	4333857	305
12	Cortes de Pallás	868	Jalance	AEMET	664520	4341554	372
13	Navarrés	220	Bicorp	AEMET	691008	4333857	305
14	Ayora	692	Jalance	AEMET	664520	4341554	372
15	Gandia	396	Gandia Marxuquera	IVIA	739781	4316502	180
			Ontinyet	AEMET	707647	4300498	350
16	Muro de Alcoi	477	Alcoy	AEMET	720841	4287680	530
17	Altea	301	Altea	IVIA	755865	4277073	175
			Alcoy	AEMET	720841	4287680	530
18	Biar	829	Camp de Mirra	IVIA	694426	4278052	690
			Alcoy	AEMET	720841	4287680	530

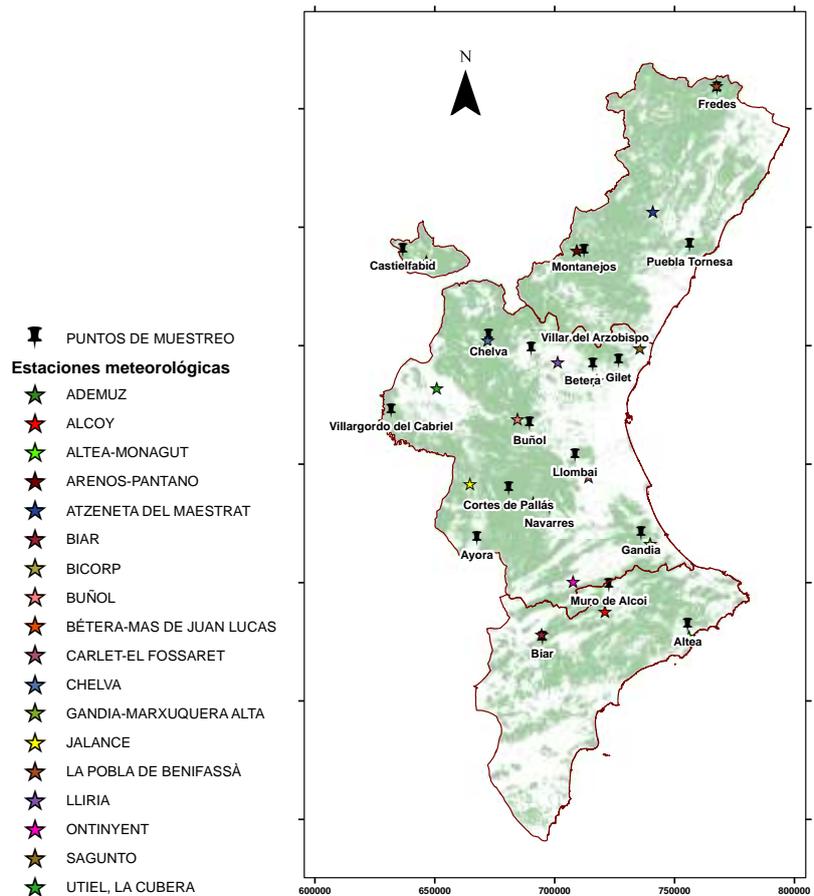


Imagen 14. Estaciones meteorológicas.

4.3.1. Cálculo de variables meteorológicas.

En el análisis de la influencia de las variables meteorológicas sobre la humedad del combustible vivo, es necesario incluir la mayor cantidad de variables con el fin de discernir cuales de ellas son las más influyentes en la variabilidad de la HCV en cada especie (Soriano Sancho, 2016).

Las estaciones meteorológicas antes mencionadas en la Tabla 6 entregan datos horarios, definiendo de ellas 28 variables meteorológicas agrupadas según el tipo de dato; precipitación, temperatura, humedad relativa, viento y punto del rocío, dichas variables se describen en la Tabla 7.

Para la elección de las variables se toma como referencia a (Soriano Sancho, 2016) el cual detalla la explicación de cada una como se muestra a continuación:

- *Precipitación*: supone el principal aporte de agua a las plantas y, por tanto, se intuye que debe tener influencia en la cantidad de agua que éstas absorben. La cantidad de lluvia también influye en la humedad del suelo, que en función de su capacidad de campo retendrá el agua distinto tiempo. En este estudio no se incluye la variable suelo, pero se analiza la precipitación hasta 60 días antes de la toma de muestra, con la idea de conocer la influencia que tiene la humedad acumulada en el suelo. Se mide en mm.
- *Temperatura máxima*: influye en la evapotranspiración de la vegetación y, a partir de ciertos umbrales, en la pérdida de humedad. Se mide en °C.
- *Temperatura media*: da una idea de la situación meteorológica de los días previos. Temperaturas medias altas indican mayores pérdidas de humedad.
- *Humedad relativa máxima*: es de interés relacionar las humedades más altas, pues son las que permiten un mayor intercambio de humedad de las plantas con el ambiente. Se mide en %.
- *Humedad relativa mínima*: a partir de ciertos umbrales mínimos de humedad, las plantas inician la pérdida de humedad por evapotranspiración, y si el proceso es duradero, pueden iniciar sistemas de protección como cerrar estomas.
- *Viento máximo*: a mayor velocidad del viento mayor evapotranspiración y pérdida de humedad si va asociado a altas temperaturas. Se mide en km/h.
- *Levante*: el viento de componente este en la Comunidad Valenciana proviene del mar, y, por lo tanto, va cargado de humedad. El análisis de esta variable pretende relacionar la humedad aportada por el viento de levante con la humedad absorbida por las plantas. Se mide en horas.
- *Punto de rocío*: se pretende analizar la influencia que tiene la condensación y la precipitación horizontal sobre la HCV de las especies. Como este valor no es cuantificable en mm de agua, se valora el número de días en que se produce condensación, entendiéndose como aquellos en los que la temperatura del punto de rocío alcanza la temperatura mínima del aire.

Tabla 7. Variables meteorológicas analizadas.

Fuente: (Soriano Sancho, 2016)

Tipo	Nº	Variable	Descripción
PRECIPITACIÓN	1	P3	precipitación acumulada en los últimos 3 días.
	2	P7	precipitación acumulada en los últimos 7 días.
	3	P15	precipitación acumulada en los últimos 15 días.
	4	P30	precipitación acumulada en los últimos 30 días.
	5	P60	precipitación acumulada en los últimos 60 días.
TEMPERATURA	6	Tmax3	media de las temperaturas máximas de los últimos 3 días
	7	Tmax7	media de las temperaturas máximas de los últimos 7 días
	8	Tmax15	media de las temperaturas máximas de los últimos 15 días
	9	Tmax30	media de las temperaturas máximas de los últimos 30 días
	10	Tmedia3	media de las temperaturas medias de los últimos 3 días
	11	Tmedia7	media de las temperaturas medias de los últimos 7 días
HUMEDAD RELATIVA	12	Tmedia15	media de las temperaturas medias de los últimos 15 días
	13	HRmax3	media de las HR máximas de los últimos 3 días
	14	HRmax7	media de las HR máximas de los últimos 7 días
	15	HRmax15	media de las HR máximas de los últimos 15 días
	16	HRmin3	media de las HR mínimas de los últimos 3 días
	17	HRmin7	media de las HR mínimas de los últimos 7 días
VIENTO	18	HRmin15	media de las HR mínimas de los últimos 15 días
	19	Vmax3	media de la velocidad máxima media del viento en 10' de los últimos 3 días
	20	Vmax7	media de la velocidad máxima media del viento en 10' de los últimos 7 días
	21	Vmax15	media de la velocidad máxima media del viento en 10' de los últimos 15 días
	22	Lev3	media de la suma de la frecuencia del viento en los octantes 2 y 3 (vientos de componente E), en los últimos 3 días
	23	Lev7	media de la suma de la frecuencia del viento en los octantes 2 y 3 (vientos de componente E), en los últimos 7 días
PUNTO DE ROCIO	24	Lev15	media de la suma de la frecuencia del viento en los octantes 2 y 3 (vientos de componente E), en los últimos 15 días
	25	Pto_r3	número de días en que se ha producido condensación en los últimos 3 días
	26	Pto_r7	número de días en que se ha producido condensación en los últimos 7 días
	27	Pto_r15	número de días en que se ha producido condensación en los últimos 15 días
	28	Pto_r30	número de días en que se ha producido condensación en los últimos 30 días

Hay que tener en cuenta que todas las variables se calculan desde el día anterior de la toma de muestra en el periodo de enero del 2016 a julio del 2018.

5. METODOLOGÍA

A continuación, se muestra en un esquema los procesos realizados desde la obtención de datos, los análisis estadísticos y la obtención de los resultados del presente trabajo.

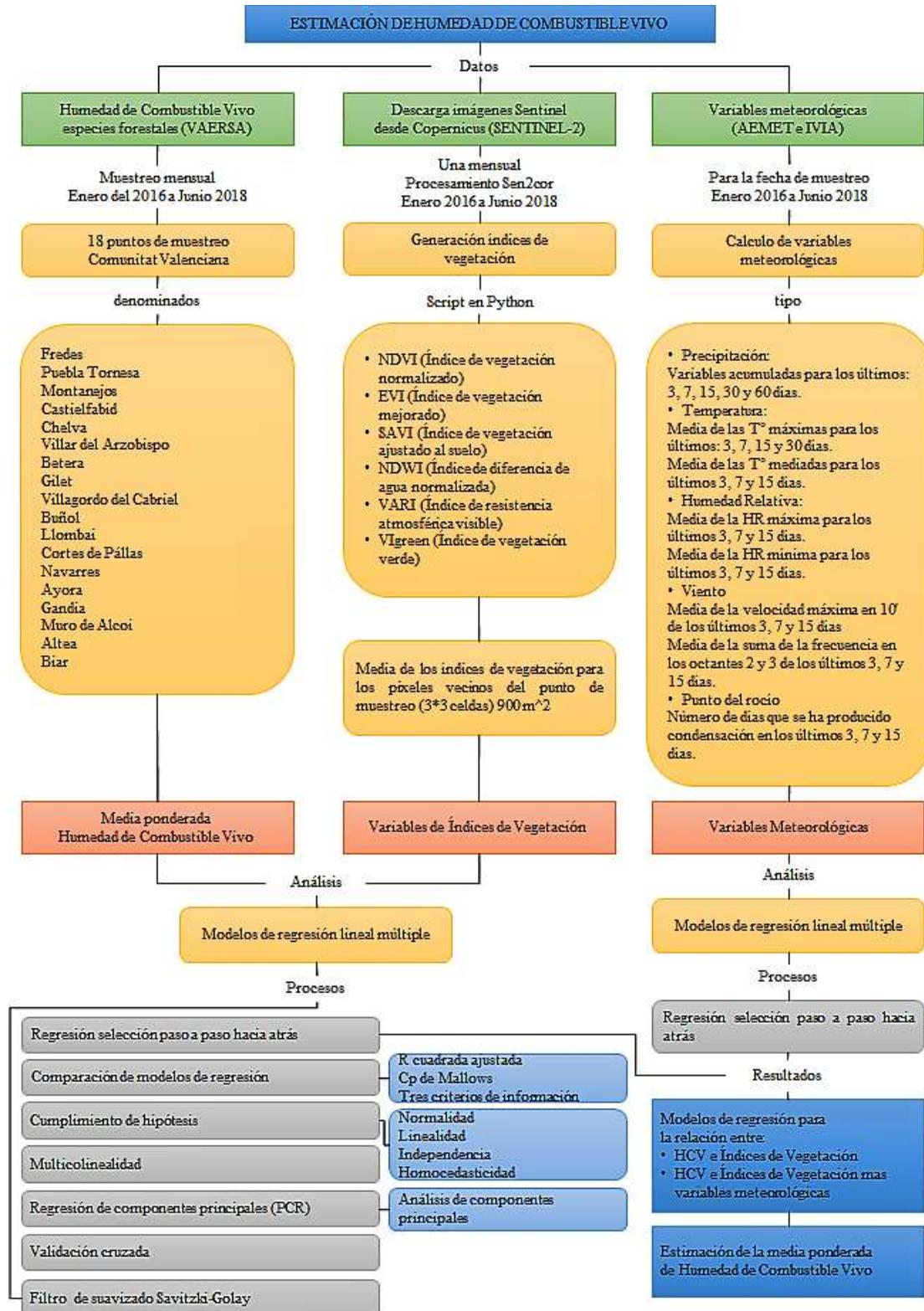


Imagen 15. Esquema de metodología de trabajo.

5.1. Análisis de modelos de regresión múltiple entre HCV e índices de vegetación.

En esta sección se explica la metodología para estimar los valores de HCV a lo largo del tiempo usando los perfiles temporales de los 6 índices de vegetación (NDVI, EVI, SAVI, NDWI, VARI, VIgreen) derivados de las imágenes satelitales Sentinel-2. Se utiliza la regresión múltiple para evaluar el potencial de cada índice de vegetación para la estimación de la HCV en cada punto de muestreo, analizando así el rendimiento de la información espectral derivada de Sentinel-2.

Las métricas de evaluación de los modelos generados para cada punto de muestreo incluyen los coeficientes de determinación (R^2), el error absoluto medio (MAE) y el error cuadrático medio (RMSE). El valor R^2 ajustado se utiliza para comparar la bondad de ajuste entre modelos con diferentes variables de entrada.

5.1.1. Método de regresión paso a paso hacia atrás usando la técnica del P-valor.

Los modelos de regresión de humedad de combustible vivo y los índices de vegetación obtenidos de Sentinel-2, han sido obtenidos mediante el software STATGRAPHICS Centurion esta herramienta permite el análisis estadístico de datos que combina procedimientos analíticos con gráficos interactivos proporcionando un entorno integrado de análisis que puede ser aplicado en cada una de las fases de un proyecto, desde los protocolos de gestión Six Sigma hasta los procesos de control de calidad.

STATGRAPHICS Centurion cuenta con funciones estadísticas avanzadas, capaces de proporcionar rigurosos análisis propios de los profesionales estadísticos más exigentes y experimentados, y al mismo tiempo ofrece un interface muy intuitivo, con funciones de asistencia exclusivas, de tal forma que proporciona la simplicidad suficiente para permitir a un analista inexperto realizar procedimientos complejos (Statgraphics, 2018).

A partir de los datos mensuales de dieciocho puntos de muestreo repartidos homogéneamente en la Comunidad Valenciana desde enero del 2016 a julio del 2018 se calcula el modelo de regresión múltiple para cada uno de los puntos de muestreo considerando como variable dependiente la humedad de combustible vivo ($Y = HCV_P$) y como variables independientes los índices de vegetación ($X1=NDVI$, $X2=EVI$, $X3=SAVI$, $X4=NDWI$, $X5=VARI$, $X6=VIgreen$) y sus respectivos polinomios de segundo grado con el fin de mejorar los modelos ($X7=NDVI^2$, $X8=EVI^2$, $X9=SAVI^2$, $X10=NDWI^2$, $X11=VARI^2$, $X12=VIgreen^2$), la variable cuadrática hace que el modelo mejore ya que su valor-P mejora, mostrando que su variable es estadísticamente significativa.

Para la selección de variables el software permite escoger cinco procedimientos de ajuste, en este caso se tomó como procedimiento la selección paso a paso hacia atrás; este comienza con un modelo que involucra todas las variables y elimina del modelo la variable que es estadísticamente menos significativa, para ello se puede elegir entre el criterio de la F o el criterio del Valor-P. En este trabajo se ha elegido el criterio del valor-

P, considerando que las variables que se tienen que eliminar en cada paso del método tendrá un Valor-P inferior a 0.05. El proceso elimina una variable en cada paso hasta que los valores de P de todas las variables que quedan son inferiores a 0.05 en este caso, aunque variables eliminadas previamente pueden ser introducidas de nuevo en el modelo en pasos posteriores.

REGRESION MULTIPLE DEL PUNTO DE MUESTREO CASTIELFABID

La aplicación de la metodología se mostrará en un solo punto de muestreo, aunque posteriormente se analizarán los resultados obtenidos en todos ellos. El punto de muestreo elegido para mostrar el procedimiento completo ha sido Castielfabid en el cual se toma 12 variables independientes (NDVI, EVI, SAVI, NDWI, VARI, VIgreen, NDVI², EVI², SAVI², NDWI², VARI², VIgreen²) de las que el método de regresión Paso a Paso hacia atrás elige 5 variables (NDVI, EVI, NDWI, VARI y NDWI²) obteniendo valores-P en la tabla ANOVA menores a 0.05 existiendo una relación estadísticamente significativa entre las variables con un nivel de confianza del 95%. La Tabla 8 muestra los valores de las estimaciones de cada una de las variables, así como el error estándar, el estadístico T-student y sus valores P, la regresión múltiple para este punto de muestreo obtiene un R-cuadrado ajustado de 57.3664%.

Tabla 8. Regresión múltiple HCV_P.

Parámetro	Estimación	Error		Valor-P
		Estándar	Estadístico T	
CONSTANTE	-453.786	113.82	-3.98687	0.0005
NDVI	-1465.8	335.2	-4.3729	0.0002
EVI	-42.6144	12.084	-3.52652	0.0017
NDWI	-4592.2	896.702	-5.1212	0.0000
VARI	1504.53	272.641	5.51834	0.0000
NDWI ²	-3798.49	793.44	-4.78737	0.0001

Análisis de Varianza

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	6054.07	5	1210.81	8.80	0.0001
Residuo	3300.6	24	137.525		
Total (Corr.)	9354.67	29			

R-cuadrada = 64.717 por ciento

R-cuadrado (ajustado para g.l.) = 57.3664 por ciento

Error estándar del est. = 11.7271

Error absoluto medio = 8.09526

Estadístico Durbin-Watson = 2.04858 (P=0.3628)

Autocorrelación de residuos en retraso 1 = -0.0623267

La salida muestra los resultados de ajustar un modelo de regresión múltiple para describir la relación entre HCV_P y 5 variables independientes. La ecuación del modelo ajustado es:

$$\text{HCV_P} = -453.786 - 1465.8 \cdot \text{NDVI} - 42.6144 \cdot \text{EVI} - 4592.2 \cdot \text{NDWI} + 1504.53 \cdot \text{VARI} - 3798.49 \cdot \text{NDWI}^2 \quad (7)$$

El estadístico R-Cuadrada indica que el modelo así ajustado explica 64.717% de la variabilidad en HCV_P. El estadístico R-Cuadrada ajustada, que es más apropiada para comparar modelos con diferente número de variables independientes, es 57.3664%. El error estándar del estimado muestra que la desviación estándar de los residuos es 11.7271. Este valor puede usarse para construir límites para nuevas observaciones. El error absoluto medio (MAE) de 8.09526 es el valor promedio de los residuos.

En la Imagen 16 se puede verificar el ajuste del modelo calculado a partir de la regresión lineal múltiple en función de la fecha de obtención del HCV.

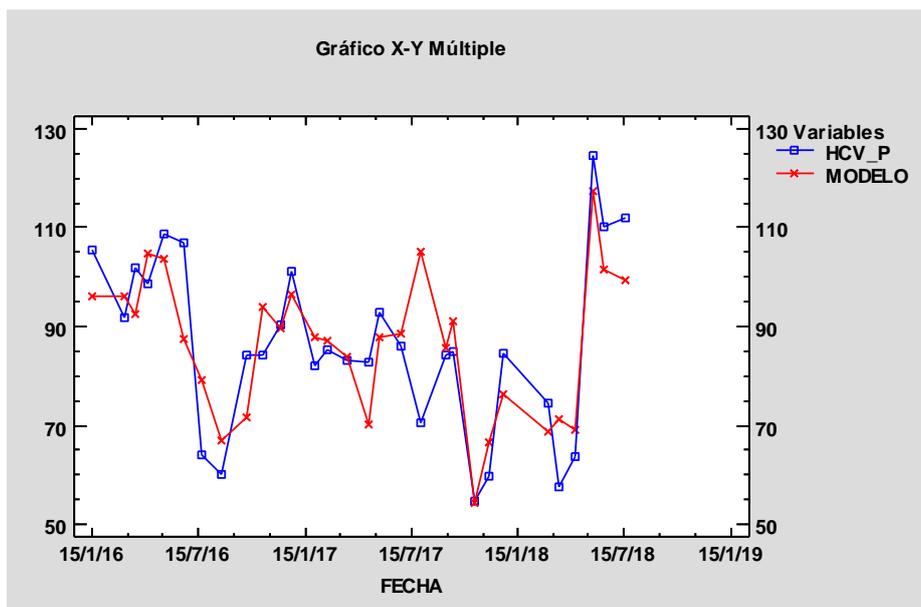


Imagen 16. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Castielfabid.

La Tabla 21 de resultados muestra los modelos para los 17 puntos de muestreo restantes.

5.1.2. Comparación de modelos de regresión con diferente número de variables.

El objetivo de este apartado es comparar modelos de regresión con diferente número de variables usando diferentes indicadores: R-cuadrada ajustada por los grados de libertad, Cp de Mallows y los tres criterios de información.

La Tabla 9 muestra los modelos que dan los valores más altos de R-cuadrada ajustada, este valor mide la proporción de variabilidad en la humedad de combustible vivo que es explicada por el modelo. Los valores grandes de R-cuadrada ajustada corresponden a valores pequeños de cuadrado medio del error (CME), según este estadístico el mejor modelo contiene cinco variables: NDVI, EVI, NDWI, VARI y NDWI². (ABDEJ), mismas que fueron obtenidas mediante la regresión paso a paso hacia atrás como se mostró en el apartado 5.1.1, esto se corrobora en la Imagen 17 de R-cuadrada que se muestra a continuación.

Selección del Modelo de Regresión - HCV_P

Variable dependiente: HCV_P

Variables independientes:

- A=NDVI
- B=EVI
- C=SAVI
- D=NDWI
- E=VARI
- F=VIgreen
- G=NDVI²
- H=EVI²
- I=SAVI²
- J=NDWI²
- K=VARI²
- L=VIgreen²

Tabla 9. Modelos con Mayor R-Cuadrada Ajustada

		<i>R-Cuadrada</i>		<i>Variables</i>
<i>CME</i>	<i>R-Cuadrada</i>	<i>Ajustada</i>	<i>Cp</i>	<i>Incluidas</i>
137.525	64.717	57.3664	1.57995	ABDEJ
138.322	68.9485	57.1193	5.23175	ABCDGFKL
138.789	68.8437	56.9747	5.28988	ACDFGHKL
139.452	65.7134	56.769	3.02705	ABCDEJ
139.453	65.7131	56.7687	3.02721	ABDEIJ
139.463	68.6923	56.7656	5.37391	ABDFGIKL
139.477	65.7073	56.7614	3.03043	ABDEFJ
139.478	65.707	56.761	3.03059	ABDEJL
139.578	68.6667	56.7302	5.38813	ABCDFJKL
140.05	67.0635	56.5837	4.27779	ABDEHJL
140.061	67.061	56.5804	4.27921	ABCDEHJ
140.065	67.06	56.5791	4.27973	ABDEHIJ
140.329	66.9979	56.4972	4.31422	ABDEFHJ
140.521	68.4549	56.4377	5.50568	ADFGHIKL
140.755	66.8977	56.3651	4.36982	ABCDEFHG
142.012	63.5658	55.9754	2.21881	ADEHJ
143.259	64.7774	55.5889	3.54646	ABDEJK
144.073	69.1977	55.3366	7.09347	ACDFGHKL
144.327	69.1433	55.2578	7.12365	ABCDGFKL
144.37	62.9608	55.2444	2.55453	BDEGJ
144.679	69.0681	55.1488	7.16535	ABCDGFKL
144.877	69.0257	55.0873	7.18889	ABCDGFKL
144.979	69.0039	55.0557	7.20097	ABCDGFKL
145.217	62.7436	54.9819	2.67507	DEGHJ
149.247	61.7097	53.7325	3.24887	BDEJL
150.909	69.3493	53.2174	9.00931	ABCDEFGJKL
151.526	69.2239	53.026	9.07891	ACDFGHJKL
151.528	69.2236	53.0255	9.07907	ABCDEFGHKL
151.645	69.1999	52.9893	9.09225	ACDFGHIKL
151.811	69.1661	52.9377	9.111	ABCDGFIJKL
159.249	69.3578	50.6321	11.0046	ABCDEFGIJKL
159.25	69.3575	50.6316	11.0047	ABCDEFGHJKL
159.921	69.2284	50.4235	11.0764	ACDFGHIJKL
159.937	69.2254	50.4186	11.0781	ABCDEFGHIKL
160.213	69.1723	50.3332	11.1075	ABCDFGHIJKL

CME	R-Cuadrada		Cp	Variables Incluidas
	R-Cuadrada	Ajustada		
168.571	69.3661	47.7422	13.0	ABCDEFGHIJKL
169.052	54.8216	47.593	5.07136	DEGH
175.511	53.0954	45.5906	6.0293	DGHK
186.322	50.2061	42.2391	7.63266	DFGK
186.344	50.2002	42.2322	7.63594	CDGK
186.346	50.1997	42.2316	7.63623	DGIK
216.412	39.8512	32.9109	11.379	AGK
225.236	37.3988	30.1756	12.74	AEG
228.646	36.4511	29.1185	13.2659	DJK
231.309	35.711	28.2931	13.6766	DGK
233.459	35.1134	27.6265	14.0082	DEJ
245.765	29.066	23.8116	15.3642	DK
251.552	27.3956	22.0175	16.2911	JK
252.721	27.0583	21.6552	16.4783	AK
255.628	26.2192	20.754	16.944	DE
258.794	25.3053	19.7724	17.4511	GK
271.615	18.7014	15.7979	19.1159	K
278.041	16.7781	13.8059	20.1832	E
308.076	7.78794	4.49466	25.1722	A
311.269	6.83238	3.50496	25.7025	G
322.575	3.44828	0.0	29.4195	L
322.575	0.0	0.0	27.4941	

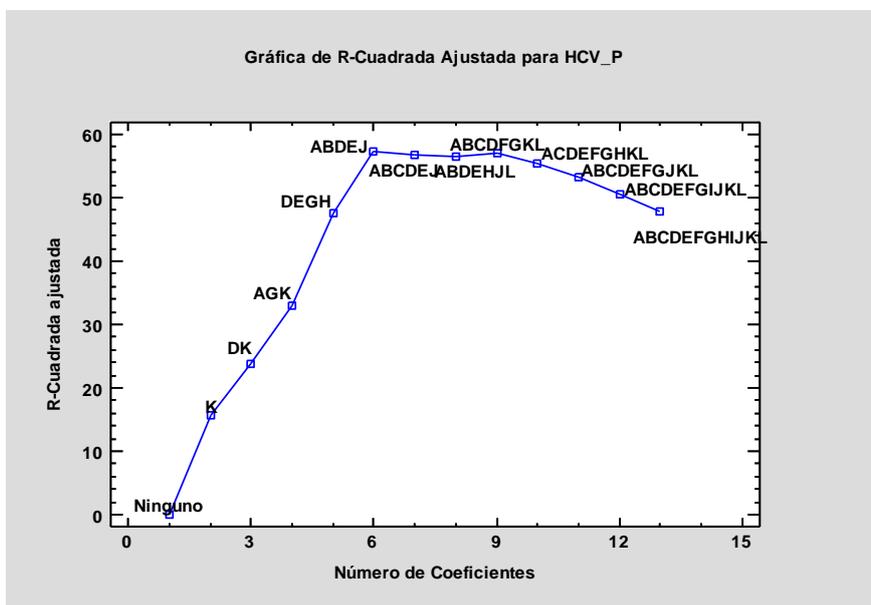


Imagen 17. Gráfica de R-Cuadrada Ajustada para HCV_P en Castielfabid.

La Tabla 10 muestra los modelos que dan los valores más pequeños del estadístico Cp de Mallows. Cp es una medida del sesgo en el modelo, basada en la comparación entre el cuadrado medio del error total y la varianza del error verdadero. La Imagen 18 de Cp contiene una línea igual a P para ayudar a seleccionar los mejores modelos. La Cp muestra el valor más cercano al número de variables en este caso selecciona el modelo (ABDEJ).

Tabla 10. Modelos con Menor Cp

CME	R-Cuadrada		Cp	Variables Incluidas
	R-Cuadrada	Ajustada		
137.525	64.717	57.3664	1.57995	ABDEJ
142.012	63.5658	55.9754	2.21881	ADEHJ
144.37	62.9608	55.2444	2.55453	BDEGJ
145.217	62.7436	54.9819	2.67507	DEGHJ
139.452	65.7134	56.769	3.02705	ABCDEJ
139.453	65.7131	56.7687	3.02721	ABDEIJ
139.477	65.7073	56.7614	3.03043	ABDEFJ
139.478	65.707	56.761	3.03059	ABDEJL
149.247	61.7097	53.7325	3.24887	BDEJL
143.259	64.7774	55.5889	3.54646	ABDEJK
140.05	67.0635	56.5837	4.27779	ABDEHJL
140.061	67.061	56.5804	4.27921	ABCDEHJ
140.065	67.06	56.5791	4.27973	ABDEHIJ
140.329	66.9979	56.4972	4.31422	ABDEFHJ
140.755	66.8977	56.3651	4.36982	ABCDEFGF
169.052	54.8216	47.593	5.07136	DEGH
138.322	68.9485	57.1193	5.23175	ABCDFGKL
138.789	68.8437	56.9747	5.28988	ACDFGHKL
139.463	68.6923	56.7656	5.37391	ABDFGIKL
139.578	68.6667	56.7302	5.38813	ABCDFJKL
140.521	68.4549	56.4377	5.50568	ADFGHIKL
175.511	53.0954	45.5906	6.0293	DGHK
144.073	69.1977	55.3366	7.09347	ACDFGHKL
144.327	69.1433	55.2578	7.12365	ABCDFGJKL
144.679	69.0681	55.1488	7.16535	ABCDEFJKL
144.877	69.0257	55.0873	7.18889	ABCDFGKL
144.979	69.0039	55.0557	7.20097	ABCDFGHKL
186.322	50.2061	42.2391	7.63266	DFGK
186.344	50.2002	42.2322	7.63594	CDGK
186.346	50.1997	42.2316	7.63623	DGIK
150.909	69.3493	53.2174	9.00931	ABCDEFGJKL
151.526	69.2239	53.026	9.07891	ACDFGHJKL
151.528	69.2236	53.0255	9.07907	ABCDFGHKL
151.645	69.1999	52.9893	9.09225	ACDFGHIKL
151.811	69.1661	52.9377	9.111	ABCDFGIJKL
159.249	69.3578	50.6321	11.0046	ABCDFGIJKL
159.25	69.3575	50.6316	11.0047	ABCDFGHJKL
159.921	69.2284	50.4235	11.0764	ACDFGHIIJKL
159.937	69.2254	50.4186	11.0781	ABCDFGHIKL
160.213	69.1723	50.3332	11.1075	ABCDFGHIJKL
216.412	39.8512	32.9109	11.379	AGK
225.236	37.3988	30.1756	12.74	AEG
168.571	69.3661	47.7422	13.0	ABCDFGHIJKL
228.646	36.4511	29.1185	13.2659	DJK
231.309	35.711	28.2931	13.6766	DGK
233.459	35.1134	27.6265	14.0082	DEJ
245.765	29.066	23.8116	15.3642	DK
251.552	27.3956	22.0175	16.2911	JK
252.721	27.0583	21.6552	16.4783	AK
255.628	26.2192	20.754	16.944	DE

Tabla 11. Modelos con el mejor criterio de información

MSE	Coefficientes	AIC	HQC	SBIC	Variables incluidas
137.525	6	5.32381	5.41346	5.60405	ABDEJ
142.012	6	5.35591	5.44557	5.63615	ADEHJ
144.37	6	5.37238	5.46203	5.65262	BDEGJ
145.217	6	5.37823	5.46788	5.65847	DEGHJ
139.452	7	5.40439	5.50898	5.73133	ABCDEJ
139.453	7	5.4044	5.50899	5.73134	ABDEIJ
139.477	7	5.40457	5.50916	5.73151	ABDEFJ
139.478	7	5.40458	5.50917	5.73152	ABDEJL
149.247	6	5.4056	5.49526	5.68584	BDEJL
143.259	7	5.43132	5.53591	5.75827	ABDEJK
169.052	5	5.46354	5.53825	5.69707	DEGH
140.05	8	5.47533	5.59487	5.84898	ABDEHJL
140.061	8	5.47541	5.59494	5.84906	ABCDEHJ
140.065	8	5.47544	5.59497	5.84909	ABDEHIJ
140.329	8	5.47732	5.59686	5.85098	ABDEFHJ
140.755	8	5.48036	5.59989	5.85401	ABCDEFGF
175.511	5	5.50104	5.57574	5.73457	DGHK
138.322	9	5.52959	5.66406	5.94995	ABCDFGKL
138.789	9	5.53295	5.66743	5.95331	ACDFGHKL
139.463	9	5.5378	5.67228	5.95816	ABDFGIKL
139.578	9	5.53862	5.6731	5.95898	ABCDFJKL
140.521	9	5.54536	5.67983	5.96572	ADFGHIKL
186.322	5	5.56081	5.63552	5.79434	DFGK
186.344	5	5.56093	5.63564	5.79446	CDGK
186.346	5	5.56094	5.63565	5.79447	DGIK
144.073	10	5.63699	5.7864	6.10405	ACDFGHKL
144.327	10	5.63875	5.78817	6.10582	ABCDFGJKL
144.679	10	5.64118	5.7906	6.10825	ABCDEFJKL
144.877	10	5.64255	5.79197	6.10962	ABCDFGKL
144.979	10	5.64326	5.79267	6.11032	ABCDFGHKL
216.412	4	5.64385	5.70362	5.83068	AGK
225.236	4	5.68382	5.74358	5.87064	AEG
228.646	4	5.69884	5.75861	5.88567	DJK
245.765	3	5.70437	5.7492	5.84449	DK
231.309	4	5.71042	5.77019	5.89725	DGK
233.459	4	5.71967	5.77944	5.9065	DEJ
251.552	3	5.72765	5.77247	5.86777	JK
252.721	3	5.73228	5.77711	5.8724	AK
271.615	2	5.73772	5.7676	5.83113	K
255.628	3	5.74372	5.78855	5.88384	DE
150.909	11	5.75001	5.91437	6.26378	ABCDEFGJKL
151.526	11	5.75409	5.91845	6.26787	ACDFGHJKL
151.528	11	5.7541	5.91846	6.26788	ABCDFGHKL
151.645	11	5.75487	5.91923	6.26865	ACDFGHKL
151.811	11	5.75597	5.92033	6.26974	ABCDFGIJKL
258.794	3	5.75603	5.80086	5.89615	GK
278.041	2	5.7611	5.79098	5.85451	E
322.575	1	5.843	5.85794	5.88971	
308.076	2	5.86368	5.89356	5.95709	A
159.249	12	5.87047	6.04977	6.43094	ABCDEFGIJKL
159.25	12	5.87048	6.04978	6.43095	ABCDFGHJKL

MSE	Coefficientes	AIC	HQC	SBIC	Variables incluidas
311.269	2	5.87399	5.90387	5.9674	G
159.921	12	5.87468	6.05398	6.43516	ACDEFGHIJKL
159.937	12	5.87478	6.05408	6.43526	ACDEFGHIJKL
160.213	12	5.8765	6.0558	6.43698	ACDEFGHIJKL
322.575	2	5.92832	5.9582	6.02173	D
168.571	13	5.99402	6.18827	6.60121	ACDEFGHIJKL

5.1.3. Análisis del cumplimiento de las hipótesis del modelo de regresión: normalidad, linealidad, homocedasticidad e independencia.

En el modelo para el punto de muestreo Castielfabid se rechaza la idea de que los residuos provienen de una distribución normal con 95% de confianza ya que su valor-P más pequeño de las pruebas realizadas es menor a 0.05, como se muestra en la Tabla 12.

Tabla 12. Pruebas de Normalidad para Residuos

Prueba	Estadístico	Valor-P
Chi-Cuadrado	14.0	0.300708
Estadístico W de Shapiro-Wilk	0.907248	0.0135977
Valor-Z para asimetría	2.06097	0.0393059
Valor-Z para curtosis	2.90532	0.00366888

En otras palabras, la hipótesis nula o H_0 es que la variable RESIDUOS posee una distribución normal y H_1 , o hipótesis alternativa que no siga una distribución normal. Puesto que el valor-P es menor que 0,05 se rechaza H_0 . El test Chi-Cuadrado indica que los residuos siguen una distribución normal. Pero otro test como el Estadístico W de Shapiro-Wilk y el de Valor-Z para curtosis indican lo contrario.

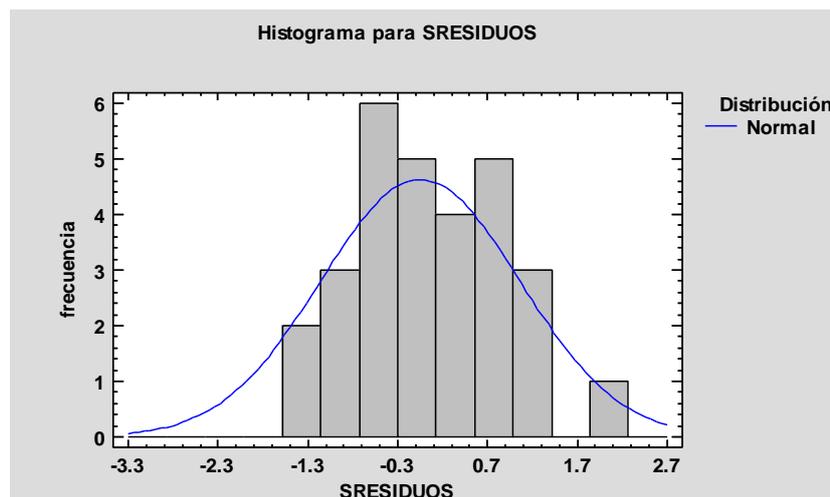


Imagen 19. Histograma para residuos de la distribución normal.

En la Imagen 20 indica que la distribución de los residuos es más apuntada que la de la curva normal, también se observa que el modelo cumple con la hipótesis de linealidad, pero la homocedasticidad es discutible ya que la varianza no es la misma para todos los

valores predichos, regularmente queda en evidencia mediante un patrón de forma en la gráfica de residuos; la linealidad se pierde en los valores más altos de la Y.

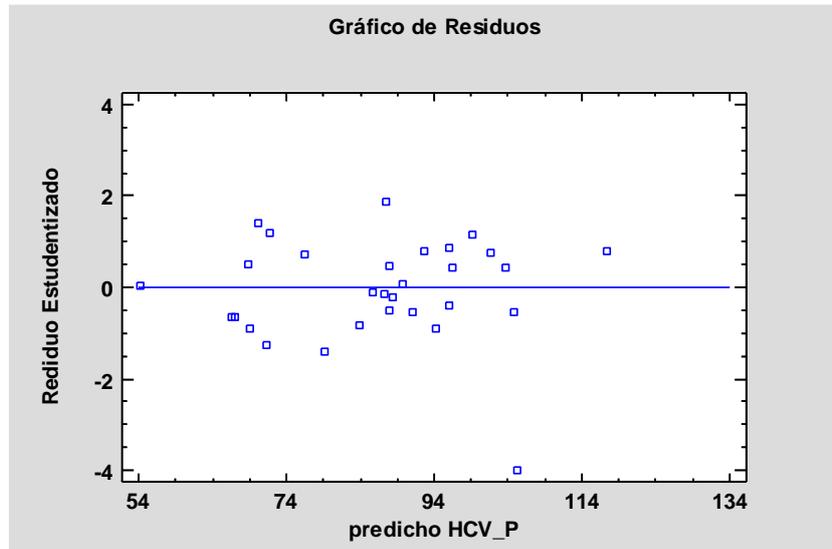


Imagen 20. Gráfica de residuos estudentizado vs predicho.

En análisis de la independencia de los residuos se utiliza el estadístico de Durbin-Watson (DW) el cual examina los residuos para determinar si hay alguna correlación significativa basada en el orden en el que se presentan en el archivo de datos. La Imagen 21 indica que hay independencia puesto que la nube de puntos no muestra ninguna tendencia alrededor de la recta $y=0$; el estadístico de Durbin Watson es 2.04858 ($P=0.3628$), puesto que el valor-P es mayor que 0.05, no hay indicación de una autocorrelación serial en los residuos con un nivel de confianza del 95.0%.

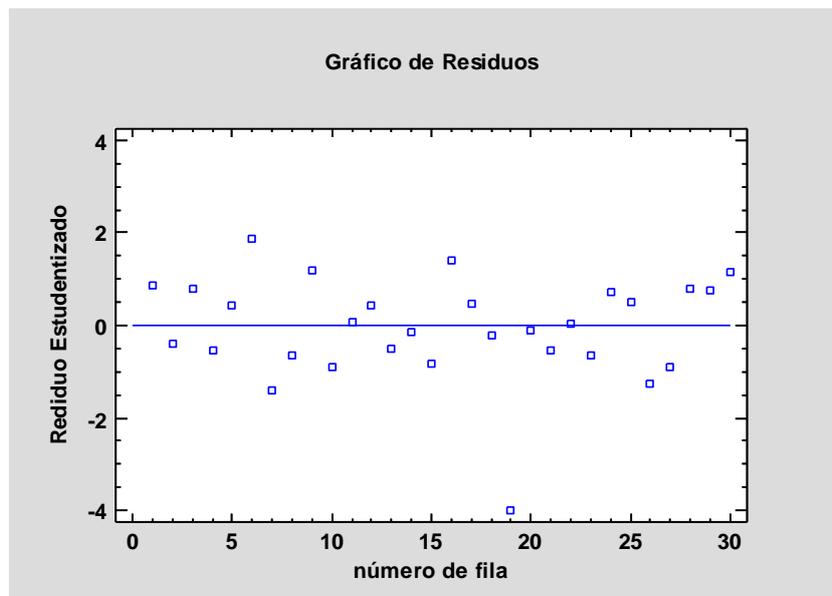


Imagen 21. Gráfica de residuos estudentizado vs número de fila.

5.1.4. Análisis de Multicolinealidad.

El caso de la linealidad perfecta es muy poco frecuente en los casos prácticos, este se presenta cuando en un modelo de regresión múltiple de dos o más variables explicativas mantienen una relación lineal. Para detectar la multicolinealidad es necesario examinar la correlación entre cada par de variables explicativas.

A partir del modelo anteriormente realizado en donde se obtuvo un R-cuadrada ajusta de 0.574 en donde todas las variables de regresión individuales son estadísticamente significativas, se obtiene la multicolinealidad mediante el desarrollo de un proceso en el lenguaje de programación para análisis estadísticos R dicho código es adjunto en el ANEXO 2.

El diagnostico se realiza a partir de la obtención de la correlación por pares de variables explicativas obtenidas en el modelo de Castielfabid.

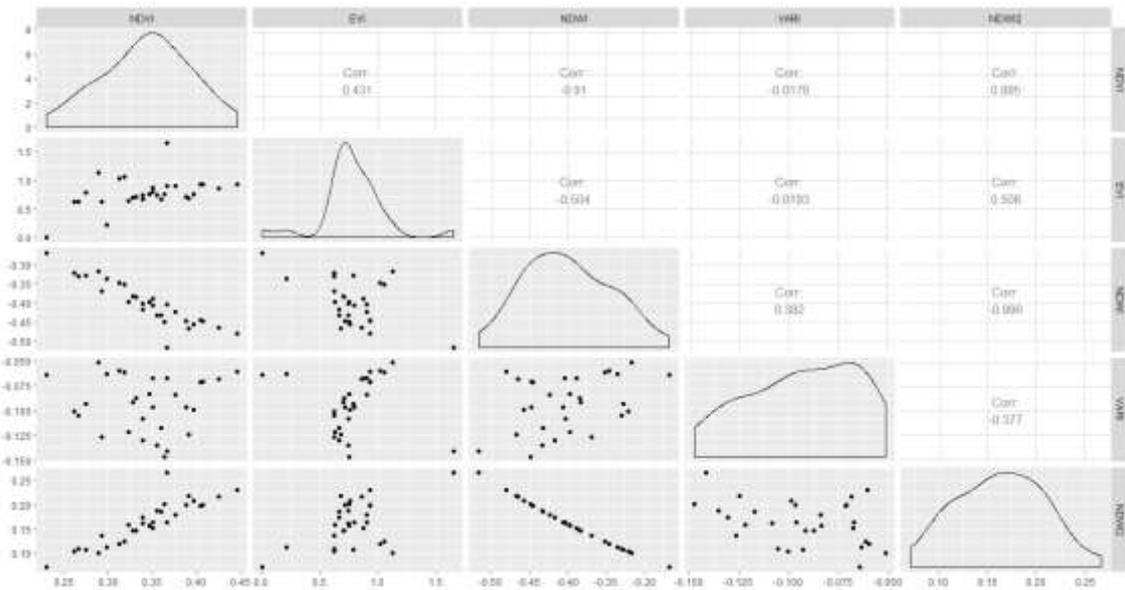


Imagen 22. Matriz de correlación.

La matriz de correlación en la Imagen 22 muestra que las correlaciones por pares entre las variables explicativas no son muy altas exceptuando los pares: NDWI-NDVI, NDWI²-NDVI y la que da correlación más alta es NDWI²-NDWI con una correlación de -0.996. La correlación entre estas dos variables puede ser la causa de multicolinealidad.

Se determina la matriz de covarianzas como se muestra en la Tabla 13, con el fin de comprender si la relación es positiva o negativa entre las variables, siendo un complemento a la matriz de correlación.

Tabla 13. Matriz de covarianzas

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	1.0000000	-0.6270097	-0.8727242	0.9471533	-0.7137295
[2,]	-0.6270097	1.0000000	-0.5421060	0.6475719	-0.4074181
[3,]	-0.8727242	-0.5421060	1.0000000	0.8368839	-0.9627230
[4,]	0.9471533	0.6475719	0.8368839	1.0000000	0.6800251
[5,]	-0.7137295	-0.4074181	-0.9627230	0.6800251	1.0000000

Se realizar la comprobación de diagnóstico global e individual de las variables para la multicolinealidad respectivamente.

Overall Multicollinearity Diagnostics

	MC Results	detection
Determinant X'X :	0.0001	1
Farrar Chi-Square:	247.5396	1
Red Indicator:	0.6007	1
Sum of Lambda Inverse:	906.5102	1
Theil's Method:	1.8770	1
Condition Number:	465.1828	1

1 --> COLLINEARITY is detected by the test
0 --> COLLINEARITY is not detected by the test

El valor del determinante estandarizado es 0.0001 siendo está muy pequeña, el valor de la prueba estadística chi-cuadrada es 247, 5396 implicando este valor la presencia de multicolinealidad en la especificación del modelo.

Se procede a ubicar la multicolinealidad siguiendo la prueba de Glauber (F – test), obtenido así individualmente la misma.

All Individual Multicollinearity Diagnostics Result

	VIF	TOL	Wi	Fi Leamer	CVIF Klein
NDVI	62.1407	0.0161	382.1296	529.8864	0.1269
EVI	2.3310	0.4290	8.3184	11.5349	0.6550
NDWI	558.3972	0.0018	3483.7324	4830.7756	0.0423
VARI	11.9395	0.0838	68.3717	94.8087	0.2894
NDWI^2	271.7019	0.0037	1691.8868	2346.0831	0.0607

1 --> COLLINEARITY is detected by the test
0 --> COLLINEARITY is not detected by the test

* all coefficients have significant t-ratios

R-square of y on all x: 0.6472

Las columnas VIF, TOL y Wi proporcionan la salida de diagnóstico para el factor de inflación de la variancia, la tolerancia y prueba F respectivamente. El estadístico F para NDWI es bastante alto (558.3972) seguido por NDWI^2, a más de los grados de libertad serán la causa de la multicolinealidad, como se observó en la matriz de correlación.

Se examina el patrón de multicolinealidad, para ello se lleva a cabo la prueba t para el coeficiente de correlación, en los segundos datos se observa que el valor-P más grande es el perteneciente a $NDWI^2$, además el coeficiente de dicha variable en el modelo de regresión podrá variar muy poco en la muestra de entrenamiento del modelo.

```
$`estimate`
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 1.0000000 -0.6270097 -0.8727242 0.9471533 -0.7137295
[2,] -0.6270097 1.0000000 -0.5421060 0.6475719 -0.4074181
[3,] -0.8727242 -0.5421060 1.0000000 0.8368839 -0.9627230
[4,] 0.9471533 0.6475719 0.8368839 1.0000000 0.6800251
[5,] -0.7137295 -0.4074181 -0.9627230 0.6800251 1.0000000

$p.value
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 0.000000e+00 0.0004653458 2.938255e-09 7.509436e-14 2.915393e-05
[2,] 4.653458e-04 0.0000000000 3.489491e-03 2.607233e-04 3.491611e-02
[3,] 2.938255e-09 0.0034894912 0.000000e+00 5.327791e-08 1.041391e-15
[4,] 7.509436e-14 0.0002607233 5.327791e-08 0.000000e+00 9.536214e-05
[5,] 2.915393e-05 0.0349161061 1.041391e-15 9.536214e-05 0.000000e+00

$statistic
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 0.000000 -4.024394 -8.937937 14.763224 -5.095000
[2,] -4.024394 0.000000 -3.225631 4.249136 -2.230615
[3,] -8.937937 -3.225631 0.000000 7.644446 -17.795960
[4,] 14.763224 4.249136 7.644446 0.000000 4.637448
[5,] -5.095000 -2.230615 -17.795960 4.637448 0.000000

$n
[1] 30

$gp
[1] 3

$method
[1] "pearson"
```

A continuación, se analiza la regresión múltiple para las variables de índices de vegetación (NDVI, EVI, SAVI, VARI, NDWI, VIgreen) sin tomar en cuenta las variables polinomiales, se puede observar que el R-cuadrada ajustada disminuye a 0.2941 y toma las 4 variables que aportan significativamente al modelo (NDVI, SAVI, VARI, NDWI) obteniendo valores de Pr cercanos a 0.05.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-30491	13491	-2.260	0.0328	*
NDVI	-4115	1787	-2.303	0.0299	*
SAVI	20399	9002	2.266	0.0324	*
NDWI	-4206	1783	-2.359	0.0264	*
VARI	2878	1103	2.609	0.0151	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.09 on 25 degrees of freedom

Multiple R-squared: 0.3914, Adjusted R-squared: 0.2941
F-statistic: 4.02 on 4 and 25 DF, p-value: 0.01186

Se realiza el diagnóstico para la obtención de la correlación por pares de variables explicativas obtenidas en el reciente modelo. En la Imagen 23 se muestra las correlaciones, en este caso no son muy altas exceptuando el par: NDWI-NDVI con una correlación de -0.91. La correlación entre estas dos variables puede ser la causa de multicolinealidad en el presente modelo.

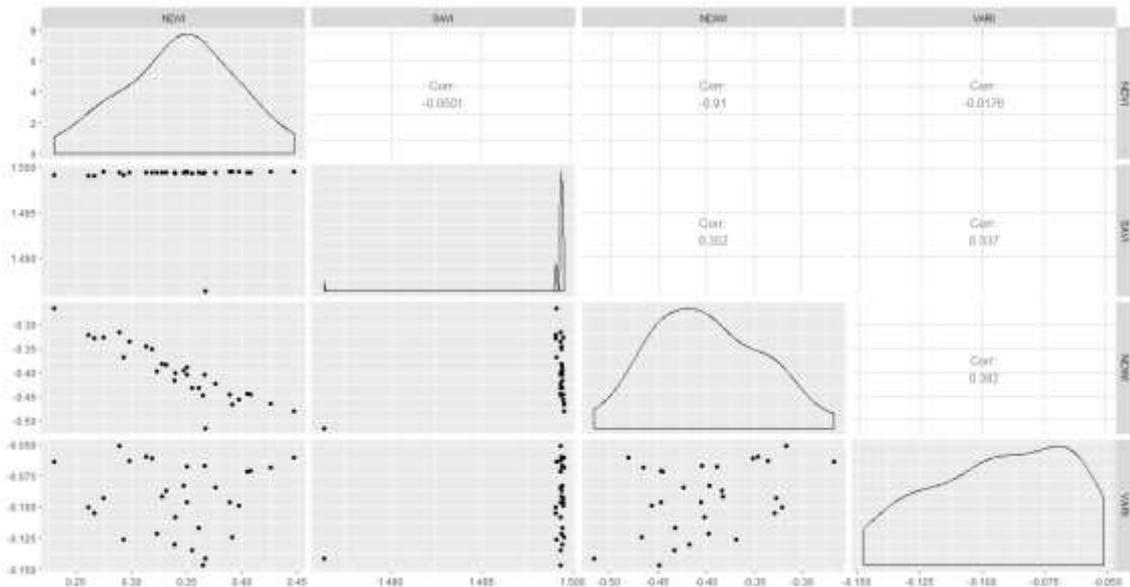


Imagen 23. Matriz de correlación sin variables polinomiales.

Se realiza el análisis de VIF, TOL y Wi para el nuevo modelo proporcionando la salida de diagnóstico para el factor de inflación de la variancia, la tolerancia y prueba F respectivamente. El estadístico F para NDWI es bastante alto (1333.38), a más de los grados de libertad serán la causa de la multicolinealidad, como se observó en la matriz de correlación.

All Individual Multicollinearity Diagnostics Result

	VIF	TOL	Wi	Fi	Leamer	CVIF	Klein
NDVI	1066.9553	0.0009	9238.2793	14390.397	0.0306	882.9268	1
SAVI	58.3944	0.0171	497.4185	774.825	0.1309	48.3226	1
NDWI	1333.3804	0.0007	11547.2969	17987.136	0.0274	1103.3989	1
VARI	117.9869	0.0085	1013.8867	1579.324	0.0921	97.6365	1

1 --> COLLINEARITY is detected by the test
0 --> COLLINEARITY is not detected by the test

* all coefficients have significant t-ratios

R-square of y on all x: 0.3914

Como era de esperarse al tener un R-cuadrada ajustada menor en comparación con el modelo generado con variables polinomiales se detecta mayor colinealidad, en este caso según el análisis VIF la variable NDWI es la causante, al igual que en el modelo polinomial, pero al ser una variable de segundo grado disminuye el valor del mismo, se obtiene un R-cuadrada ajustada más alta y las variables ingresadas al modelo son estadísticamente más significativas que al analizarlas sin ser polinomiales, mejorando de esta manera el modelo.

5.1.5. Modelos con una muestra de entrenamiento y una de test.

Se procede a realizar el ajuste del modelo de regresión de Castielfabid con el método denominado Principal Components Regression (PCR) el mismo que utiliza mínimos cuadrados empleados como predictores las componentes generadas a partir de un Principal Component Analysis (PCA), obteniendo así un número reducido de componentes que expliquen la mayor parte de la varianza de los datos, se procede a realizar el proceso en el lenguaje de programación para análisis estadísticos R código adjunto en el ANEXO 3.

La metodología utilizada a continuación es útil para eliminar la información redundante, si se emplea como predictores las componentes principales se mejora el modelo de regresión. Es importante acotar que, si bien la regresión de componentes principales reduce el número de predictores del modelo, no es considerado como un método de selección de variables ya que todas las variables se necesitan para el cálculo de las componentes. El número óptimo de componentes principales que se empleará como predictores en PCR se identificará por la validación cruzada.

El conjunto de datos viene dado por el modelo ya calculado es decir por las cinco variables que aportan al mismo (NDVI, EVI, NDWI, VARI, $NDWI^2$) y la HCV_P que pertenece a la humedad de combustible vivo, todas estas observadas en 30 fechas.

Para poder evaluar la capacidad predictiva del modelo, se divide las observaciones disponibles en dos grupos: uno de entrenamiento para ajustar el modelo (86% de los datos) y uno de test (14% de los datos), es decir 26 y 4 datos respectivamente estos últimos datos han sido escogidos uno para cada estación del año, para invierno se tomó los datos del 30 de enero del 2017, para primavera se tomó los datos del 3 de mayo del 2017, para verano se tomó los datos del 31 de julio del 2017 y para otoño se tomó los datos del 27 de noviembre del 2017.

En primer lugar, se procede a realizar el ajuste del modelo incluyendo las cinco variables como predictores y utilizando los datos de entrenamiento.

Call:

```
lm(formula = as.formula(equation), data = training)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-17.5450	-4.7904	0.6823	6.4740	16.1012

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-570.74	96.56	-5.911	8.82e-06	***
NDVI	-1706.18	285.71	-5.972	7.71e-06	***
EVI	-52.67	11.11	-4.739	0.000126	***
NDWI	-5478.75	770.42	-7.111	6.84e-07	***
VARI	1711.03	234.10	7.309	4.57e-07	***
I(NDWI^2)	-4575.90	668.51	-6.845	1.19e-06	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.137 on 20 degrees of freedom

Multiple R-squared: 0.7982, Adjusted R-squared: 0.7477

F-statistic: 15.82 on 5 and 20 DF, p-value: 2.35e-06

Se obtiene que el valor de R-cuadrada ajustada es 0.7477 lo que indica que el modelo es capaz de predecir la humedad de combustible vivo ponderado de las observaciones con las que se ha entrenado. El P-valor (2.35e-06) indica que el modelo en conjunto es significativo. Al comparar este modelo con respecto al modelo usando todos los datos como se muestra en el ítem 5.1.1., el valor de R-cuadrada ajusta aumenta significativamente.

Para poder comparar los errores en los datos de entrenamiento y en el conjunto de datos tipo test, la estimación del error de predicción se obtendrá mediante el Error Medio Cuadrático (MSE). Primero se obtiene el MSE de las observaciones de entrenamiento un valor de 64.22399 y el MSE para las nuevas observaciones es decir el test es 469.6426. Se observa que el MSE del modelo es alto cuando predice las mismas observaciones con las que se ha entrenado, siendo el MSE siete veces más grande al predecir nuevas observaciones, esto puede significar que la utilidad del modelo es cuestionable ya que el objetivo es predecir la humedad de combustible vivo de futuras muestra.

Se recurre a la selección de predictores mediante stepwise selection empleando el criterio de información de Akaike (AIC) como criterio de evaluación, de esta manera se da una medida de calidad relativa del modelo estadístico, para el conjunto de datos; el número de predictores del modelo resultante es 6 es decir que se trabajara con todas las variables (NDVI, EVI, NDWI, VARI, NDWI^2) y la HCV_P.

Se vuelve a sacar el MSE para el nuevo modelo generado por el proceso stepwise selection para los 6 predictores disponibles, el MSE de las observaciones de entrenamiento obteniendo un valor de 64.22399 y el MSE para las nuevas observaciones es decir el test es 469.6426; al no haber eliminado predictores los resultados son los mismos valores que se obtuvieron con el anterior modelo.

Se procede a ajustar el modelo empleando las componentes principales para las cinco variables analizadas, a continuación, se muestra la varianza explicada y acumulada de las 5 primeras componentes.

	PC1	PC2	PC3	PC4	PC5
Standard deviation	1.83702	1.020255	0.7437219	0.1738888	0.03284757
Proportion of Variance	0.67493	0.208180	0.1106200	0.0060500	0.00022000
Cumulative Proportion	0.67493	0.883110	0.9937400	0.9997800	1.00000000

El estudio de la proporción de varianza explicada muestra que la primera componente recoge la mayor parte de la información (67.493%), decayendo drásticamente la varianza en las sucesivas componentes.

Una vez que se obtuvo el valor de las componentes para cada observación se ajusta el modelo lineal empleando dichos valores junto con la variable respuesta que le corresponde a cada observación. Según la proporción de varianza acumulada se emplea las 4 primeras componentes ya que en conjunto explica el 99.978% de varianza.

Se obtiene el MSE para el test es decir las nuevas observaciones (285.9778), para modelo que emplea como predictores las 4 primeras componentes, este valor es menor que el obtenido con el modelo generado por stepwise selection y el obtenido para todos los predictores (469.6426), esto indica que emplear componentes principales como predictores es útil para este caso, mientras que el MSE para los datos de entrenamiento es (268.4061) siendo mayor en comparación a los modelos antes calculados.

Por último, se recurre a cross validation en los componentes principales para identificar el número óptimo de componentes con el que se minimizaría el MSE, siendo en este caso 6, empleando este número en el PCR se consigue el test-MSE a 285.9778, siendo el mismo valor que el modelo anterior.

La Tabla 14 muestra los valores predichos con los modelos antes explicados tanto para los datos de entrenamiento como para el test, los datos obtenidos con el MSE más bajo (285.9778) son los datos que tienen menor residual para los datos test caso contrario para los datos de entrenamiento los valores que tienen menor residual son los obtenidos en Stepwise Selection obteniendo un MSE (64.22399).

Tabla 14. Comparación de modelos datos de entrenamiento y test

	Nº	Fecha toma	HCV_P	Modelo (Training)	Modelo (Stepwise Selection)	Modelo (PCR)	Modelo (PCR_CV)
Entrenamiento	1	15/1/2016	105.4850	97.9542	97.9542	101.3842	101.3842
	2	10/3/2016	91.6828	98.1817	98.1817	100.8708	100.8708
	3	29/3/2016	101.7750	94.6801	94.6801	93.7737	93.7737
	4	18/4/2016	98.5750	107.9714	107.9714	97.2706	97.2706
	5	16/5/2016	108.5872	107.7005	107.7005	91.5280	91.5280
	6	21/6/2016	107.0711	90.9699	90.9699	78.6689	78.6689
	7	19/7/2016	64.1345	81.6795	81.6795	77.9285	77.9285
	8	24/8/2016	60.1100	64.0292	64.0292	79.6736	79.6736
	9	5/10/2016	84.1725	72.4033	72.4033	86.4577	86.4577
	10	1/11/2016	84.1535	97.5140	97.5140	93.5949	93.5949
	11	2/12/2016	90.3325	89.8547	89.8547	109.6662	109.6662
	12	22/12/2016	101.1775	98.5170	98.5170	105.1515	105.1515
	13	21/2/2017	85.3635	90.4443	90.4443	84.4321	84.4321
	14	28/3/2017	82.9920	83.4735	83.4735	77.1739	77.1739
	15	22/5/2017	93.0545	86.8989	86.8989	91.9693	91.9693
	16	26/6/2017	86.1840	87.3969	87.3969	92.0993	92.0993
	17	12/9/2017	84.3225	87.9462	87.9462	83.4230	83.4230
	18	26/9/2017	84.7820	93.6118	93.6118	87.8030	87.8030
	19	27/11/2017	59.7600	52.9454	52.9454	75.6879	75.6879
	20	20/12/2017	84.4220	77.8418	77.8418	73.5323	73.5323
	21	7/3/2018	74.5535	66.6058	66.6058	77.2131	77.2131
	22	27/3/2018	57.5300	69.9160	69.9160	76.4571	76.4571
	23	22/4/2018	63.5705	64.9291	64.9291	88.6825	88.6825
	24	23/5/2018	124.5685	123.5157	123.5157	95.0335	95.0335
	25	12/6/2018	110.2055	105.5156	105.5156	88.8018	88.8018
	26	19/7/2018	111.8815	102.9288	102.9288	87.1477	87.1477
Test	27	30/1/2017	82.2540	90.1000	90.1000	93.8000	93.8000
	28	3/5/2017	82.7590	64.5000	64.5000	90.8000	90.8000
	29	31/7/2017	70.6100	108.5000	108.5000	95.4000	95.4000
	30	2/11/2017	54.7385	67.0000	67.0000	77.9000	77.9000

En la Imagen 24 se observa el comportamiento de los modelos calculados con datos de entrenamiento y test, el comportamiento de los 4 datos estimados de test muestra que no varía en gran cantidad a excepción de dato de la fecha 31/7/2017 existiendo una diferencia de 37.89 entre el dato real de HCV y el dato del modelo (Training).

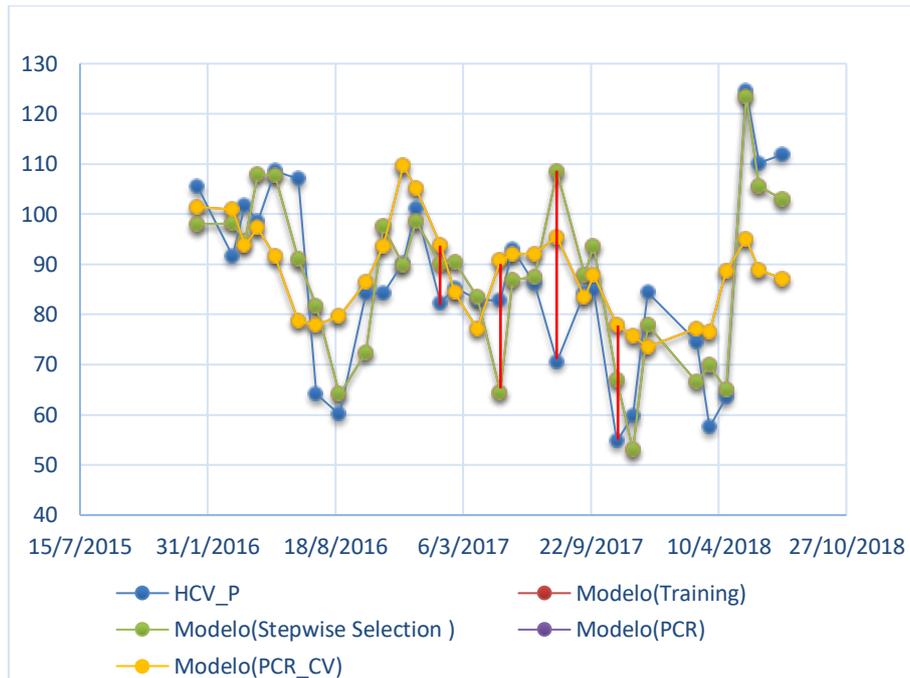


Imagen 24. Gráfica comparativa de modelos calculados por Regresión de Componentes Principales (PCR).

5.1.6. Análisis de validación cruzada.

El resultado del modelo antes presentado se verifica al realizar la regresión lineal múltiple para obtener la validación cruzada del mismo en el lenguaje de programación para análisis estadísticos R.

La validación cruzada se emplea para estimar el error que comete un modelo al predecir observaciones o datos que no han sido consideradas en dicho modelo y así evaluar la capacidad predictiva. La función para realizar la validación cruzada de la regresión lineal proporciona medidas de validación interna y cruzada de precisión predictiva, los datos se asignarán aleatoriamente a un número de “folds”, cada folds se elimina, a su vez, mientras que los datos restantes se utilizan para volver a ajustar el modelo de regresión y para predecir las observaciones eliminadas.

Se utiliza la validación cruzada para evaluar los resultados del análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento. Se predice y estima la precisión del modelo.

En R se obtiene el resumen de la regresión lineal el cual corrobora los resultados obtenidos en el software Statgraphics, como se muestra a continuación.

La sección de Residuals entrega 5 estadísticos sobre la distribución de los residuos del modelo: valores mínimos, 1er, 2do y 3er cuartil y valor máximo. Si los residuos se distribuyen conforme a una distribución normal se espera que la mediana sea 0 o muy cerca de 0, y que el 1er y 3er cuartil sean simétricos. Lo mismo para los valores mínimos o máximos.

La sección de Coefficients; muestra los coeficientes estimados por el modelo, es decir, los parámetros ocultos β_0 y β_1 de la ecuación. Estas son la ordenada al origen y las pendientes estimadas para cada variable. En la tercera columna se registra el error estándar para la estimación de cada variable, al que se puede interpretar como el promedio de los residuos. A partir del coeficiente estimado y el error estándar se computa un valor T, un estadístico de la divergencia entre el estimado que produce el modelo y un estimado hipotético con valor 0.

En la última columna se muestra el error estándar de los residuos y los valores de R^2 , múltiple y ajustado. El R^2 de 0.574 indica el modelo explica un 57.4% de la varianza de la variable dependiente.

Residuals:

Min	1Q	Median	3Q	Max
-34.56	-5.99	-0.18	7.73	19.51

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-453.8	113.8	-3.99	0.00054	***
NDVI	-1465.8	335.2	-4.37	0.00020	***
EVI	-42.6	12.1	-3.53	0.00173	**
NDWI	-4592.2	896.7	-5.12	3.1e-05	***
VARI	1504.5	272.6	5.52	1.1e-05	***
I(NDWI^2)	-3798.5	793.4	-4.79	7.1e-05	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11.7 on 24 degrees of freedom

Multiple R-squared: 0.647, Adjusted R-squared: 0.574

F-statistic: 8.8 on 5 and 24 DF, p-value: 7.48e-05

La tabla de análisis de varianza entrega la información de la suma y la media de los cuadrados al calcular los datos predichos. Para la presente validación cruzada se utilizó un $m=10$ es decir se divide a las muestras en 3 grupos y se predice los valores como se muestra en la sección fold 1, el valor m se elige según convenga, para este trabajo se ha elegido 3 grupos, se elimina un dato y ajusta de nuevo el modelo, dando la predicción de ese dato y se observa el error cometido (CV Residual).

Analysis of Variance Table

Response: HCV_P

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
NDVI	1	729	729	5.30	0.0303	*
EVI	1	82	82	0.59	0.4489	
NDWI	1	782	782	5.69	0.0253	*
VARI	1	1310	1310	9.53	0.0051	**
I(NDWI^2)	1	3152	3152	22.92	7.1e-05	***
Residuals	24	3301	138			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

fold 1

Observations in test set: 3

	[,1]	[,2]	[,3]
Predicted	92.60	79.2	87.94
cvpred	91.93	81.1	85.63
HCV_P	101.78	64.1	93.05
CV residual	9.85	-17.0	7.43

Sum of squares = 441 Mean square = 147 n = 3

Esta función muestra la suma de los errores al cuadrado (MS) que son comparables con los valores en los anteriores análisis en este caso se obtiene (147) siendo menor que los antes analizados obteniendo así residuales menores.

5.1.7. Filtro de suavizado Savitsky-Golay.

Una vez obtenidos los mejores modelos de regresión para los 18 puntos de muestreo de la humedad de combustible vivo y los índices de vegetación se aplica el filtro Savitsky-Golay para suavizar el ruido típico de estas series de tiempo con el fin de mejorar los modelos. Este método se basa en el cálculo de una regresión polinomial local de grado n , con al menos $n+1$ puntos equidistantes, con el fin de determinar el nuevo valor de cada punto. El resultado será una función equivalente a los datos de entrada, pero suavizada. Esta técnica de suavizado se usa ampliamente ya que conserva las características de la distribución inicial, como los máximos y mínimos relativos y el ancho máximo (Jönsson & Eklundh, 2004). El valor suavizado para una fecha determinada se calculó utilizando una ventana temporal de ± 3 fechas tanto para humedad de combustible vivo como para los índices de vegetación. Este proceso se lo realizó en R utilizando el paquete “sgolayfilt” como se muestra en el código adjunto en el ANEXO 4.

En la Imagen 25 se muestra los datos de humedad de combustible vivo reales del punto de muestreo de Castielfabid y los datos suavizados de la misma variable con el filtro antes mencionado. Se observa que los datos no pierden sus características de serie temporal.

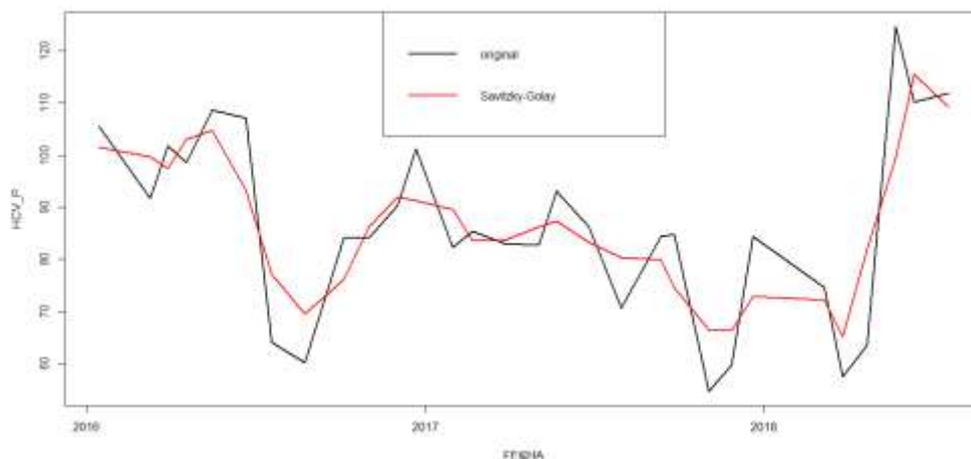


Imagen 25. Gráfica HCV_P vs HCV_P filtrado de Castielfabid.

Este proceso se realiza de igual forma para los 6 índices de vegetación obteniendo así nuevos valores para el cálculo del modelo.

Como modelo se toma el punto de muestreo Castielfabid del cual el software toma seis variables NDVI_F, NDWI_F, VARI_F, Vigreen_F, EVI_F² y NDWI_F² obteniendo valores-P en la tabla ANOVA menores a 0.05 existiendo una relación estadísticamente significativa entre las variables con un nivel de confianza del 95%. La Tabla 15 muestra los valores de las estimaciones de cada una de las variables, así como el error estándar, el estadístico T-student y sus valores-P, la regresión múltiple para este punto de muestreo obtiene un R-cuadrado ajustado de 70.8273%.

Tabla 15. Regresión múltiple HCV_P y las 12 variables de índices de vegetación aplicadas el filtro Savitzki Golay.

Parámetro	Estimación	Error		Valor-P
		Estándar	Estadístico T	
CONSTANTE	-569.464	125.657	-4.5319	0.0001
NDVI_F	-1554.73	333.733	-4.65862	0.0001
NDWI_F	-5230.78	906.971	-5.7673	0.0000
VARI_F	1536.69	250.513	6.13418	0.0000
Vigreen_F	307.803	147.758	2.08315	0.0485
EVI_F ²	-50.7102	10.1934	-4.97479	0.0000
NDWI_F ²	-4367.7	818.043	-5.33921	0.0000

Análisis de Varianza

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	3968.91	6	661.485	12.73	0.0000
Residuo	1194.7	23	51.9436		
Total (Corr.)	5163.61	29			

R-cuadrada = 76.863 por ciento

R-cuadrado (ajustado para g.l.) = 70.8273 por ciento

Error estándar del est. = 7.20719

Error absoluto medio = 4.57932

Estadístico Durbin-Watson = 0.925647 (P=0.0000)

Autocorrelación de residuos en retraso 1 = 0.390384

La salida muestra los resultados de ajustar un modelo de regresión múltiple para describir la relación entre HCV_P suavizada y 6 variables independientes suavizadas. La ecuación del modelo ajustado es:

$$\text{HCV_P_FILTRO} = -569.464 - 1554.73 \cdot \text{NDVI_F} - 5230.78 \cdot \text{NDWI_F} + 1536.69 \cdot \text{VARI_F} + 307.803 \cdot \text{Vigreen_F} - 50.7102 \cdot \text{EVI_F}^2 - 4367.7 \cdot \text{NDWI_F}^2 \quad (8)$$

El estadístico R-Cuadrada indica que el modelo así ajustado explica 76.863% de la variabilidad en HCV_P suavizada. El estadístico R-Cuadrada ajustada, que es más apropiada para comparar modelos con diferente número de variables independientes, es 70.8273%. El error estándar del estimado muestra que la desviación estándar de los residuos es 7.20719. Este valor puede usarse para construir límites para nuevas observaciones. El error absoluto medio (MAE) de 4.57932 es el valor promedio de los residuos.

En la Imagen 26 se puede verificar el ajuste del modelo calculado a partir de la regresión lineal múltiple para los datos suavizados.

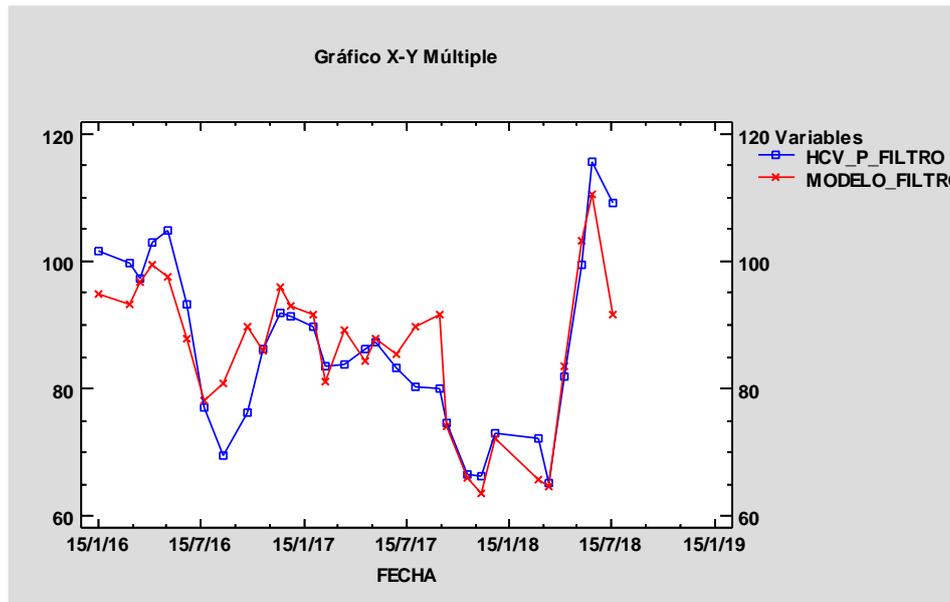


Imagen 26. Gráfica ajuste del modelo para índices de vegetación suavizados en el punto de muestro Castielfabid.

Con este método de datos suavizados a partir del filtro Savitsky-Golay se obtuvo los modelos para los 17 puntos de muestreo restantes como se muestra en la Tabla 16 obteniendo los mejores resultados de R-cuadrada ajustado para Betera, Biar, Villar del Arzobispo, Castielfabid y Muela Cortes, este último con el R² ajustada mayor se debe a que en dicho punto de muestreo cuenta con menos observaciones.

Tabla 16. Modelos de regresión múltiple para los 18 puntos de muestreo con datos suavizados por el filtro Savitsky Golay.

Nº	Punto de muestreo	R-Cuadrada ajustada	Modelo
1	Fredes	30.2229	$HCV_P_FILTRO = 137.549 + 3015.16*VARI_F - 5392.46*Vigreen_F - 7780.93*VARI_F^2 + 22401.4*Vigreen_F^2$
2	Puebla Tornesa	57.1631	$HCV_P_FILTRO = 31600.9 - 64.8588*VARI_F + 46.4794*NDVI_F^2 - 14021.4*SAVI_F^2 + 607.215*VARI_F^2.$
3	Montanejos	58.4401	$HCV_P_FILTRO = -119.453 - 228.285*EVI_F - 1331.86*NDWI_F + 4740.87*VARI_F - 6828.62*Vigreen_F + 66.8174*EVI_F^2 - 1069.71*NDWI_F^2 - 16449.1*VARI_F^2 + 33324.8*Vigreen_F^2$
4	Castielfabid	70.8273	$HCV_P_FILTRO = -569.464 - 1554.73*NDVI_F - 5230.78*NDWI_F + 1536.69*VARI_F + 307.803*Vigreen_F - 50.7102*EVI_F^2 - 4367.7*NDWI_F^2$
5	Chelva	48.5268	$HCV_P_FILTRO = 368.852 + 3324.0*NDVI_F + 4226.81*NDWI_F - 2737.55*NDVI_F^2 - 24.5644*EVI_F^2 + 3554.63*NDWI_F^2 + 14381.9*VARI_F^2 - 23135.9*Vigreen_F^2$

Nº	Punto de muestreo	R-Cuadrada ajustada	Modelo
6	Villar del Arzobispo	75.4658	$HCV_P_FILTRO = -344013. + 6545.35*NDVI_F - 381.788*EVI_F + 229724.*SAVI_F + 6728.46*NDWI_F - 3295.83*VARI_F + 151.863*EVI_F^2$
7	Betera	88.1827	$HCV_P_FILTRO = 110147. + 529.825*EVI_F - 73666.8*SAVI_F - 2091.11*Vigreen_F + 1915.84*NDVI_F^2 - 154.11*EVI_F^2 - 1965.51*NDWI_F^2$
8	Gilet	71.0852	$HCV_P_FILTRO = 63.5441 - 5085.04*VARI_F + 7107.38*Vigreen_F + 23.677*EVI_F^2 - 2826.77*VARI_F^2$
9	Villargordo del Cabriel	40.0543	$HCV_P_FILTRO = 80.8731 + 3876.45*NDVI_F + 3854.04*NDWI_F - 2779.13*Vigreen_F$
10	Buñol	50.8378	$HCV_P_FILTRO = -139148. + 374.406*EVI_F + 92721.7*SAVI_F - 169.418*EVI_F^2 + 34252.4*VARI_F^2 - 60989.6*Vigreen_F^2$
11	Llombai	36.6372	$HCV_P_FILTRO = 159879. - 106566.*SAVI_F + 26358.4*VARI_F^2 - 68570.9*Vigreen_F^2$
12	Cortes de Pallás	88.8313	$HCV_P_FILTRO = -487275. - 14740.5*NDVI_F + 886.72*EVI_F + 327347.*SAVI_F + 8350.82*VARI_F - 385.805*EVI_F^2 + 13415.5*NDWI_F^2$
13	Navarrés	42.0542	$HCV_P_FILTRO = -193.555 - 1751.83*NDVI_F - 2562.46*NDWI_F + 1310.9*NDVI_F^2 + 21.4682*EVI_F^2 - 1992.42*NDWI_F^2$
14	Ayora	68.5476	$HCV_P_FILTRO = -97.2047 + 3760.12*NDVI_F - 43.1947*EVI_F + 2813.92*NDWI_F - 1722.42*VARI_F - 1096.51*NDVI_F^2 + 9.9651*EVI_F^2 + 1622.26*VARI_F^2 - 2578.15*Vigreen_F^2$
15	Gandia	63.6167	$HCV_P_FILTRO = -359810. - 1800.63*NDVI_F + 240360.*SAVI_F + 5169.42*VARI_F - 4650.57*Vigreen_F + 1480.03*NDWI_F^2 - 23845.4*VARI_F^2 + 32624.1*Vigreen_F^2$
16	Muro de Alcoi	57.3054	$HCV_P_FILTRO = 117166. + 967.56*EVI_F + 2178.06*NDWI_F - 7746.21*Vigreen_F - 52294.1*SAVI_F^2 - 11411.6*Vigreen_F^2$
17	Altea	57.3054	$HCV_P_FILTRO = 37483.4 - 3438.94*NDVI_F + 46.9777*EVI_F - 24864.5*SAVI_F - 3063.92*NDWI_F - 687.824*VARI_F + 1334.32*Vigreen_F + 2537.57*NDVI_F^2 - 2468.29*NDWI_F^2 - 1612.37*Vigreen_F^2$
18	Biar	75.7419	$HCV_P_FILTRO = -208.274 - 1965.4*NDVI_F + 523.089*EVI_F - 1923.16*NDWI_F + 880.314*VARI_F - 204.016*EVI_F^2 + 131.953*VARI_F^2$

5.2. Análisis de modelos de regresión múltiple entre HCV e índices de vegetación más variables meteorológicas.

5.2.1. Modelo de regresión múltiple general para la Comunidad Valenciana.

En este apartado se explica la generación y análisis de un modelo de regresión múltiple para todos los puntos de muestreo, considerándolo único para toda la Comunidad Valenciana. La regresión se trabaja con los 18 puntos de muestreo obteniendo un total de 552 observaciones y 40 variables pertenecientes a índices de vegetación y meteorológicas; en este caso se tomó como procedimiento la selección paso a paso hacia atrás; este comienza con un modelo que involucra las variables y elimina del modelo las variables que son estadísticamente menos significativas, es decir las que tienen un valor P determinado por el α (valor-P) siendo menor a 0.05, obteniendo así las variables que más aporten al modelo.

El software Statgraphics toma dos variables de índices de vegetación (NDVI, NDWI) y once variables meteorológicas (P7, P60, Tmax7, Tmax30, Tmedia7, Tmedia15, HRmax7, HRmax15, HRmin7, HRmin15, lev3) obteniendo valores-P en la tabla ANOVA menores a 0.05 existiendo una relación estadísticamente significativa entre las variables con un nivel de confianza del 95%. En la Tabla 17 se muestra los resultados de la regresión múltiple para este punto de muestreo obteniendo un R-cuadrada ajustada de 22.2183%, siendo una bondad de ajuste baja para considerarla un buen modelo.

Tabla 17. Regresión múltiple HCV_P y las variables de índices de vegetación y meteorológicas para la Comunidad Valenciana.

		Error	Estadístico	
Parámetro	Estimación	Estándar	T	Valor-P
CONSTANTE	77.0417	6.37141	12.0918	0.0000
NDVI	56.8198	12.8781	4.41211	0.0000
NDWI	34.903	17.4279	2.00271	0.0457
P7	-0.0859835	0.0368407	-2.33392	0.0200
P60	0.0630818	0.0125731	5.0172	0.0000
Tmax7	1.57162	0.448395	3.50499	0.0005
Tmax30	-1.99378	0.407336	-4.89468	0.0000
Tmedia7	-2.06043	0.756051	-2.72525	0.0066
Tmedia15	1.97371	0.727117	2.71443	0.0069
HRmax7	-0.367499	0.147922	-2.48441	0.0133
HRmax15	0.500985	0.159024	3.15037	0.0017
HRmin7	0.501531	0.173289	2.89419	0.0040
HRmin15	-0.59437	0.183949	-3.23116	0.0013
lev3	0.534026	0.189686	2.81532	0.0051

Análisis de Varianza

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	37355.5	13	2873.5	13.11	0.0000
Residuo	117947.	538	219.232		
Total (Corr.)	155302.	551			

R-cuadrada = 24.0534 por ciento
 R-cuadrado (ajustado para g.l.) = 22.2183 por ciento
 Error estándar del est. = 14.8065
 Error absoluto medio = 11.0402
 Estadístico Durbin-Watson = 1.06507 (P=0.0000)
 Autocorrelación de residuos en retraso 1 = 0.467155

La salida muestra los resultados de ajustar un modelo de regresión múltiple para describir la relación entre HCV_P y 13 variables independientes. La ecuación del modelo ajustado es:

$$\begin{aligned} \text{HCV_P} = & 77.0417 + 56.8198 \cdot \text{NDVI} + 34.903 \cdot \text{NDWI} - 0.0859835 \cdot \text{P7} + 0.0630818 \cdot \text{P60} + \\ & 1.57162 \cdot \text{Tmax7} - 1.99378 \cdot \text{Tmax30} - 2.06043 \cdot \text{Tmedia7} + 1.97371 \cdot \text{Tmedia15} - 0.367499 \cdot \text{HRmax7} + \\ & 0.500985 \cdot \text{HRmax15} + 0.501531 \cdot \text{HRmin7} - 0.59437 \cdot \text{HRmin15} + 0.534026 \cdot \text{lev3} \end{aligned} \quad (9)$$

El estadístico R-Cuadrada indica que el modelo así ajustado explica 24.0534% de la variabilidad en HCV_P. El estadístico R-Cuadrada ajustada, que es más apropiada para comparar modelos con diferente número de variables independientes, es 22.2183%. El error estándar del estimado muestra que la desviación estándar de los residuos es 14.8065. Este valor puede usarse para construir límites para nuevas observaciones. El error absoluto medio (MAE) de 11.0402 es el valor promedio de los residuos.

El modelo calculado para la Comunidad Valenciana ha sido aplicado a los puntos de muestreo, obteniendo así modelos que no se ajustan a los valores de HCV por tener un R-cuadrada muy baja. En la Imagen 27 se puede observar como ejemplo el ajuste del modelo calculado a partir de la regresión múltiple para el punto de muestreo Altea, el modelo sigue la tendencia de los valores de HCV sobre todo en los datos pertenecientes de enero a abril del 2017 en donde se forma un pico predominante. Las variables meteorológicas de precipitación aportan al modelo y hacen que sigan la tendencia ya que en este rango de tiempo hubo lluvias constantes.

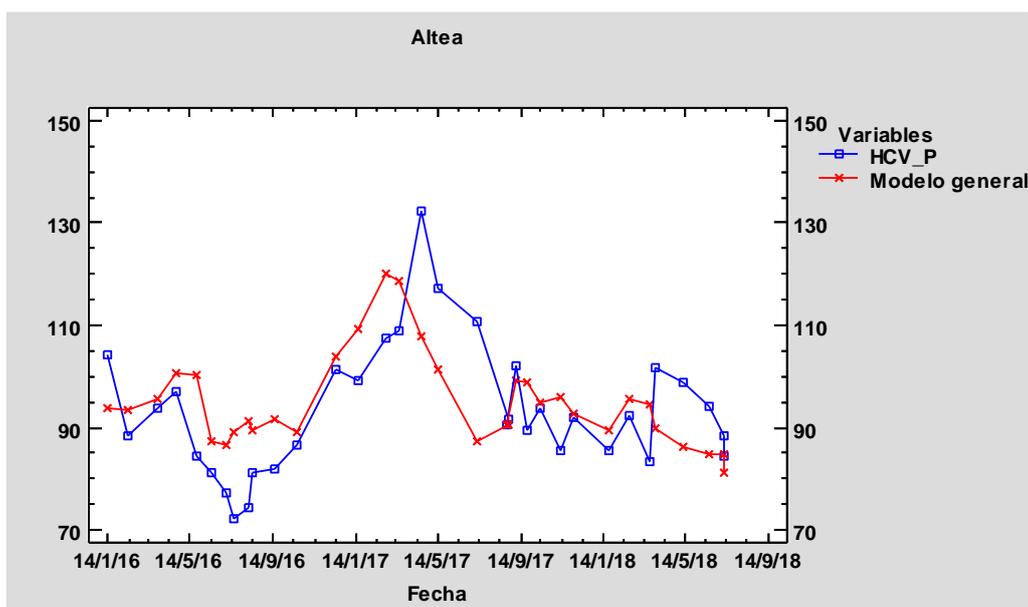


Imagen 27. Gráfica ajuste del modelo para la Comunidad Valenciana en el punto de muestreo Altea.

En la Imagen 28 se observa el ajuste del modelo calculado a partir de la regresión múltiple para el punto de muestreo Castielfabid, el modelo intenta seguir la tendencia de los valores de HCV, pero al existir un cambio constante en los valores de HCV y al tener una bondad de ajuste tan baja en el modelo no se obtiene buenos resultados.

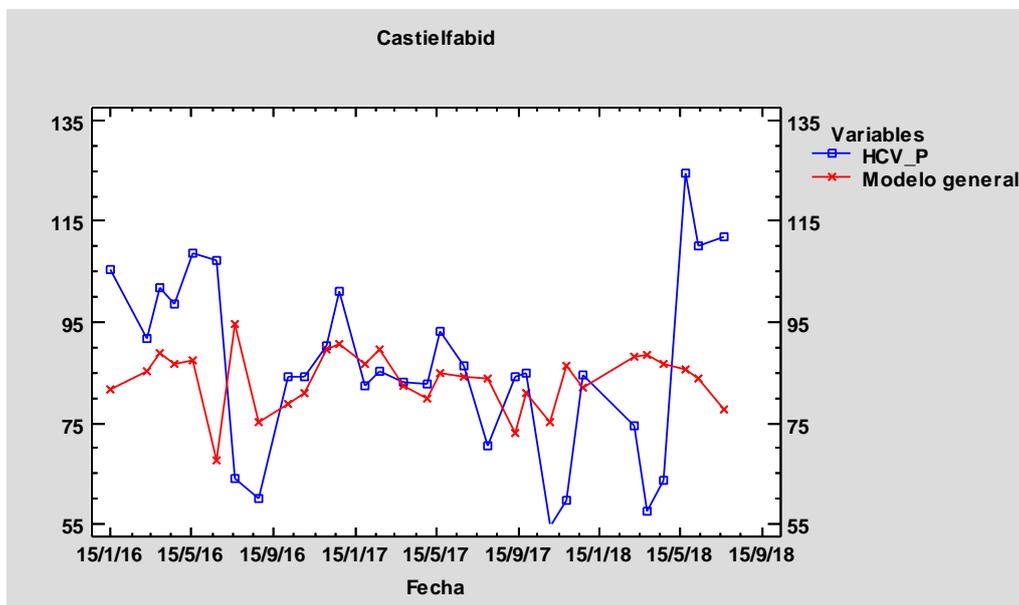


Imagen 28. Gráfica ajuste del modelo para la Comunidad Valenciana en el punto de muestreo Castielfabid.

Después de obtener este resultado se procede a realizar el análisis de modelos diferentes para cada punto de muestreo. Un único modelo para toda la Comunidad Valenciana no entrega buenos resultados, se puede decir que, la extensión del territorio y la topografía afectan tanto en las variables meteorológicas como en los índices de vegetación y por tal razón se considera un modelo diferente para cada punto de muestreo.

5.2.2. Modelo de regresión múltiple para los 18 puntos de muestreo.

La humedad de combustible vivo en este caso se ve afectado por los índices de vegetación y por las variables meteorológicas, al ser mayor el número de variables (40) que de observaciones (30) se opta por realizar la regresión múltiple en grupos, es decir, ingresando datos de manera independiente de índices de vegetación, precipitación, temperatura, humedad relativa, viento, levante y punto del rocío; en este caso se tomó como procedimiento la selección paso a paso hacia atrás; este comienza con un modelo que involucra las variables de cada grupo y elimina del modelo la variable que es estadísticamente menos significativa, es decir la que tiene un valor P determinado por el α (valor-P) siendo menor a 0.05 para cada uno de los grupos, obteniendo así las variables que más aporten al modelo.

El software Statgraphics toma ocho variables de índices de vegetación (NDVI, EVI, NDWI, VARI, VIgreen, NDVI², VARI², VIgreen²) y ocho variables meteorológicas (P15, P60, Tmax7, Tmax15, Tmedia7, Tmedia15, HRmax7, HRmin15) obteniendo valores-P en la tabla ANOVA menores a 0.05 existiendo una relación estadísticamente

significativa entre las variables con un nivel de confianza del 95%. En la Tabla 18 se muestra los resultados de la regresión múltiple para este punto de muestreo obteniendo un R-cuadrada ajustada de 94.6268%.

Tabla 18. Regresión múltiple HCV_P y las variables de índices de vegetación y meteorológicas.

Parámetro	Estimación	Error		Estadístico	Valor-P
		Estándar	T		
CONSTANTE	-613.364	69.5677	-8.81679	0.0000	
NDVI	2997.81	514.291	5.82902	0.0001	
EVI	-69.733	7.53737	-9.25163	0.0000	
NDWI	-1170.95	315.524	-3.71113	0.0026	
VARI	-7633.03	2537.79	-3.00774	0.0101	
Vigreen	10242.8	3178.18	3.22286	0.0067	
NDVI^2	-5832.32	481.191	-12.1206	0.0000	
VARI^2	-40137.3	10020.8	-4.00538	0.0015	
Vigreen^2	65169.7	18872.0	3.45326	0.0043	
P15	0.30657	0.0616895	4.96956	0.0003	
P60	-0.219474	0.0435537	-5.03916	0.0002	
Tmax7	-7.00221	0.841493	-8.32118	0.0000	
Tmax15	8.71042	1.37584	6.33098	0.0000	
Tmedia7	8.81067	1.55384	5.67028	0.0001	
Tmedia15	-10.44	2.09519	-4.98281	0.0003	
HRmax7	-0.818569	0.195737	-4.182	0.0011	
HRmin15	1.30837	0.262355	4.98701	0.0002	

Análisis de Varianza

Fuente	Suma de Cuadrados	Gl	Cuadrado Medio	Razón-F	Valor-P
Modelo	9129.35	16	570.584	32.92	0.0000
Residuo	225.324	13	17.3326		
Total (Corr.)	9354.67	29			

R-cuadrada = 97.5913 por ciento

R-cuadrado (ajustado para g.l.) = 94.6268 por ciento

Error estándar del est. = 4.16324

Error absoluto medio = 2.24996

Estadístico Durbin-Watson = 2.49431 (P=0.6875)

Autocorrelación de residuos en retraso 1 = -0.275567

La salida muestra los resultados de ajustar un modelo de regresión múltiple para describir la relación entre HCV_P y 16 variables independientes. La ecuación del modelo ajustado es:

$$\begin{aligned}
 \text{HCV_P} = & -613.364 + 2997.81 \cdot \text{NDVI} - 69.733 \cdot \text{EVI} - 1170.95 \cdot \text{NDWI} - 7633.03 \cdot \text{VARI} + \\
 & 10242.8 \cdot \text{Vigreen} - 5832.32 \cdot \text{NDVI}^2 - 40137.3 \cdot \text{VARI}^2 + 65169.7 \cdot \text{Vigreen}^2 + 0.30657 \cdot \text{P15} - \\
 & 0.219474 \cdot \text{P60} - 7.00221 \cdot \text{Tmax7} + 8.71042 \cdot \text{Tmax15} + 8.81067 \cdot \text{Tmedia7} - 10.44 \cdot \text{Tmedia15} - \\
 & 0.818569 \cdot \text{HRmax7} + 1.30837 \cdot \text{HRmin15}
 \end{aligned}
 \tag{10}$$

El estadístico R-Cuadrada indica que el modelo así ajustado explica 97.5913% de la variabilidad en HCV_P. El estadístico R-Cuadrada ajustada, que es más apropiada para comparar modelos con diferente número de variables independientes, es 94.6268%. El error estándar del estimado muestra que la desviación estándar de los residuos es 4.16324.

Este valor puede usarse para construir límites para nuevas observaciones. El error absoluto medio (MAE) de 2.24996 es el valor promedio de los residuos.

En la Imagen 29 se puede verificar el ajuste del modelo calculado a partir de la regresión lineal múltiple.

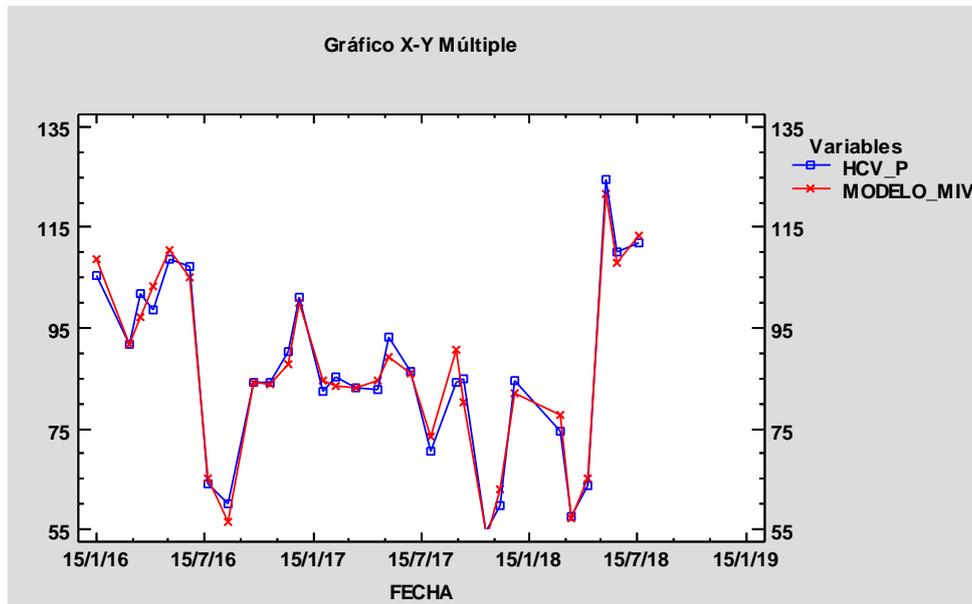


Imagen 29. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestreo Castielfabid.

Por este método se obtuvo los modelos para los 17 puntos de muestreo restantes como se muestra en la Tabla 22 de resultados.

5.2.3. Análisis del cumplimiento de las hipótesis del modelo de regresión: normalidad, linealidad, homocedasticidad e independencia.

En el modelo para el punto de muestreo Castielfabid no se rechaza la idea de que los residuos provienen de una distribución normal con 95% de confianza ya que su valor-P más pequeño de las pruebas realizadas es mayor a 0.05, como se muestra en la Tabla 19.

Tabla 19. Pruebas de Normalidad para Residuos

Prueba	Estadístico	Valor-P
Chi-Cuadrado	15.0	0.241436
Estadístico W de Shapiro-Wilk	0.972098	0.640501
Valor-Z para asimetría	0.28835	0.773075
Valor-Z para curtosis	-0.843028	0.399211

En otras palabras, la hipótesis nula o H_0 es que la variable RESIDUOS posee una distribución normal y H_1 , o hipótesis alternativa que no siga una distribución normal. Puesto que el valor-P es mayor que 0,05 no se rechaza H_0 . El test Chi-Cuadrado indica que los residuos siguen una distribución normal. Pero otro test como el Estadístico W de

Shapiro-Wilk y el de Valor-Z para curtosis son similares ya que los P-valores también son mayores que 0,05.

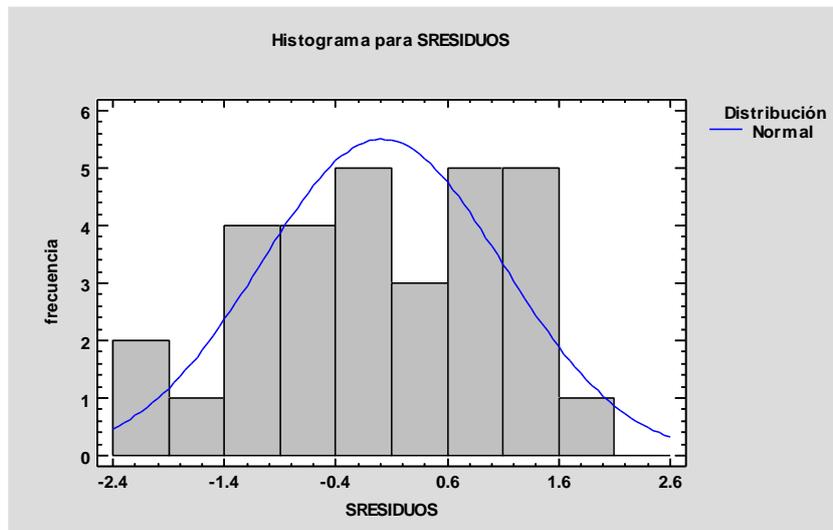


Imagen 30. Histograma para residuos de la distribución normal.

En la Imagen 30 indica que la distribución de los residuos no es más apuntada que la de la curva normal, también se observa que el modelo cumple con la hipótesis de linealidad, pero la homocedasticidad es discutible ya que la varianza no es la misma para todos los valores predichos, regularmente queda en evidencia mediante un patrón de forma en la gráfica de residuos; la linealidad se pierde en los valores más altos de la Y.

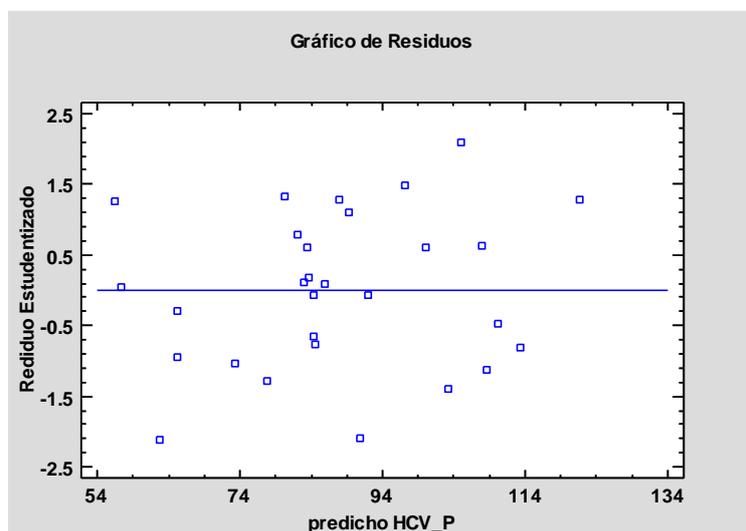


Imagen 31. Gráfica de residuos estudentizado vs predicho.

En análisis de la independencia de los residuos se utiliza el estadístico de Durbin-Watson (DW) el cual examina los residuos para determinar si hay alguna correlación significativa basada en el orden en el que se presentan en el archivo de datos. La Imagen 32 indica que hay independencia puesto que la nube de puntos no muestra ninguna tendencia alrededor de la recta $y=0$; el estadístico de Durbin Watson es 2.49431 ($P=0.6875$), puesto

que el valor-P es mayor que 0.05, no hay indicación de una autocorrelación serial en los residuos con un nivel de confianza del 95.0%.

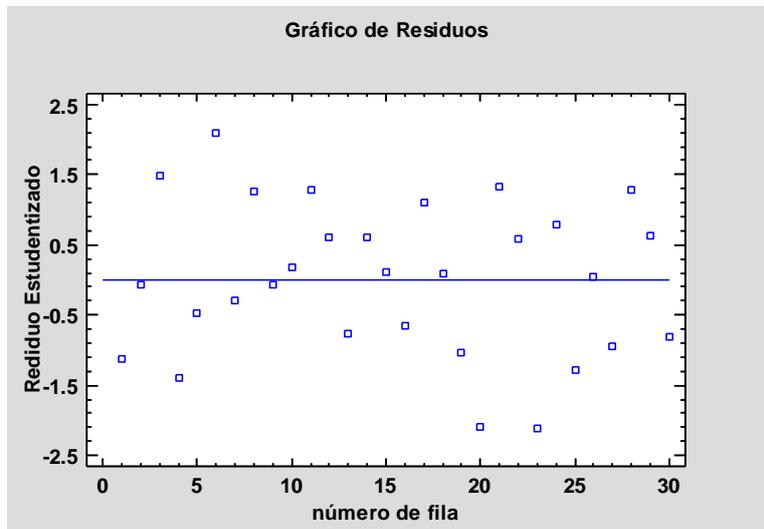


Imagen 32. Gráfica de residuos estudentizado vs número de fila.

5.2.4. Análisis de Multicolinealidad.

A partir del modelo anteriormente realizado en donde se obtuvo un R-cuadrada ajusta de 0.9462 en donde todas las variables de regresión individuales son estadísticamente significativas.

El diagnostico se realiza a partir de la obtención de la correlación por pares de variables explicativas obtenidas en el modelo de Castielfabid.



Imagen 33. Matriz de correlación.

La matriz de correlación en la Imagen 33 muestra que las correlaciones por pares entre las variables explicativas no son muy altos exceptuando los pares que pertenecen a los índices de vegetación NDVI y VARI teniendo correlaciones altas con sus respectivos polinomiales obteniendo correlaciones de 0.996 y -0.992 respectivamente. Las variables meteorológicas en las que se encuentra correlación alta son las pertenecientes a temperatura de las cuales Tmedia15-Tmax15 tiene una correlación de 0.973. La correlación entre estas variables puede ser la causa de multicolinealidad, en el apartado 5.1.4 se observa el motivo de la multicolinealidad al tener variables polinomiales, el presente modelo nos ofrece un R-cuadrada ajustada alto y sus variables son estadísticamente significativas.

Se realizar la comprobación de diagnóstico global e individual de las variables para el análisis de multicolinealidad.

Overall Multicollinearity Diagnostics

	MC Results	detection
Determinant $ X'X $:	0.0000	1
Farrar Chi-Square:	1163.7831	1
Red Indicator:	0.4034	0
Sum of Lambda Inverse:	26408.1106	1
Theil's Method:	0.5057	1
Condition Number:	1946.4946	1

1 --> COLLINEARITY is detected by the test
0 --> COLLINEARITY is not detected by the test

El valor del determinante estandarizado es 0.0000 siendo está nula, el valor de la prueba estadística chi-cuadrada es 1163.78 implicando este valor la presencia de multicolinealidad en la especificación del modelo.

Se procede a ubicar la multicolinealidad siguiendo la prueba de Glauber (F – test), obtenido así individualmente la misma.

All Individual Multicollinearity Diagnostics Result

	VIF	TOL	Wi	Fi	Leamer	CVIF	Klein
NDVI	1160.6588	0.0009	1082.3482	1242.4915	0.0294	-509.7376	1
EVI	7.1957	0.1390	5.7826	6.6382	0.3728	-3.1602	0
NDWI	548.5675	0.0018	511.0630	586.6794	0.0427	-240.9196	1
VARI	8207.9069	0.0001	7659.7798	8793.1145	0.0110	-3604.7452	1
Vigreen	6318.5233	0.0002	5896.3550	6768.7749	0.0126	-2774.9665	1
NDVI2	471.0269	0.0021	438.6917	503.6002	0.0461	-206.8654	1
VARI2	4981.3577	0.0002	4648.3338	5336.0975	0.0142	-2187.7106	1
Vigreen2	3887.3510	0.0003	3627.2609	4163.9475	0.0160	-1707.2452	1
P15	2.6865	0.3722	1.5740	1.8069	0.6101	-1.1798	0
P60	4.4713	0.2236	3.2399	3.7193	0.4729	-1.9637	0
Tmax7	90.1582	0.0111	83.2143	95.5266	0.1053	-39.5956	1
Tmax15	203.3161	0.0049	188.8283	216.7672	0.0701	-89.2923	1
Tmedia7	174.1905	0.0057	161.6445	185.5613	0.0758	-76.5009	1
Tmedia15	305.8574	0.0033	284.5336	326.6329	0.0572	-134.3263	1
HRmax7	24.7001	0.0405	22.1201	25.3930	0.2012	-10.8478	0
HRmin15	20.1429	0.0496	17.8667	20.5102	0.2228	-8.8463	0

1 --> COLLINEARITY is detected by the test
 0 --> COLLINEARITY is not detected by the test
 * all coefficients have significant t-ratios
 R-square of y on all x: 0.9759

Las columnas VIF, TOL y W_i proporcionan la salida de diagnóstico para el factor de inflación de la variancia, la tolerancia y prueba F respectivamente. El estadístico F para VARI es bastante alto (8207.9069) seguido por Vgreen, a más de los grados de libertad serán la causa de la multicolinealidad, como se observó en la matriz de correlación. A pesar de existir colinealidad entre algunas variables todos los coeficientes de este modelo tienen significancia y su bondad de ajuste es alta permitiendo ser un modelo predictor de HCV.

5.2.5. Modelos con una muestra de entrenamiento y una de test.

El conjunto de datos viene dado por el modelo ya calculado es decir por las 16 variables que aportan al mismo (NDVI, EVI, NDWI, VARI, Vgreen, $NDVI^2$, $VARI^2$, $Vgreen^2$, P15, P60, Tmax7, Tmax15, Tmedia7, Tmedia15, HRmax7, HRmin15) y la HCV_P que pertenece a la humedad de combustible vivo, todas estas observadas en 30 fechas.

Para poder evaluar la capacidad predictiva del modelo, se divide las observaciones disponibles en dos grupos: uno de entrenamiento para ajustar el modelo (86% de los datos) y uno de test (14% de los datos), es decir 26 y 4 datos respectivamente estos últimos datos han sido escogidos uno para cada estación del año, para invierno se tomó los datos del 30 de enero del 2017, para primavera se tomó los datos del 3 de mayo del 2017, para verano se tomó los datos del 31 de julio del 2017 y para otoño se tomó los datos del 27 de noviembre del 2017.

En primer lugar, se procede a realizar el ajuste del modelo incluyendo las 16 variables como predictores y utilizando los datos de entrenamiento.

Residuals:

Min	1Q	Median	3Q	Max
-5.6737	-1.0616	-0.0058	1.2706	4.5233

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-5.521e+02	8.014e+01	-6.889	7.16e-05	***
NDVI	2.536e+03	5.527e+02	4.589	0.001311	**
EVI	-6.445e+01	8.997e+00	-7.164	5.29e-05	***
NDWI	-1.378e+03	3.104e+02	-4.440	0.001624	**
VARI	-4.414e+03	3.269e+03	-1.350	0.209848	
Vgreen	5.936e+03	4.161e+03	1.427	0.187474	
I(NDVI^2)	-5.492e+03	5.156e+02	-10.652	2.11e-06	***
I(VARI^2)	-2.794e+04	1.305e+04	-2.141	0.060913	.
I(Vgreen^2)	3.960e+04	2.493e+04	1.588	0.146651	
P15	3.537e-01	8.436e-02	4.192	0.002334	**
P60	-2.323e-01	4.469e-02	-5.197	0.000566	***
Tmax7	-5.911e+00	9.335e-01	-6.332	0.000136	***
Tmax15	1.048e+01	2.324e+00	4.508	0.001471	**

Tmedia7	8.358e+00	1.545e+00	5.410	0.000427	***
Tmedia15	-1.341e+01	3.272e+00	-4.099	0.002679	**
HRmax7	-1.043e+00	3.139e-01	-3.322	0.008911	**
HRmin15	1.292e+00	3.720e-01	3.474	0.007007	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.891 on 9 degrees of freedom

Multiple R-squared: 0.9835, Adjusted R-squared: 0.9542

F-statistic: 33.58 on 16 and 9 DF, p-value: 4.46e-06

Se obtiene que el valor de R-cuadrada ajustada es 0.9542 lo que indica que el modelo es capaz de predecir la humedad de combustible vivo ponderado de las observaciones con las que se ha entrenado. El P-valor (4.46e-06) indica que el modelo en conjunto es significativo. Al comparar este modelo con respecto al modelo usando todos los datos como se muestra en el ítem 5.2., el valor de R-cuadrada ajusta apenas aumenta.

Para poder comparar los errores en los datos de entrenamiento y en el conjunto de datos tipo test, la estimación del error de predicción se obtendrá mediante el Error Medio Cuadrático (MSE). Primero se obtiene el MSE de las observaciones de entrenamiento un valor de 5.2418 y el MSE para las nuevas observaciones es decir el test es 128.1751. Se observa que el MSE del modelo es bajo cuando predice las mismas observaciones con las que se ha entrenado, siendo el MSE más grande al predecir nuevas observaciones, esto puede significar que la utilidad del modelo es cuestionable ya que el objetivo es predecir la humedad de combustible vivo de futuras muestra. Este modelo al presentar una bondad de ajuste mayor en comparación al modelo generado con IVeg, tiene un MSE mucho menor por lo que entrega más confiabilidad en las estimaciones.

Se recurre a la selección de predictores mediante stepwise selection empleando el criterio de información de Akaike (AIC) como criterio de evaluación, de esta manera se da una medida de calidad relativa del modelo estadístico, para el conjunto de datos; el número de predictores del modelo resultante es 17 es decir que se trabajara con todas las variables (NDVI, EVI, NDWI, VARI, Vigreen, NDVI², VARI², Vigreen², P15, P60, Tmax7, Tmax15, Tmedia7, Tmedia15, HRmax7, HRmin15) y la HCV_P, es decir con las mismas que se trabajó para el modelo de entrenamiento.

Se vuelve a sacar el MSE para el nuevo modelo generado por el proceso stepwise selection para los 17 predictores disponibles, el MSE de las observaciones de entrenamiento obteniendo un valor de 5.2418 y el MSE para las nuevas observaciones es decir el test es 128.1751; al no haber eliminado predictores los resultados son los mismos valores que se obtuvieron con el anterior modelo.

Se procede a ajustar el modelo empleando las componentes principales para las 16 variables analizadas, a continuación, se muestra la varianza explicada y acumulada de las 16 primeras componentes.

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
Standard deviation	2.132086	2.122405	1.786454	1.221323	1.024494	0.863213	0.5125174
Proportion of Variance	0.284110	0.281540	0.199460	0.093230	0.065600	0.046570	0.0164200
Cumulative Proportion	0.284110	0.565650	0.765110	0.858340	0.923940	0.970510	0.9869300
	PC8	PC9	PC10	PC11	PC12	PC13	
Standard deviation	0.3210313	0.2363611	0.1589532	0.1194501	0.08073959	0.04935024	
Proportion of Variance	0.0064400	0.0034900	0.0015800	0.0008900	0.00041000	0.00015000	
Cumulative Proportion	0.9933700	0.9968600	0.9984400	0.9993300	0.99974000	0.99989000	
	PC14	PC15	PC16				
Standard deviation	0.03254675	0.02562938	0.005149966				
Proportion of Variance	0.00007000	0.00004000	0.00000000				
Cumulative Proportion	0.99996000	1.00000000	1.00000000				

El estudio de la proporción de varianza explicada muestra que la primera componente recoge la mayor parte de la información (28.41%), decayendo drásticamente la varianza en las sucesivas componentes.

Una vez que se obtuvo el valor de las componentes para cada observación se ajusta el modelo lineal empleando dichos valores junto con la variable respuesta que le corresponde a cada observación. Según la proporción de varianza acumulada se emplea las 14 primeras componentes ya que en conjunto explica el 99.99% de varianza.

Se obtiene el MSE para el test es decir las nuevas observaciones (119.7799), para modelo que emplea como predictores las 14 primeras componentes, este valor es mayor que el obtenido con el modelo generado por stepwise selection y el obtenido para todos los predictores (5.2418), esto indica que emplear componentes principales como predictores no es útil para este caso, mientras que el MSE para los datos de entrenamiento es (202.4051) siendo también mayor en comparación a los modelos antes calculados.

Por último, se recurre a cross validation en los componentes principales para identificar el número óptimo de componentes con el que se minimizaría el MSE, siendo en este caso 17 es decir todas las variables utilizadas en el presente modelo, empleando este número en el PCR se consigue el test-MSE de 119.7799, siendo el mismo valor que el modelo calculado para componente principales.

La Tabla 20 muestra los valores predichos con los modelos antes explicados tanto para los datos de entrenamiento como para el test, los datos obtenidos con el MSE más bajo (128.1751) son los datos que tienen menor residual para los datos test caso contrario para los datos de entrenamiento los valores que tienen menor residual son los obtenidos en Stepwise Selection y Training obteniendo un MSE (5.2418).

Tabla 20. Comparación de modelos datos de entrenamiento y test

	Nº	Fecha toma	HCV_P	Modelo (Training)	Modelo (Stepwise Selection)	Modelo (PCR)	Modelo (PCR_CV)
Entrenamiento	1	15/1/2016	105.4850	105.91247	105.91247	105.91247	105.91247
	2	10/3/2016	91.6828	93.62856	93.62856	93.62856	93.62856
	3	29/3/2016	101.7750	97.65345	97.65345	97.65345	97.65345
	4	18/4/2016	98.5750	102.47839	102.47839	102.47839	102.47839
	5	16/5/2016	108.5872	108.98621	108.98621	108.98621	108.98621
	6	21/6/2016	107.0711	106.56511	106.56511	106.56511	106.56511
	7	19/7/2016	64.1345	64.22709	64.22709	64.22709	64.22709
	8	24/8/2016	60.1100	57.07192	57.07192	57.07192	57.07192
	9	5/10/2016	84.1725	83.34075	83.34075	83.34075	83.34075
	10	1/11/2016	84.1535	86.16441	86.16441	86.16441	86.16441
	11	2/12/2016	90.3325	88.58074	88.58074	88.58074	88.58074
	12	22/12/2016	101.1775	101.19442	101.19442	101.19442	101.19442
	13	21/2/2017	85.3635	85.67597	85.67597	85.67597	85.67597
	14	28/3/2017	82.9920	82.98663	82.98663	82.98663	82.98663
	15	22/5/2017	93.0545	90.76765	90.76765	90.76765	90.76765
	16	26/6/2017	86.1840	88.75388	88.75388	88.75388	88.75388
	17	12/9/2017	84.3225	89.99617	89.99617	89.99617	89.99617
	18	26/9/2017	84.7820	83.42147	83.42147	83.42147	83.42147
	19	27/11/2017	59.7600	54.42635	54.42635	54.42635	54.42635
	20	20/12/2017	84.4220	83.67574	83.67574	83.67574	83.67574
	21	7/3/2018	74.5535	77.5901	77.5901	77.5901	77.5901
	22	27/3/2018	57.5300	56.52912	56.52912	56.52912	56.52912
	23	22/4/2018	63.5705	64.72595	64.72595	64.72595	64.72595
	24	23/5/2018	124.5685	122.72894	122.72894	122.72894	122.72894
	25	12/6/2018	110.2055	105.68216	105.68216	105.68216	105.68216
	26	19/7/2018	111.8815	112.66146	112.66146	112.66146	112.66146
Test	27	30/1/2017	82.2540	89.4000	89.4000	90.0000	90.0000
	28	3/5/2017	82.7590	81.7000	81.7000	90.7000	90.7000
	29	31/7/2017	70.6100	74.1000	74.1000	89.4000	89.4000
	30	2/11/2017	54.7385	80.9000	80.9000	61.1000	61.1000

En la Imagen 34 se observa el comportamiento de los modelos calculados con datos de entrenamiento y test, el comportamiento de los 4 datos estimados de test muestra que no varía en gran cantidad a excepción de dato de la fecha 2/11/2017 existiendo una diferencia de 26.13 entre el dato real de HCV y el dato del modelo (Training), siendo las predicciones mejores que la realizada en el apartado 5.1.5, en el que se genera un modelo solo con índices de vegetación y su bondad de ajuste es menor.

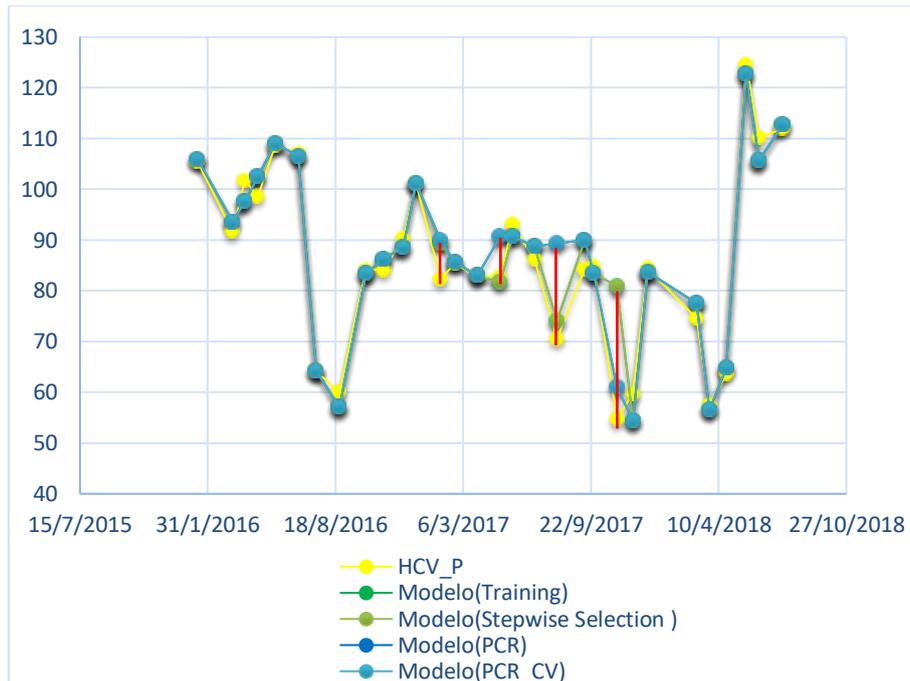


Imagen 34. Gráfica comparativa de modelos calculados por Regresión de Componentes Principales (PCR).

6. RESULTADOS

6.1. Modelos de regresión lineal múltiple.

6.1.1. Análisis entre la media ponderada de la humedad de combustible vivo y los índices de vegetación.

Al modelar la relación existente entre la media ponderada de HCV de las especies forestales más representativas de la Comunidad Valenciana y los índices de vegetación correspondientes a los 18 puntos de muestreo como se detalla en el apartado 5.1.1., se obtiene cuáles son las variables independientes es decir índices de vegetación que están relacionadas con la variable dependiente (HCV) obteniendo un proceso estadístico para estimar la relación entre ellas. Los valores R-cuadrada ajustada se utilizaron para evaluar la bondad del ajuste para las parcelas muestreadas, este estadístico redujo el efecto de variables explicativas adicionales que inflan automáticamente la métrica R^2 . La Tabla 21 muestra los modelos para cada punto de muestreo obteniendo los mejores resultados de R-cuadrada ajustada para Gandia (55.31%), Gilet (55.42%), Castielfabid (57.36%) y Muela Cortes (65.83%), este último con el R^2 ajustada mayor se debe a que en dicho punto de muestreo cuenta con menos observaciones y la variación depende de menos factores. Los valores de R-cuadrada ajustada no son tan altos por lo que las estimaciones no se ajustan bien a las variables reales en cada punto de muestreo. Los modelos que menos explican a las variables reales son Ayora y Alcoy con un 11.50 y 11.76% respectivamente, además en el punto de muestreo Chelva no se obtuvo ningún modelo estadísticamente significativo.

Los índices de vegetación NDVI y NDWI son los predictores que aparecen en una gran cantidad de los modelos calculados, el modelo del punto de muestreo Gandia es el que abarca más variables de índices de vegetación y es uno de los modelos que explica mejor a la variable real.

Tabla 21. Modelos de regresión múltiple para los 18 puntos de muestreo entre la media ponderada de HCV y las variables de índices de vegetación.

Nº	Punto de muestreo	R-Cuadrada ajustada	Modelo
1	Fredes	34.35	$HCV_P = 99.1683 + 2653.64 * VARI^2 - 5675.09 * Vigreen^2$
2	Puebla Tornesa	17.43	$HCV_P = 47110.4 - 155.576 * NDVI - 31340.7 * SAVI + 195.708 * NDWI^2$
3	Montanejos	27.83	$HCV_P = -168.218 - 434.668 * NDVI + 27.0365 * EVI - 1254.06 * NDWI - 701.687 * NDWI^2.$
4	Castielfabid	57.37	$HCV_P = -453.786 - 1465.8 * NDVI - 42.6144 * EVI - 4592.2 * NDWI + 1504.53 * VARI - 3798.49 * NDWI^2$
5	Chelva	No se obtuvo ningún modelo estadísticamente significativo.	
6	Villar del Arzobispo	32.57	$HCV_P = -113532. - 114.48 * NDVI + 50567.8 * SAVI^2 - 1719.1 * VARI^2$
7	Betera	42.67	$HCV_P = -9.22598 + 71.2763 * EVI - 563.873 * Vigreen$
8	Gilet	55.43	$HCV_P = -203.313 - 6460.15 * NDVI + 462.019 * EVI - 6423.57 * NDWI + 4476.39 * Vigreen - 164.795 * EVI^2 - 7930.79 * Vigreen^2$
9	Villargordo del Cabriel	14.58	$HCV_P = 78.9771 + 2313.06 * NDVI + 2295.16 * NDWI - 1661.88 * Vigreen$
10	Buñol	30.55	$HCV_P = -137921. + 1094.19 * NDWI - 736.902 * VARI + 1174.62 * NDVI^2 + 61493.6 * SAVI^2$
11	Llombai	21.61	$HCV_P = 98.3957 + 18423.2 * VARI^2 - 48120.8 * Vigreen^2$
12	Cortes de Pallás	65.83	$HCV_P = -718078. + 478957. * SAVI + 25948.4 * VARI - 40421.4 * Vigreen - 271166. * VARI^2 + 639260. * Vigreen^2$
13	Navarrés	34.10	$HCV_P = 87.8385 + 88.0028 * VARI + 1115.95 * VARI^2$
14	Ayora	11.51	$HCV_P = 177.466 + 1703.61 * NDVI + 1845.64 * NDWI - 1798.09 * NDVI^2 + 1721.39 * NDWI^2 + 2171.64 * VARI^2 - 3026.41 * Vigreen^2$

Nº	Punto de muestreo	R-Cuadrada ajustada	Modelo
15	Gandia	55.32	$HCV_P = -259010. - 4429.82*NDVI + 1020.18*EVI + 172842.*SAVI - 1787.13*NDWI + 1566.86*VARI + 2086.67*NDVI^2 - 346.129*EVI^2 - 4817.29*VARI^2$
16	Muro de Alcoi	11.77	$HCV_P = 34430.6 - 22911.6*SAVI$
17	Altea	30.22	$HCV_P = 39.9501 + 26.3182*EVI - 286.391*VARI + 235.865*Vigreen + 56.43*NDVI^2$
18	Biar	50.92	$HCV_P = -223.675 - 2411.18*NDVI + 545.982*EVI - 2393.38*NDWI + 1613.16*Vigreen - 219.709*EVI^2$

La Imagen 35 muestra los resultados de R-cuadrada ajustada para los modelos de regresión lineal múltiple calculados para relacionar la media ponderada de HCV y las variables de índices de vegetación.

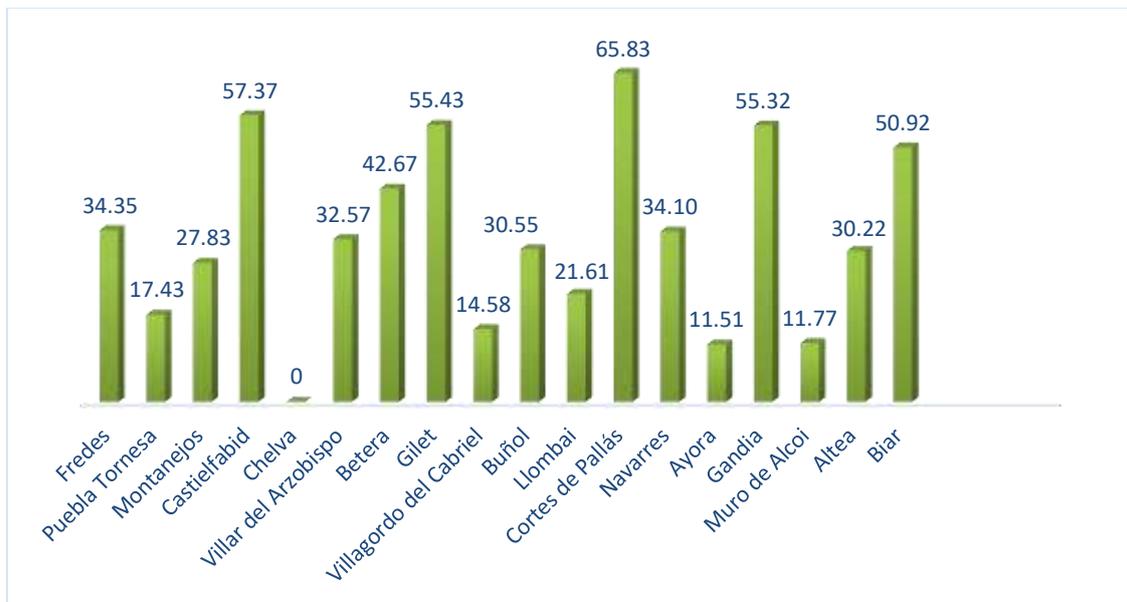


Imagen 35. Gráfica de resultados de R-cuadrada ajustada para los modelos relacionados entre HCV e IVeg.

A continuación, se presenta el ajuste de los modelos calculados en relación a los valores reales de Humedad de combustible vivo, en los puntos de muestreo que presentan un R-cuadrada ajustada por encima del 50%.

CATIELFABID

En el modelo del punto de muestreo Castielfabid se obtiene un R-cuadrada ajustada de 57.37%, en la Imagen 36 se observa que los datos estimados por el modelo se ajustan en su mayoría a la tendencia de los datos reales. La estimación del dato de HCV para julio del 2017 es el único que difiere de la tendencia, este desfase se debe a que los índices de vegetación cambian de una manera considerable en comparación con las otras fechas de muestreo perteneciendo este a verano, entregándonos así el modelo calculado un dato fuera de rango.

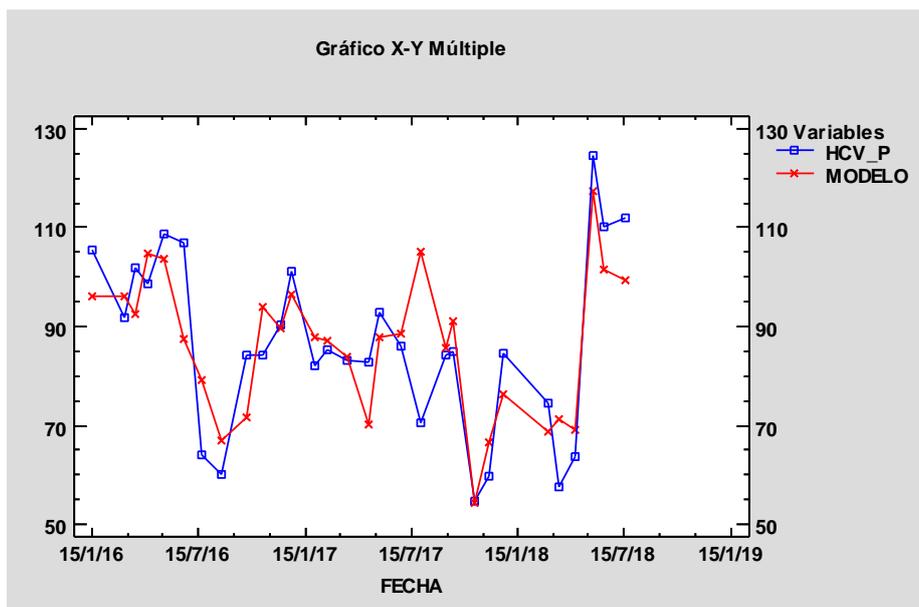


Imagen 36. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Castielfabid.

GILET

En el modelo del punto de muestreo Gilet se obtiene un R-cuadrada ajustada de 55.43%, en la Imagen 37 se observa que el modelo no recoge las oscilaciones alrededor de julio del 2016, esto ocurre ya que el muestreo y cálculo de HCV para verano se realizó en dos fechas del mismo mes, teniendo una diferencia en la recogida de especies forestales ya que en los últimos días de julio, julio y agosto no se obtuvo la especie *Erica multiflora* y *Pistacea lentiscus*, esto provocó que el cálculo de la media ponderada de HCV baje por la falta de dichas especies. En los valores máximos de primavera del 2017 el modelo tampoco se ajusta a los valores de HCV, sin embargo, los datos siguen la tendencia.

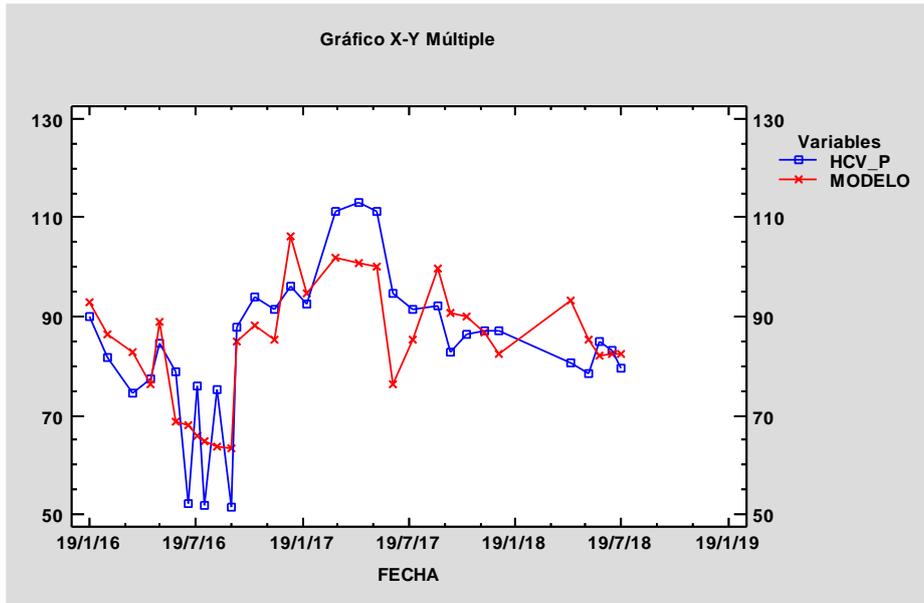


Imagen 37. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Gilet.

CORTES DE PALLÁS

En la Imagen 38 se puede observar que este punto de muestreo cuenta con una menor cantidad de datos a diferencia de los otros, ya que el muestreo por parte de la empresa VAERSA empezó en marzo del 2017, por lo cual se tiene menos observaciones para el análisis y la variación de HCV depende de menos factores, teniendo así un R-cuadrada ajustada de 65.83% siendo el mayor de todos por el motivo antes expuesto.

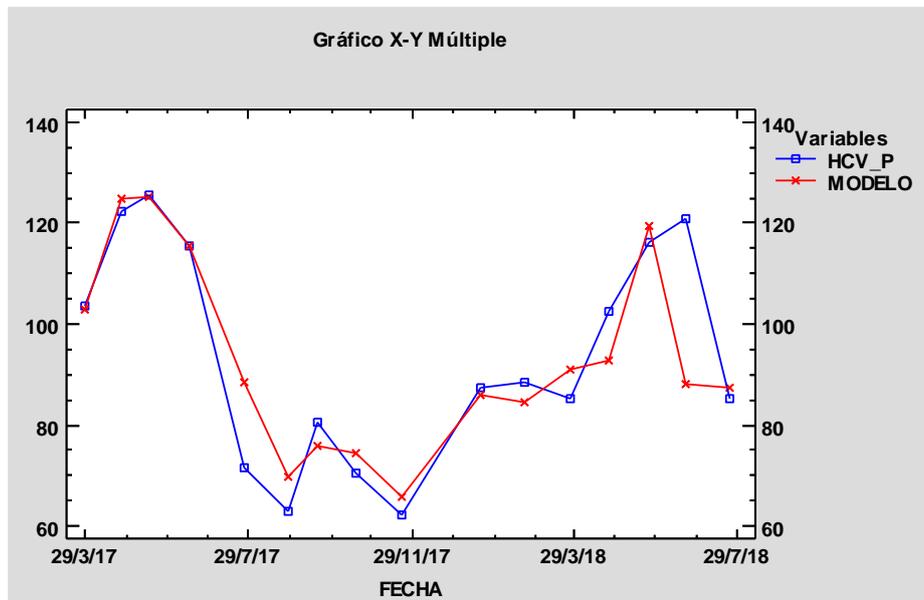


Imagen 38. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Cortes de Pallás.

GANDIA

En el modelo del punto de muestreo Gandia se obtiene un R-cuadrada ajustada de 55.32%, en la Imagen 39 se muestra que el modelo se ajusta a la tendencia de los datos de HCV a excepción de los valores altos presentes en primavera del 2016 ya que existe un cambio drástico de valores de HCV entre primavera y verano.

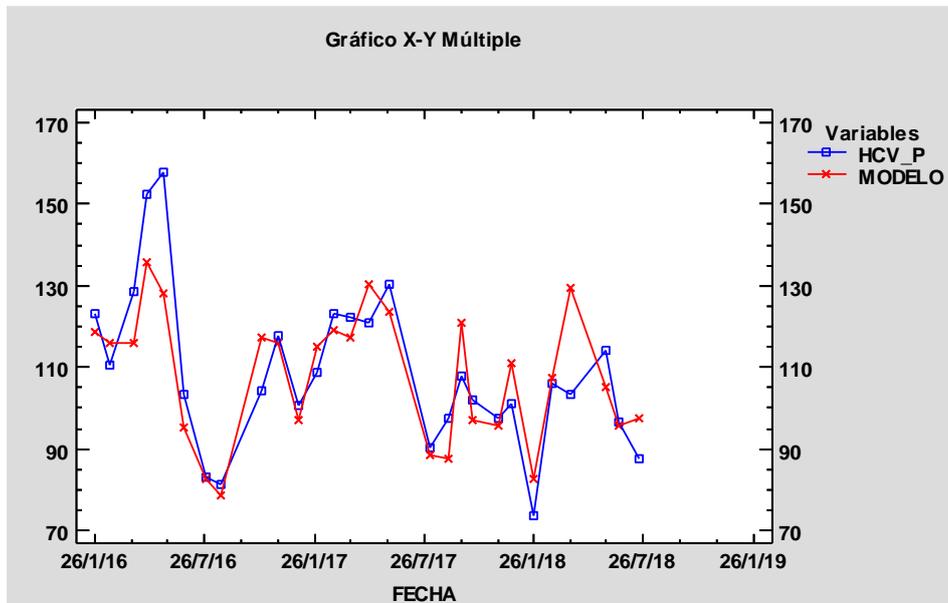


Imagen 39. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Gandia.

BIAR

En el modelo del punto de muestreo Biar se obtiene un R-cuadrada ajustada de 50.92%, en la Imagen 40 se observa que el modelo en el lapso del año 2017 tiene un comportamiento casi plano, esto se debe a que los datos de índice de vegetación en este punto de muestreo no tienen una variación considerable entre las diferentes fechas. A pesar de que el modelo es mayor a 50% el ajuste no es bueno.

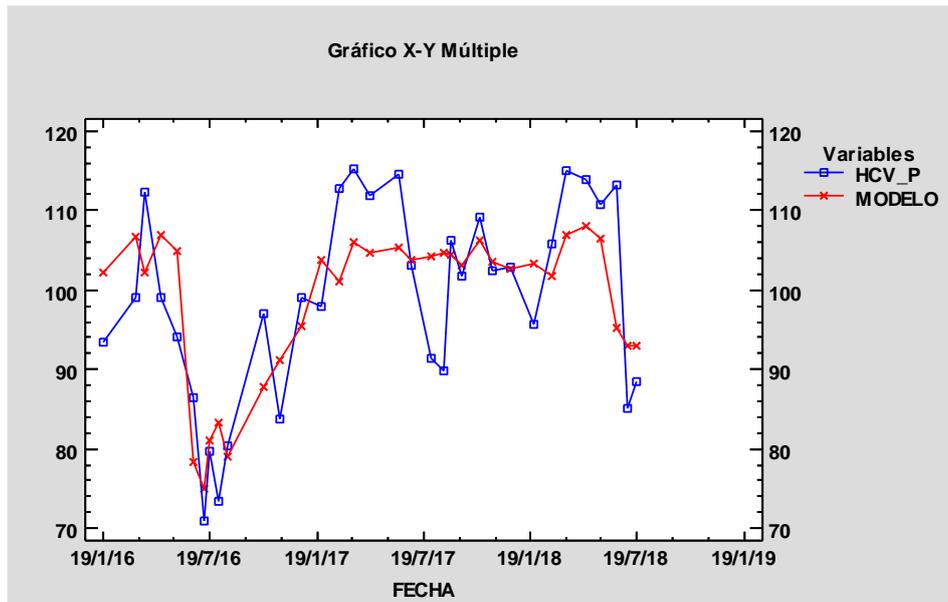


Imagen 40. Gráfica ajuste del modelo para índices de vegetación en el punto de muestro Biar.

6.1.2. Análisis entre la media ponderada de la humedad de combustible vivo y los índices de vegetación más las variables meteorológicas.

En este apartado se muestra los resultados de modelar la relación existente entre la media ponderada de HCV, los índices de vegetación y las variables meteorológicas correspondientes a los 18 puntos de muestreo como se detalla en el apartado 5.2. La Tabla 22 se muestra que en todos los casos el R-cuadrada ajustada son mayores al 90% a excepción de Villargordo de Cabriel y Puebla Tornesa, con R-cuadrada ajustada de 56.9597 y 84.2069% respectivamente. Los valores de la bondad de ajuste son altos por lo que las estimaciones se ajustan bien a las variables reales en cada punto de muestreo a excepción de Villargordo de Cabriel.

Las variables predictoras que contribuyen significativamente a los modelos pertenecen al grupo de índices de vegetación y a las meteorológicas en los 18 puntos de muestreo. Siendo las variables meteorológicas las que contribuyen significativamente y existiendo por lo menos una variable de índice de vegetación como es el caso de Villargordo de Cabriel y Muro de Alcoi.

Tabla 22. Modelos de regresión múltiple para los 18 puntos de muestreo entre la media ponderada de HCV y las variables de índices de vegetación más meteorológicas.

Nº	Punto de muestreo	R-Cuadrada ajustada	Modelo
1	Fredes	94.0766	$\text{HCV}_P = 591.429 + 243.316 \cdot \text{EVI} + 1686.85 \cdot \text{NDWI} - 5849.3 \cdot \text{VARI} + 8724.97 \cdot \text{Vigreen} - 86.3054 \cdot \text{EVI}^2 + 1074.27 \cdot \text{NDWI}^2 + 22039.1 \cdot \text{VARI}^2 - 46933.5 \cdot \text{Vigreen}^2 - 0.00538665 \cdot \text{P60}^2 + 1.28201 \cdot \text{P60} - 7.18817 \cdot \text{Tmax7} + 1.13933 \cdot \text{Tmax15} + 5.80864 \cdot \text{Tmedia7} + 0.48229 \cdot \text{HRmax3} - 0.810022 \cdot \text{HRmin3} - 1.16369 \cdot \text{HRmin15}$
2	Puebla Tornesa	84.2069	$\text{HCV}_P = 773.938 - 2893.35 \cdot \text{NDVI} + 104.291 \cdot \text{EVI} + 837.964 \cdot \text{VARI} + 330.576 \cdot \text{Vigreen} - 65.2734 \cdot \text{EVI}^2 + 2793.29 \cdot \text{NDWI}^2 - 3086.98 \cdot \text{Vigreen}^2 - 2.32521 \cdot \text{HRmin7} + 3.23541 \cdot \text{HRmin15} + 3.7622 \cdot \text{Tmax3} - 11.7082 \cdot \text{Tmax7} + 8.09774 \cdot \text{Tmax30} + 3.80564 \cdot \text{lev3} - 4.99935 \cdot \text{lev15} + 4.63319 \cdot \text{Vmax3} - 5.70762 \cdot \text{Vmax7}$
3	Montanejos	96.516	$\text{HCV}_P = -415.212 + 816.384 \cdot \text{NDVI} - 151.709 \cdot \text{EVI} - 1902.15 \cdot \text{NDWI} + 838.191 \cdot \text{VARI} - 2415.94 \cdot \text{NDVI}^2 + 60.9133 \cdot \text{EVI}^2 - 7860.54 \cdot \text{VARI}^2 + 15799.7 \cdot \text{Vigreen}^2 + 0.246702 \cdot \text{P7} - 0.149824 \cdot \text{P60} + 5.78903 \cdot \text{Tmax3} + 4.89996 \cdot \text{Tmax15} - 8.83328 \cdot \text{Tmax30} - 7.70082 \cdot \text{Tmedia7} + 4.46002 \cdot \text{Tmedia15} - 2.3653 \cdot \text{HRmax7} + 2.96072 \cdot \text{HRmin7} - 1.24027 \cdot \text{HRmin15} - 22.2272 \cdot \text{lev15} - 0.839998 \cdot \text{pto}_r15$
4	Castielfabid	94.6268	$\text{HCV}_P = -613.364 + 2997.81 \cdot \text{NDVI} - 69.733 \cdot \text{EVI} - 1170.95 \cdot \text{NDWI} - 7633.03 \cdot \text{VARI} + 10242.8 \cdot \text{Vigreen} - 5832.32 \cdot \text{NDVI}^2 - 40137.3 \cdot \text{VARI}^2 + 65169.7 \cdot \text{Vigreen}^2 + 0.30657 \cdot \text{P15} - 0.219474 \cdot \text{P60} - 7.00221 \cdot \text{Tmax7} + 8.71042 \cdot \text{Tmax15} + 8.81067 \cdot \text{Tmedia7} - 10.44 \cdot \text{Tmedia15} - 0.818569 \cdot \text{HRmax7} + 1.30837 \cdot \text{HRmin15}$
5	Chelva	98.4644	$\text{HCV}_P = -52825.6 - 2803.56 \cdot \text{NDVI} + 129.006 \cdot \text{EVI} - 2888.36 \cdot \text{NDWI} - 1204.08 \cdot \text{VARI} + 3212.67 \cdot \text{Vigreen} - 37.2666 \cdot \text{EVI}^2 + 23416.5 \cdot \text{SAVI}^2 + 9505.06 \cdot \text{VARI}^2 - 22450.3 \cdot \text{Vigreen}^2 + 2.69504 \cdot \text{P3} - 0.18537 \cdot \text{P30} + 0.123776 \cdot \text{P60} - 3.69556 \cdot \text{Tmax3} + 22.2363 \cdot \text{Tmax7} - 10.4002 \cdot \text{Tmax15} - 2.82673 \cdot \text{Tmax30} + 6.96497 \cdot \text{Tmedia3} - 33.7836 \cdot \text{Tmedia7} + 19.4467 \cdot \text{Tmedia15} - 0.554443 \cdot \text{HRmax3} - 3.58454 \cdot \text{HRmax7} + 5.59042 \cdot \text{HRmax15} - 1.08522 \cdot \text{HRmin3} + 5.29055 \cdot \text{HRmin7} - 4.6204 \cdot \text{HRmin15} - 4.00613 \cdot \text{Vmax3} + 2.4444 \cdot \text{Vmax7} + 2.49861 \cdot \text{Vmax15}$
6	Villar del Arzobispo	99.2345	$\text{HCV}_P = -210253. + 2016.27 \cdot \text{NDVI} - 90.043 \cdot \text{EVI} + 140526. \cdot \text{SAVI} + 2695.65 \cdot \text{NDWI} - 1895.47 \cdot \text{Vigreen} + 596.841 \cdot \text{NDWI}^2 + 19135.7 \cdot \text{VARI}^2 - 46351.6 \cdot \text{Vigreen}^2 - 2.53904 \cdot \text{P3} + 0.57347 \cdot \text{P15} - 0.427613 \cdot \text{P60} + 5.31457 \cdot \text{Tmax3} + 5.80078 \cdot \text{Tmax7} - 16.4302 \cdot \text{Tmax30} - 22.9291 \cdot \text{Tmedia7} + 24.8568 \cdot \text{Tmedia15} + 0.698997 \cdot \text{HRmax3} - 4.98 \cdot \text{HRmax7} + 5.30443 \cdot \text{HRmax15} - 0.957261 \cdot \text{HRmin3} + 5.28483 \cdot \text{HRmin7} - 5.23256 \cdot \text{HRmin15} - 0.815138 \cdot \text{Vmax7}$

Nº	Punto de muestreo	R-Cuadrada ajustada	Modelo
7	Betera	99.9978	$\begin{aligned} \text{HCV_P} = & 175.232 - 10.3061*\text{EVI} + 202.131*\text{Vigreen} + 1.14183*\text{P3} - \\ & 3.60919*\text{P7} - 0.232195*\text{P15} + 0.695189*\text{P30} + 0.872188*\text{P60} - \\ & 7.57112*\text{Tmax7} - 33.1225*\text{Tmax15} + 12.2891*\text{Tmax30} + \\ & 7.92936*\text{Tmedia3} - 5.24649*\text{Tmedia7} + 20.3231*\text{Tmedia15} - \\ & 1.17896*\text{HRmax3} + 11.94*\text{HRmax7} - 8.13417*\text{HRmax15} - \\ & 3.68112*\text{HRmin3} + 0.32477*\text{HRmin7} + 1.16338*\text{HRmin15} + \\ & 0.695701*\text{Vmax3} \end{aligned}$
8	Gilet	95.8278	$\begin{aligned} \text{HCV_P} = & -232779. + 41191.1*\text{NDVI} + 33094.5*\text{NDWI} - \\ & 71507.6*\text{Vigreen} + 56074.6*\text{NDVI}^2 + 102305.*\text{SAVI}^2 - \\ & 63127.0*\text{NDWI}^2 + 145974.*\text{VARI}^2 - 252171.*\text{Vigreen}^2 - \\ & 3.71221*\text{P3} - 2.24314*\text{P7} + 6.65885*\text{P15} - 4.27972*\text{P30} + \\ & 0.199029*\text{P60} - 7.94008*\text{Tmax3} + 80.0087*\text{Tmax15} - 41.468*\text{Tmax30} - \\ & 14.9297*\text{Tmedia3} - 7.04845*\text{Tmedia15} - 5.58904*\text{HRmax3} + \\ & 1.69544*\text{HRmax7} + 3.83039*\text{HRmax15} + 6.42342*\text{HRmin3} + \\ & 3.8979*\text{HRmin7} - 4.64583*\text{HRmin15} + 1.4725*\text{Vmax3} \end{aligned}$
9	Villargordo del Cabriel	56.9597	$\begin{aligned} \text{HCV_P} = & 35.2851 - 63.3356*\text{NDWI} + 0.116366*\text{P60} - \\ & 0.397519*\text{Tmedia3} + 0.189735*\text{HRmax7} \end{aligned}$
10	Buñol	91.6147	$\begin{aligned} \text{HCV_P} = & -242980. - 261.327*\text{EVI} - 1642.37*\text{NDWI} + 3067.9*\text{VARI} - \\ & 2481.38*\text{NDVI}^2 + 108125.*\text{SAVI}^2 - 1.63297*\text{P7} - 0.767136*\text{P15} + \\ & 0.470151*\text{P30} + 7.95651*\text{Tmax3} - 12.3178*\text{Tmax7} - 23.1503*\text{Tmax15} + \\ & 6.20419*\text{Tmax30} + 16.2012*\text{Tmedia15} - 6.82465*\text{HRmax3} + \\ & 8.87034*\text{HRmax7} + 3.01131*\text{HRmax15} + 8.23832*\text{HRmin3} - \\ & 5.6655*\text{HRmin7} - 9.80253*\text{HRmin15} + 9.94758*\text{Vmax15} + \\ & 8.94385*\text{lev3} - 25.7627*\text{lev7} + 30.009*\text{lev15} - 67.6201*\text{pto_r3} + \\ & 35.427*\text{pto_r7} - 25.7661*\text{pto_r15} + 26.2958*\text{pto_r30} \end{aligned}$
11	Llombai	99.052	$\begin{aligned} \text{HCV_P} = & -290.623 + 76.3654*\text{EVI} - 15035.8*\text{VARI}^2 + \\ & 36032.5*\text{Vigreen}^2 - 7.45696*\text{P3} + 4.74203*\text{P7} - 7.28112*\text{P15} + \\ & 3.32652*\text{P30} + 232.559*\text{Tmax3} - 506.187*\text{Tmax15} + 168.716*\text{Tmax30} - \\ & 223.369*\text{Tmedia3} - 62.5412*\text{Tmedia7} + 396.001*\text{Tmedia15} + \\ & 1.79692*\text{HRmax3} - 6.93677*\text{HRmax7} + 8.83389*\text{HRmax15} + \\ & 3.47818*\text{HRmin3} + 3.53892*\text{HRmin7} - 8.59467*\text{HRmin15} + \\ & 13.9508*\text{Vmax3} - 18.331*\text{Vmax7} + 14.9481*\text{Vmax15} + 6.72987*\text{lev3} - \\ & 15.7779*\text{lev7} + 11.8165*\text{lev15} \end{aligned}$
12	Cortes de Pallás	99.7406	$\begin{aligned} \text{HCV_P} = & -650275. + 433700.*\text{SAVI} + 23093.2*\text{VARI} - \\ & 33985.7*\text{Vigreen} - 269521.*\text{VARI}^2 + 577856.*\text{Vigreen}^2 + \\ & 0.28237*\text{P30} - 1.38773*\text{Tmedia7} - 0.68297*\text{HRmax3} + \\ & 1.20607*\text{HRmax15} + 1.0386*\text{Vmax3} \end{aligned}$

Nº	Punto de muestreo	R-Cuadrada ajustada	Modelo
13	Navarrés	96.2359	$\begin{aligned} \text{HCV_P} = & -197500. + 1884.09*\text{NDVI} - 297.391*\text{EVI} + 132045.*\text{SAVI} + \\ & 2474.5*\text{NDWI} - 594.152*\text{VARI} + 4923.45*\text{Vigreen} - 6589.97*\text{NDVI}^2 \\ & + 83.2641*\text{EVI}^2 + 7029.48*\text{NDWI}^2 + 14325.9*\text{VARI}^2 - \\ & 25925.6*\text{Vigreen}^2 + 0.397411*P3 - 0.219232*P15 + 0.0832265*P60 + \\ & 6.87538*Tmax7 - 14.4*Tmax15 + 6.60519*Tmax30 + 1.3256*Tmedia3 \\ & - 7.90168*Tmedia7 + 6.99053*Tmedia15 + 0.233993*HRmax3 - \\ & 0.951953*HRmax7 + 0.34141*HRmax15 + 0.975358*HRmin7 + \\ & 1.07304*Vmax7 \end{aligned}$
14	Ayora	99.6835	$\begin{aligned} \text{HCV_P} = & 1441.61 + 9002.59*\text{NDVI} - 35.0722*\text{EVI} + 12944.5*\text{NDWI} + \\ & 4672.87*\text{VARI} - 4218.17*\text{Vigreen} - 8315.98*\text{NDVI}^2 + \\ & 7.44507*\text{EVI}^2 + 11019.7*\text{NDWI}^2 + 20150.6*\text{VARI}^2 - \\ & 20929.4*\text{Vigreen}^2 - 2.00276*P3 - 0.473573*P7 - 0.381472*P15 - \\ & 0.0660117*P30 + 0.300597*P60 + 1.39717*Tmax3 - 1.50969*Tmax7 - \\ & 24.6689*Tmax15 - 7.74815*Tmax30 - 20.7381*Tmedia7 + \\ & 55.4997*Tmedia15 + 0.531363*HRmax3 + 3.27033*HRmax7 + \\ & 1.32986*HRmax15 + 1.9429*HRmin3 - 3.40369*HRmin7 - \\ & 1.71459*HRmin15 \end{aligned}$
15	Gandia	98.1694	$\begin{aligned} \text{HCV_P} = & -704488. - 2549.11*\text{NDVI} + 39.8504*\text{EVI} + 469671.*\text{SAVI} - \\ & 3704.85*\text{NDWI} + 1974.29*\text{VARI} - 1372.57*\text{NDVI}^2 - \\ & 1847.73*\text{VARI}^2 + 13.3678*P3 - 5.24379*P15 + 0.727045*P60 - \\ & 358.546*Tmax3 + 576.51*Tmax7 - 331.341*Tmax15 + \\ & 288.882*Tmedia3 - 504.207*Tmedia7 + 341.133*Tmedia15 + \\ & 4.32423*HRmax3 - 2.24892*HRmax7 - 1.12283*HRmax15 - \\ & 4.43201*HRmin3 + 4.47259*HRmin7 \end{aligned}$
16	Muro de Alcoi	97.2314	$\begin{aligned} \text{HCV_P} = & 52540.6 - 34956.5*\text{SAVI} + 0.403262*P7 + 0.777119*P15 - \\ & 0.663054*P30 + 19.5213*Tmax3 + 10.2119*Tmax7 - 39.4156*Tmax15 \\ & - 6.90537*Tmax30 - 31.2249*Tmedia3 + 49.0199*Tmedia15 + \\ & 1.75698*HRmax3 - 6.09516*HRmax7 + 4.93807*HRmax15 + \\ & 0.760057*HRmin3 + 4.40909*HRmin7 - 5.38745*HRmin15 + \\ & 1.70399*Vmax3 - 3.32055*Vmax7 - 4.29693*lev3 + 17.4262*lev7 - \\ & 9.03871*lev15 - 5.26171*pto_r7 \end{aligned}$
17	Altea	99.4039	$\begin{aligned} \text{HCV_P} = & 77.8782 + 11.4393*\text{EVI} + 131.53*\text{Vigreen} - \\ & 75.3547*\text{NDVI}^2 + 0.299481*P3 - 1.09142*P7 + 1.65749*P15 - \\ & 0.521858*P30 + 0.0617653*P60 - 3.68942*Tmax3 + 17.2743*Tmax7 - \\ & 7.65312*Tmax30 - 5.13551*Tmedia3 - 3.96305*Tmedia15 - \\ & 0.828238*HRmax7 + 2.27547*HRmax15 + 2.07979*HRmin3 - \\ & 1.0309*HRmin7 - 3.68029*HRmin15 - 16.3107*Vmax3 + \\ & 14.6089*Vmax7 + 8.27495*Vmax15 - 3.1331*lev3 + 3.30565*lev7 - \\ & 330.939*pto_r3 - 107.107*pto_r7 + 130.097*pto_r15 - 15.354*pto_r30 \end{aligned}$

Nº	Punto de muestreo	R-Cuadrada ajustada	Modelo
18	Biar	99.0199	$\begin{aligned} \text{HCV}_P = & -53.1496 + 2496.24 \cdot \text{NDVI} - 406.952 \cdot \text{EVI} + 2477.61 \cdot \text{NDWI} \\ & - 1278.51 \cdot \text{VARI} + 190.92 \cdot \text{EVI}^2 - 2105.02 \cdot \text{VARI}^2 - 1.28591 \cdot \text{P3} + \\ & 1.16524 \cdot \text{P15} - 0.358376 \cdot \text{P30} - 12.6151 \cdot \text{Tmax3} + 20.9159 \cdot \text{Tmax7} + \\ & 6.92527 \cdot \text{Tmax15} - 3.04196 \cdot \text{Tmax30} + 7.92669 \cdot \text{Tmedia3} - \\ & 19.7015 \cdot \text{Tmedia7} + 1.89378 \cdot \text{HRmax3} - 2.84261 \cdot \text{HRmax7} + \\ & 3.76154 \cdot \text{HRmax15} - 2.0129 \cdot \text{HRmin3} + 2.2981 \cdot \text{HRmin7} - \\ & 1.41955 \cdot \text{Vmax3} + 5.36809 \cdot \text{Vmax7} + 1.87407 \cdot \text{Vmax15} - 3.18745 \cdot \text{lev7} \\ & - 1.20105 \cdot \text{lev15} \end{aligned}$

La Imagen 41 muestra los resultados de R-cuadrada ajustada para los modelos de regresión lineal múltiple calculados para relacionar la media ponderada de HCV y las variables de índices de vegetación más las variables meteorológicas. El mejor modelo en este caso es el obtenido para el punto de muestreo Betera el cual llega a un R-cuadrada ajustada de 99.998%, el cual se ajusta con 20 variables predictoras, 2 pertenecientes a índices de vegetación (EVI y VIgreen) y 18 variables meteorológicas (P3, P7, P15, P30, P60, Tmax7, Tmax15, Tmax30, Tmedia3, Tmedia7, Tmedia15, HRmax3, HRmax7, HRmax15, HRmin3, HRmin7, HRmin15, Vmax3) pertenecientes al grupo de precipitación, temperatura, humedad relativa y viento.

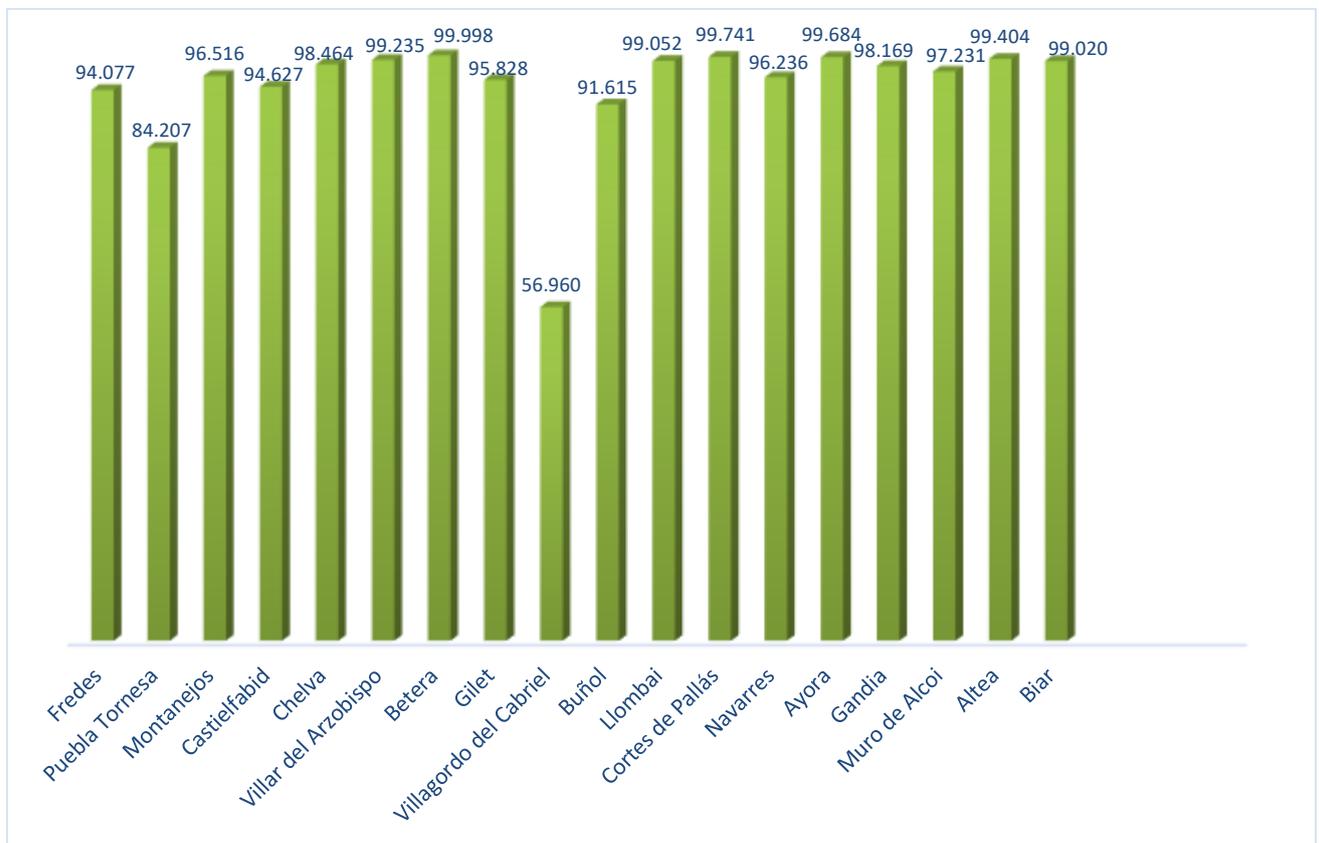


Imagen 41. Gráfica de resultados de R-cuadrada ajustada para los modelos relacionados entre HCV, IV y variables meteorológicas.

A continuación, se presenta el ajuste de los modelos calculados en relación a los valores reales de Humedad de combustible vivo en algunos puntos de muestreo.

PUEBLA TORNESA

En el modelo del punto de muestreo Puebla Tornesa se obtiene un R-cuadrada ajustada de 84.20%, estando por debajo del 90% como muestra la mayoría de puntos de muestreo, en la Imagen 42 se observa que el modelo no se ajusta totalmente a los datos de HCV a pesar que sigue la tendencia de los datos; en verano del 2017 se muestra un pico de descenso de HCV haciendo que el modelo pierda el ajuste en dicho dato.

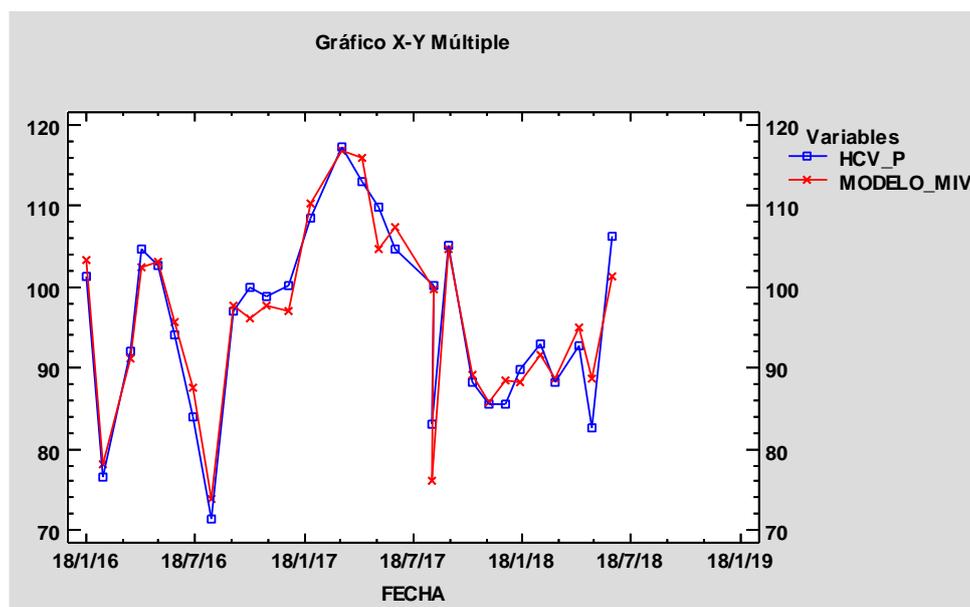


Imagen 42. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Puebla Tornesa.

CHELVA

El análisis de la relación del punto de muestreo Chelva entre HCV e IVeg como se muestra en la Tabla 21, no obtuvo ningún modelo estadísticamente significativo, al incluir en los análisis variables meteorológicas vuelven a ingresar todas las variables independientes y hace que los datos de índices de vegetación sean significativos y aporten al modelo. En la Imagen 43 se observa como en los meses de verano del 2016 se producen oscilaciones entre las medidas de HCV ya que los muestreos se realizaban dos veces por mes existiendo diferencias entre ellas, el modelo se logró ajustar con un R-cuadrada ajustada de 98.46%, a pesar de las oscilaciones el modelo en los meses de julio, julio y agosto del 2016, se adapta a los valores ya que las variables meteorológicas actúan de la misma manera para dichas fechas, la precipitación acumulada para los últimos tres días sube y baja al igual que la HCV y de la misma manera las temperaturas máximas y medias, teniendo una relación muy importante entre las variables.

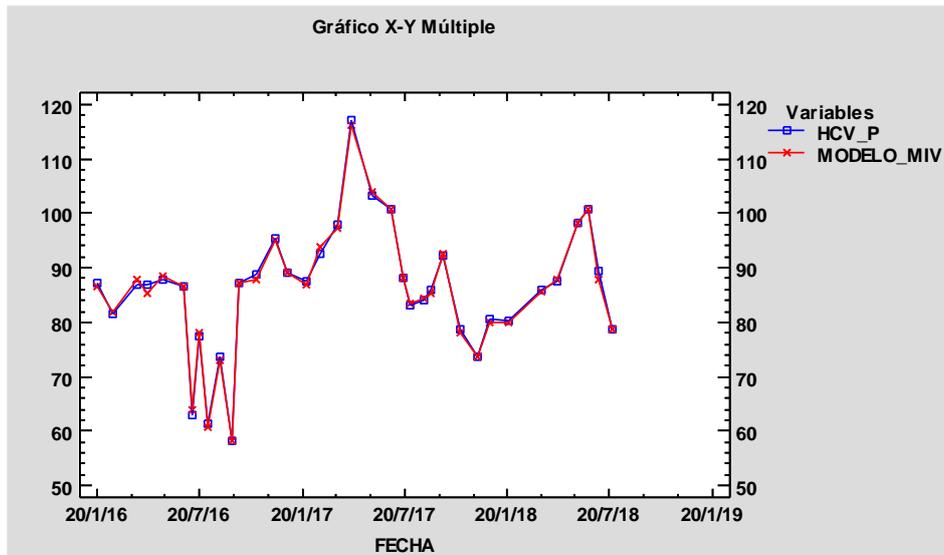


Imagen 43. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Chelva.

GILET

En el modelo del punto de muestreo Gilet se obtiene un R-cuadrada ajustada de 95.82%, al igual que el punto de muestreo Chelva antes explicado en los meses de verano se produce las oscilaciones por el mismo motivo, a diferencia de este, en Gilet las variables meteorológicas de precipitación, temperatura, humedad relativa y viento tienen el mismo comportamiento que la HCV haciendo que el modelo se ajuste en la mayoría de datos y sigan la misma tendencia.

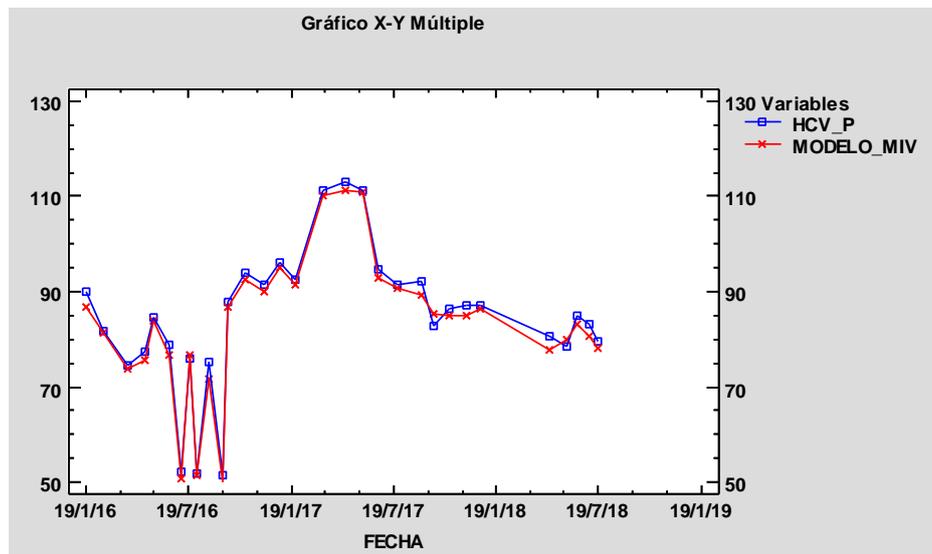


Imagen 44. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Gilet.

VILLARGORDO DEL CABRIEL

En el modelo del punto de muestreo Villargordo del Cabriel se obtiene un R-cuadrada ajustada de 56.96%, siendo el ajuste más bajo entre los 18 puntos de muestreo, se observa en la Imagen 45 que el HCV muestra unos picos muy destacados en verano haciendo que el modelo no se ajuste bien en este lapso, en los máximos de primavera también el modelo no alcanza a los valores de HCV, a pesar que el modelo sigue la tendencia en este rango, el modelo no se lo considera bueno.

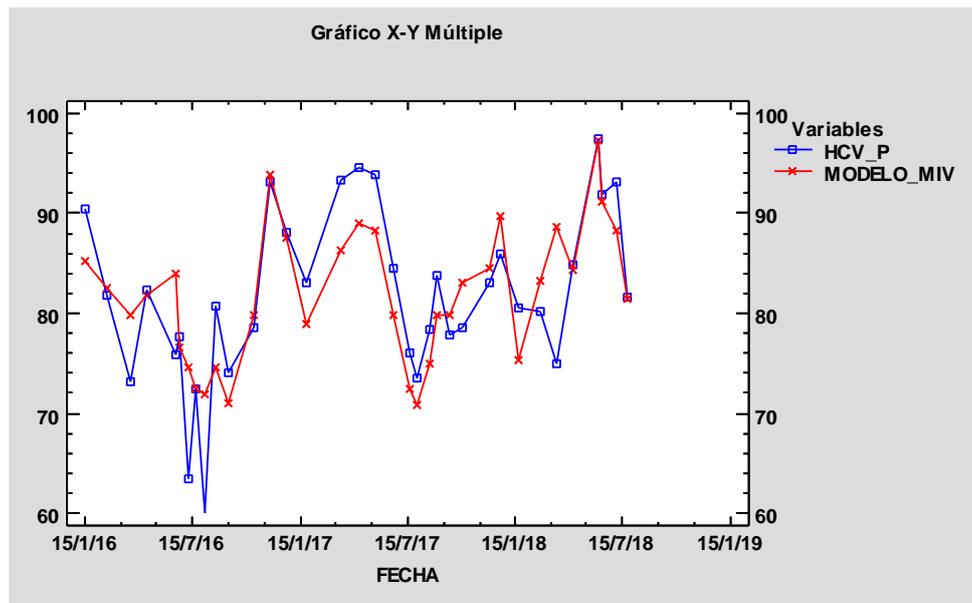


Imagen 45. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Villargordo del Cabriel.

BUÑOL

En el modelo del punto de muestreo Buñol se obtiene un R-cuadrada ajustada de 91.61%, en la Imagen 46 se muestra que en el transcurso de los tres años de análisis, en primavera hay picos relevantes en donde el porcentaje de HCV es alto, el modelo se ajusta bien en estos picos ya que las variables meteorológicas pertenecientes a punto del rocío aportan significativamente en los datos pertenecientes a estas fechas; estos datos toman valores distintos a cero entregando condensación para los últimos 3, 7, 15 y 30 días, aportando al modelo, reproduciendo estos máximos relativos en las tres fechas.

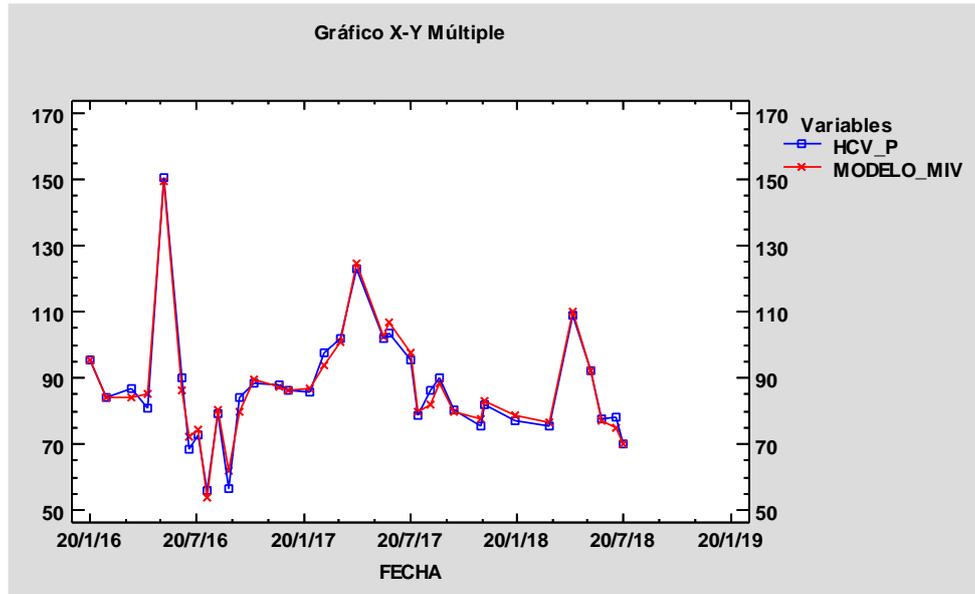


Imagen 46. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Buñol.

ALTEA

En el modelo del punto de muestreo Altea se obtiene un R-cuadrada ajustada de 99.40%, en la Imagen 47 se muestra que el modelo se ajusta casi perfectamente a los datos de HCV, los datos pertenecientes de enero a abril del 2017 hacen que la gráfica tenga un pico superior muy prominente, el ocasional del mismo son las lluvias que se registraron en este rango de tiempo, luego se observa el descenso del HCV también causado por las mismas precipitaciones. A partir de agosto del 2017 hasta primavera del 2018 se observa que existen oscilaciones en los datos, esto se debe a los cambios que existen en estas fechas en la humedad relativa y también en las precipitaciones, ajustado muy bien el modelo a los datos de HCV.

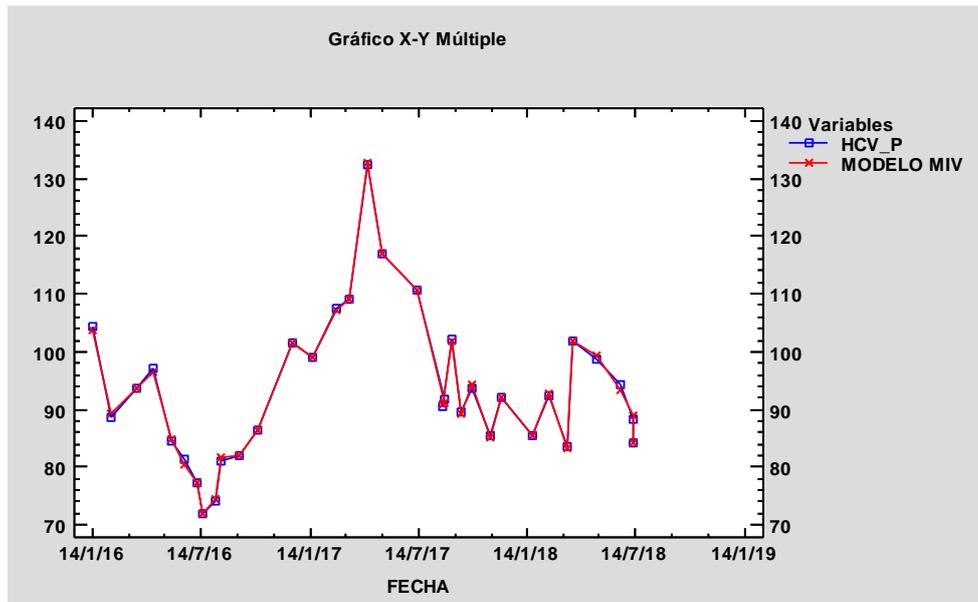


Imagen 47. Gráfica ajuste del modelo con índices de vegetación y variables meteorológicas en el punto de muestro Altea.

6.2. Estimación del valor de la media ponderada de HCV con los modelos de regresión utilizando el intervalo de confianza.

Una vez que se ha realizado el análisis y se ha generado los modelos de regresión lineal múltiple para los 18 puntos de muestreo, es posible predecir la media ponderada de humedad de combustible vivo, es importante recalcar que toda predicción tiene asociado un error y por lo tanto un intervalo; en el apartado 5.1.1., para el análisis de las regresiones en este trabajo se ha elegido el criterio del valor-P, considerando su intervalo de confianza del 95%, este devolverá un intervalo para el valor promedio de todas las medias promedio de humedad de combustible vivo.

La obtención de la precisión de las predicciones se solicita en intervalos en torno a su predicción, este proceso se realiza en el lenguaje de programación para análisis estadísticos R, el cual entrega una matriz con la predicción del modelo y un intervalo de confianza del 95% alrededor de la predicción media, el análisis se realiza para todos los puntos de muestreo a partir de los modelos de regresión obtenidos entre la humedad de combustible vivo y las variables de índices de vegetación más meteorológicas que se muestran en la Tabla 23.

Como ejemplo para la estimación se toma el punto de muestreo Fredes, del cual se sabe que este modelo explica un 94.077% de la varianza de la media ponderada de HCV, en la Tabla 23 se muestra el valor predicho por el modelo (fit) de HCV, sus respectivos intervalos de confianza (lwr y upr) siendo el intervalo bajo y alto respectivamente, se verifica que el valor real de la media ponderada de humedad de combustible vivo este dentro de dicho intervalo de confianza para comprobar la precisión del modelo. Todos los valores de HCV se encuentran en el rango del intervalo de confianza a excepción de dos valores pertenecientes a las fechas de muestreo 13/2/2016 y 23/7/2017 teniendo una

diferencia mínima positiva entre el valor real de HCV y el rango “upr” del intervalo de confianza, es decir que el valor de HCV es mayor que el intervalo de confianza.

Tabla 23. Estimación de la HCV en el punto de muestreo Fredes.

Fecha de muestreo	fit	lwr	upr	Diferencia HCV_P y fit
18/1/2016	96.36129	91.20317	101.51941	
13/2/2016	92.89962	88.62025	97.17899	0.38
30/3/2016	101.96919	95.79423	108.14416	
17/4/2016	106.09347	98.84648	113.34047	
15/5/2016	116.39807	111.63687	121.15928	
12/6/2016	129.48817	122.84846	136.12788	
14/7/2016	105.51814	99.97209	111.06418	
16/8/2016	93.22456	87.67063	98.77848	
19/9/2016	77.70624	70.94288	84.4696	
17/10/2016	103.45096	96.78779	110.11414	
16/11/2016	93.50962	85.95749	101.06174	
23/12/2016	103.222	97.19095	109.25304	
20/3/2017	91.09795	84.03256	98.16334	
23/4/2017	110.82974	103.95862	117.70087	
22/5/2017	116.53778	109.7782	123.29737	
18/6/2017	112.01451	104.83457	119.19445	
23/7/2017	106.90478	101.23932	112.57024	0.75
20/8/2017	88.63198	80.16652	97.09744	
17/9/2017	92.96677	85.24877	100.68477	
24/10/2017	91.38609	83.0909	99.68129	
21/11/2017	87.81704	81.52272	94.11137	
20/12/2017	84.2705	78.0457	90.4953	
16/1/2018	88.42794	83.07019	93.78568	
20/2/2018	94.59904	89.31066	99.88741	
13/3/2018	92.96895	87.15171	98.7862	
22/4/2018	81.10235	74.73133	87.47337	
15/5/2018	111.64267	106.11869	117.16664	
17/6/2018	155.90236	148.23288	163.57183	
3/7/2018	110.98399	102.63498	119.33299	

En la Imagen 48 se puede observar que los datos de la media ponderada de humedad de combustible vivo del punto de muestreo Fredes entran en el rango del intervalo de confianza a excepción de los dos datos antes mencionados.

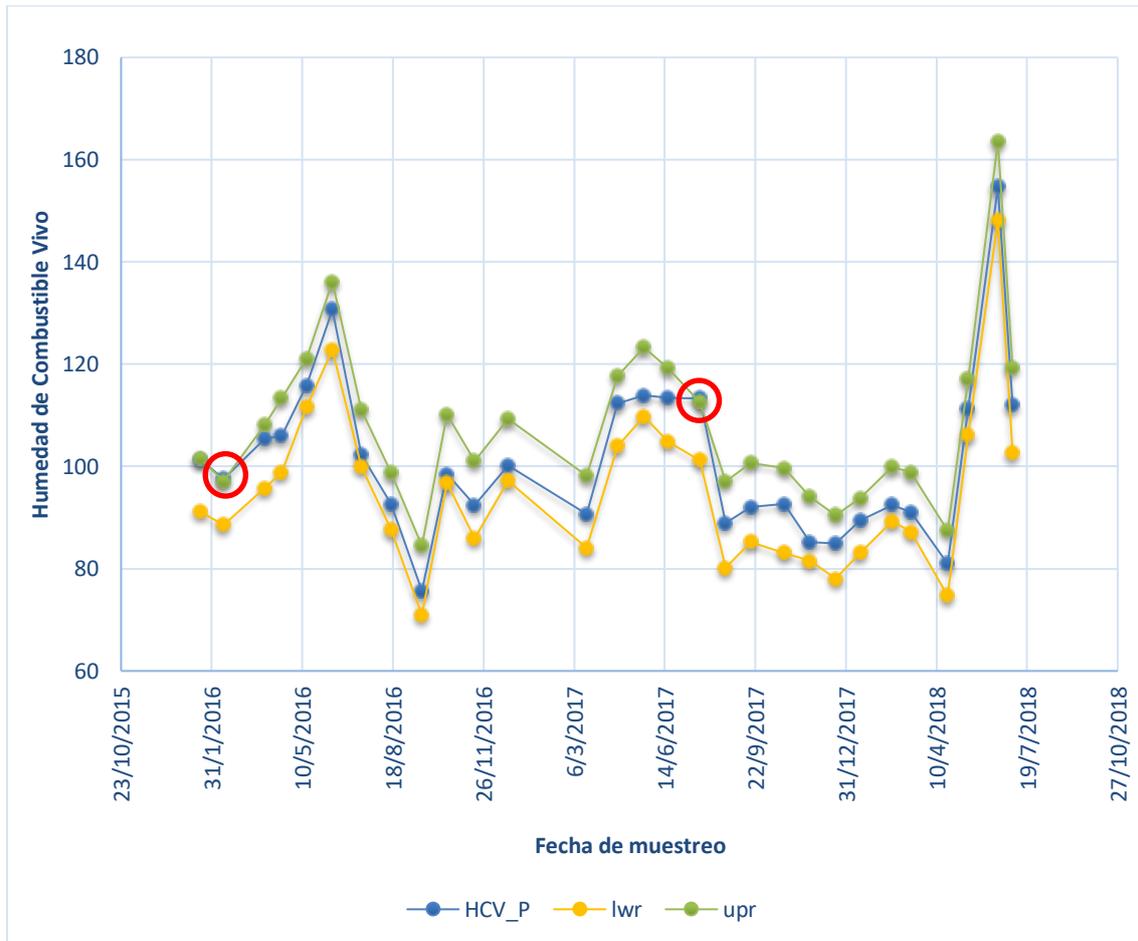


Imagen 48. Representación de la HCV de Fredes en el rango de intervalo de confianza.

La estimación de la media ponderada de HCV se analiza en todos los puntos de muestreo, de los cuales al tener modelos buenos con un R-cuadrada ajustada de más del 90% los valores de HCV ingresan en el intervalo de confianza calculado para cada modelo, en la Tabla 24 se observa las fechas de toma de muestra de HCV que no ingresan en el intervalo de confianza calculado para cada modelo, en los puntos de muestreo Fredes, Puebla Tornesa, Montanejos y Castielfabid son pocos los datos que no ingresan en el intervalo de confianza, teniendo una diferencia mínima entre el rango y el valor real, la mayoría de estos valores pertenecen a verano siendo los datos más variantes en la temporalidad de HCV analizado en los tres años. El punto de muestreo Villargordo del Cabriel al tener un ajuste bajo con un R cuadrado de 56.96% es el modelo que más fechas deja fuera del rango del intervalo de confianza, siendo 15 fechas de las 35 observaciones realizadas.

Tabla 24. Fechas de toma de muestreo que no ingresan en el intervalo de confianza.

Nº	Punto de muestreo	R-Cuadrada Ajustada	Fecha de muestreo	fit	lwr	upr	Diferencia HCV_P y rango
1	Fredes	94.0766	13/2/2016	92.89962	88.62025	97.17899	0.38
			23/7/2017	106.90478	101.23932	112.57024	0.75
2	Puebla Tornesa	84.207	30/8/2017	76.08841	69.12198	83.05484	0.07
3	Montanejos	96.516	5/7/2016	93.67014	91.77927	95.56101	-1.35
			20/6/2017	103.58723	101.25746	105.917	0.67
4	Castielfabid	94.627	12/9/2017	90.81323	85.88224	95.74421	-1.55
5	Villargordo del Cabriel	56.960	15/1/2016	85.1758	80.72039	89.63121	0.72
			30/3/2016	79.82401	77.5054	82.14263	-4.35
			16/6/2016	83.92228	79.66844	88.17611	-3.77
			7/7/2016	74.56361	70.96406	78.16317	-7.56
			4/8/2016	71.97304	68.11713	75.82895	-8.67
			24/8/2016	74.56084	71.31287	77.80881	2.95
			23/3/2017	86.27917	83.82326	88.73508	4.51
			25/4/2017	89.05153	86.08106	92.02199	2.50
			22/5/2017	88.26526	84.45281	92.07771	1.81
			22/6/2017	79.83296	76.27255	83.39337	1.08
			20/7/2017	72.47759	68.91832	76.03686	0.06
			21/8/2017	74.89273	71.45588	78.32959	0.01
			4/9/2017	79.76414	76.98064	82.54764	1.28
			15/10/2017	82.98606	79.58247	86.38964	-0.96
27/3/2018	88.62557	84.91463	92.33651	-9.94			

7. CONCLUSIONES

- Las especies forestales muestreadas son típicas de los montes mediterráneos, destacando de ellas *Pinus halepensis* y *Rosmarinus officinalis* siendo las más representativas a nivel general. De las especies muestreadas se obtuvo un HCV considerable por lo que se ha optado en realizar el análisis estadístico de todas ellas determinando la media ponderada de HCV.
- El muestreo temporal de los datos de humedad de combustible vivo está muy limitado, ya que el trabajo de campo para las mediciones de HCV requiere mucho esfuerzo, se mide de manera manual en un número limitado de sitios. Para superar las limitaciones espaciales y temporales de este tipo de mediciones se han utilizado índices de vegetación derivados de datos satelitales del sensor Sentinel-2 y variables meteorológicas pertenecientes a estaciones cercanas a los puntos de muestreo.
- La generación de sensores Sentinel-2 proporcionó información espacial con una resolución alta (10 m de resolución), ofreciendo la oportunidad de mejorar las estimaciones de HCV por medio del cálculo de índices de vegetación recuperados de este sensor, obteniendo información sobre el estado de la vegetación y las características de las mismas.
- Las fechas de toma de imágenes satelitales descargadas son las más cercanas a la fecha de toma de muestreo de HCV en las parcelas, por tal razón para que no exista este desfase se realizó una interpolación lineal de datos de índices de vegetación a la fecha de muestreo, obteniendo así las variables en un mismo periodo, se pudo comprobar que los datos no variaban en grandes cantidades ya que los índices cambian de una manera representativa en un periodo amplio mas no en el transcurso de pocos días.
- Los índices de vegetación calculados (NDVI, EVI, SAVI, NDWI, VARI y VIgreen) están relacionados con el verdor de la superficie y la biomasa de la vegetación representada por el índice de área de la hoja verde que se ven afectadas y por lo tanto se pudo correlacionar con la HCV. Se observó que estas variables son interdependientes con tendencias interanuales similares, lo que insinuó una posible estimación de HCV a partir de los índices de vegetación. El análisis de la regresión múltiple entrega el rendimiento de la información espectral derivada del sensor obteniendo modelos con un R-cuadrada ajustada hasta de 65% siendo los mejores modelos los pertenecientes a los puntos de muestreo Gandia, Gilet, Castielfabid y Muela Cortes. Cabe recalcar que al evaluar la bondad del ajuste los resultados son muy bajos por tanto no se consideran buenos modelos para la estimación de HCV.
- La humedad de combustible vivo también depende físicamente de la humedad del suelo, la evapotranspiración y la fisiología de las plantas. Por tal motivo interviene la humedad, temperatura y precipitaciones, siendo variables meteorológicas que varían durante el desarrollo de la planta. Fue necesario añadir al estudio dichas variables y analizar la relación en conjunto con los índices de vegetación,

obteniendo mejores modelos para la estimación, con R-cuadrada ajustada mayores al 90% en todos los puntos de muestreo a excepción de Villargordo y Puebla Tornesa con una bondad de ajuste de 56 y 84% respectivamente. Las variables predictoras de índices de vegetación y meteorológicas al relacionarse entre sí hacen que contribuyan de una manera significativa a los modelos.

- El procedimiento de ajuste paso a paso hacia atrás empleado en las regresiones múltiples permitió que el modelo involucre todas las variables calculadas y vaya eliminando del modelo las variables que no sean estadísticamente significativas según el criterio del Valor-P inferior a 0.05, obteniendo los mejores modelos para cada punto de muestreo en el análisis de relación entre HCV, los índices de vegetación y las variables meteorológicas.
- El análisis de la comparación de modelos de regresión con diferentes números de variables, el cumplimiento de las hipótesis, la multicolinealidad, componentes principales y validación cruzada verifican y confirman los resultados obtenidos en las regresiones múltiples realizadas. Dichos análisis proporcionan las condiciones para obtener resultados confiables.
- Se utilizó validación cruzada para evaluar los resultados de los análisis estadísticos en los 18 puntos de muestreo y garantizar que son independientes de la partición entre datos de entrenamiento y test. Se predijo y estimó la precisión del modelo, obteniendo residuales bajos entre el valor de HCV, el predicho por el modelo y el predicho por la validación cruzada.
- Después de obtener los resultados del análisis de los modelos de regresión múltiple para la relación entre HCV y las variables de índices de vegetación, se optó por probar un método de suavizado de los datos denominado filtro Savitsky-Golay, el cual es empleado para disminuir el ruido típico de series de tiempo con el fin de mejorar los modelos. Esta técnica conservó las características de la distribución inicial de los dos grupos de variables obteniendo nuevos modelos. Efectivamente el análisis de los modelos mejora significativamente obteniendo R-cuadradas ajustadas hasta de 88%, sin embargo, no se lo considerará como resultados fiables ya que no se los obtiene mediante valores reales de humedad de combustible vivo.
- Las variables de índices de vegetación más influyentes en los modelos de regresión múltiple son NDVI y EVI siendo los índices que mejor estiman la cantidad, calidad y desarrollo de la vegetación, sin embargo, los otros cuatro índices también contribuyen a las regresiones aportando por lo menos a un punto de muestreo. En cuanto a las variables meteorológicas los datos de precipitación, temperatura, humedad relativa y viento son los predictores que más contribuyen a los modelos; son pocos los puntos de muestreo (Montanejos, Puebla Tornesa, Buñol, Muro de Alcoi, Altea y Biar) en que las variables de levante y punto del rocío aportan significativamente siendo las variables que menos relación tiene con el HCV.

- Los modelos de regresión para la estimación de la media ponderada de HCV en los 18 puntos de muestreo han tenido buenos resultados por lo que se comprobó si los HCV están dentro del intervalo de confianza de cada modelo generado; de los 18 puntos de muestreo, Fredes, Montanejos y Castielfabid son los modelos que a pesar de tener un R-cuadrada ajustada mayor a 90% dos de sus valores de HCV no ingresan en el rango de intervalo de confianza y Puebla Tornesa y Villargordo del Cabriel al tener una bondad de ajuste menor al 90% tiene más datos de HCV que no ingresan al rango. Con este análisis el modelo que no se lo considera bueno pertenece al punto de muestreo Villargordo del Cabriel ya que tiene un R-cuadrada ajustada de 56% y la mitad de los valores de HCV se quedan fuera del rango de confianza del modelo generado.

8. BIBLIOGRAFIA

- Agency, E. S. (24 de 07 de 2015). *SENTINEL-2 User Handbook*. Obtenido de https://sentinel.esa.int/documents/247904/685211/Sentinel-2_User_Handbook
- Argañaraz, J., Landi, M. A., Bravo, S. J., Gavier-Pizarro, G. I., Scavuzzo, C. M., IEEE, & Bellis, L. M. (2016). Estimation of Live Fuel Moisture Content From MODIS Images for Fire Danger Assessment in Southern Gran Chaco. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(12)(DOI: 10.1109/JSTARS.2016.2575366), 5339-5349.
- Boksoon, M., Seung Hee, K., Shenyue, J., Kristen, W., & Menas C, K. (2018). Estimating Live Fuel Moisture from MODIS Satellite Data for Wildfire Danger Assessment in Southern California USA. *Remote Sensing*, 10(1), 87-107.
- Carvacho Bart, L., & Sánchez Martínez, M. (2010). Comparacion de índices de vegetación a partir de imágenes MODIS en la región del Libertador Bernardo O'Higgins, Chile, en el período 2001-2005. *Tecnologías de la Información Geográfica: La Información Geográfica al servicio de los ciudadanos.*, 728-737.
- COPERNICUS, E. (2018). *Copernicus Open Access Hub*. Obtenido de <https://scihub.copernicus.eu/dhus/#/home>
- ESA. (2018). *Sentinel OnLine*. Obtenido de <https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/data-formats>
- European Space Agency. (2018). *Science toolbox exploitation platform*. Obtenido de <http://step.esa.int/main/third-party-plugins-2/sen2cor/>
- Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, 7(11), 1417-1434.
- Huete, A. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25(3), 295-309.

- Huete, A. R. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 59 (3), 440-451.
- Huete, A. R., Hua, G., Qi, J., Chehbouni, A., & Van Leeuwen, W. J. (1992). Normalization of Multidirectional Red and NIR reflectances with the SAVI. *Remote Sensing of Environment*, 41(2-3), 143-154.
- Jönsson, P., & Eklundh, L. (2004). A program for analyzing time-series of satellite sensor data. *Computers & Geosciences*, 30(8), 833-845.
- Kaplan, G., & Avdan, U. (2017). Object-based water body extraction model using. *European Journal of Remote Sensing*, 50(1), 137-143.
- Kaufman, Y., & Tanre, D. (1992). Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, 30(2), 261 - 270.
- Marino, E., Guillén, C., Nur, A., Tomé, J., & Hernando, C. (19 de 11 de 2018). Estimation of live fuel moisture content of shrubland using MODIS and Sentinel-2 images. *Advances in Forest Fire Research*, 2018-226.
- McFeeters, S. K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), 1425–1432.
- PATFOR. (2011). *Conselleria de Agricultura Medio Ambiente, Cambio Climático y Desarrollo Rural*. Obtenido de Plan de Acción Territorial Forestal de la Comunitat Valenciana : <http://www.agroambient.gva.es/es/web/medio-natural/patfor>
- Peterson, S. H., Roberts, D. A., & Dennisin, P. E. (2008). Mapping live fuel moisture with MODIS data: A multiple regression approach. *Remote Sensing of Environment* 112, 4272–4284.
- Rouse Jr., J. (1973). *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*. Maryland: Goddard Space Flight Center.

SENTINEL. (2018). *ESA SENTINEL ONLINE*. Obtenido de <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial>

Soriano Sancho, J. (21 de 06 de 2016). Análisis de la Humedad del Combustible Vivo (HCV) en la Comunitat Valenciana y la influencia de las variables meteorológicas. Valencia.

Statgraphics. (18 de abril de 2018). *Statgraphics.Net*. Recuperado el 1 de mayo de 2018, de <https://www.statgraphics.net/>

Takeshi, M., Kenlo, N. N., Hiroyuki, O., & Satoshi, T. (2010). Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology. *Remote Sensing of Environment*, 2(10), 2369-2387.

Tello, J., Gómez-Báguena, R., & Casterad, M. (2017). Comparación y ajuste en zonas agrícolas de índices de vegetación derivados de. *Nuevas plataformas y sensores de teledetección. XVII Congreso de la Asociación Española de Teledetección.*, 81-84.

Tucker, C., Falkowski, M., Gessler, P., Morgan, P., & Hudak, A. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2), 127–150.

9. PRESUPUESTO

La Tabla 25, muestra el presupuesto aproximado del presente proyecto, se toma en cuenta que los datos de HCV fueron provistos por la empresa VAERSA que trabaja para la Dirección de Prevención de Incendios Forestales de la Generalitat Valenciana obteniéndose de manera gratuita por el convenio con la Generalitat Valenciana, los datos meteorológicos también fueron provistos por VAERSA a pesar de ello en la Tabla 25 refleja el coste de los datos, las imágenes satelitales fueron descargadas gratuitamente en la página de Copernicus <https://scihub.copernicus.eu/>.

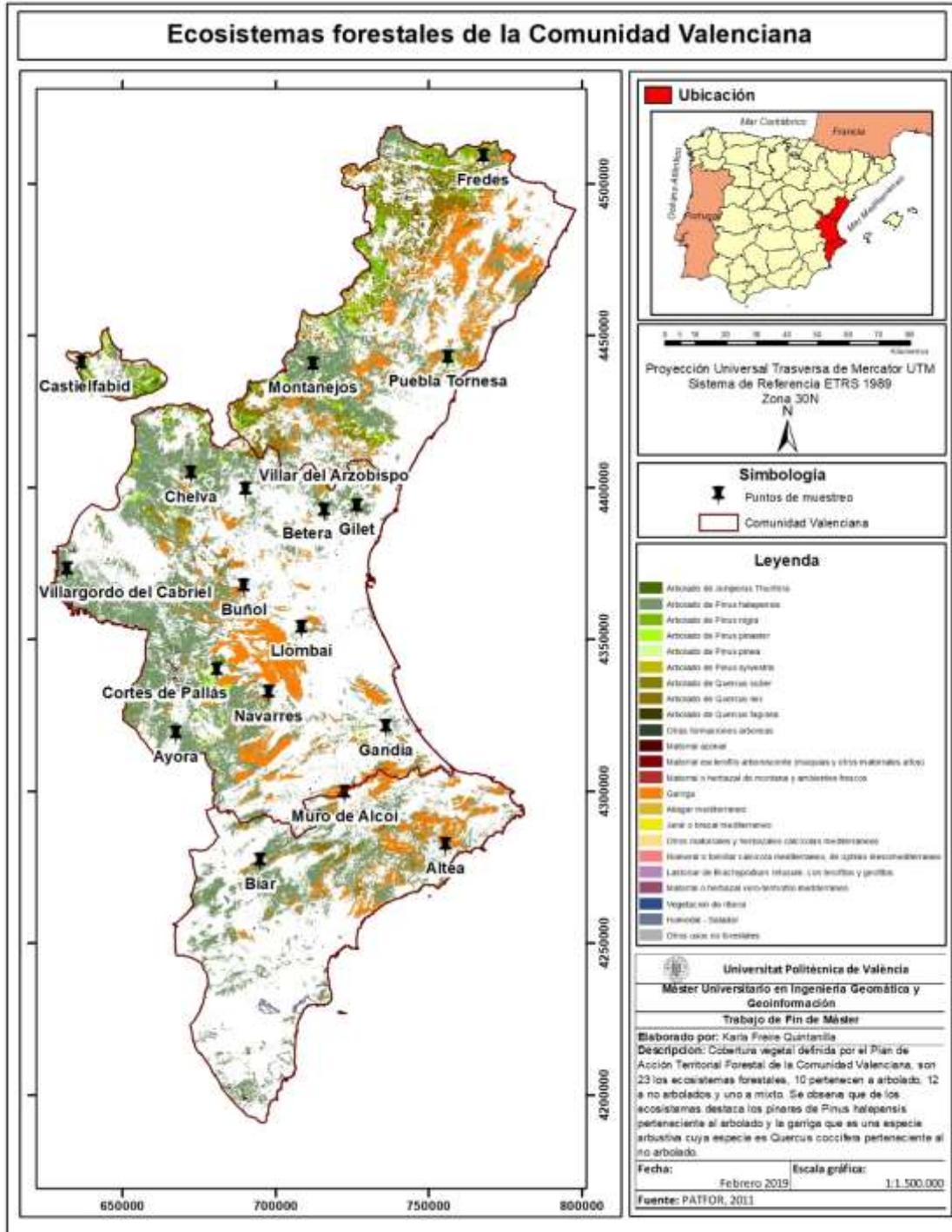
Para el desarrollo del estudio se utilizó algunos softwares como se detalla en la Tabla 25 la mayoría de ellos son libres por lo que no represento un gasto, el único software de paga fue StatGraphics XVII-X64 que a pesar de contar con la licencia de estudiante de la UPV se plasma el precio de una licencia estudiantil para un año.

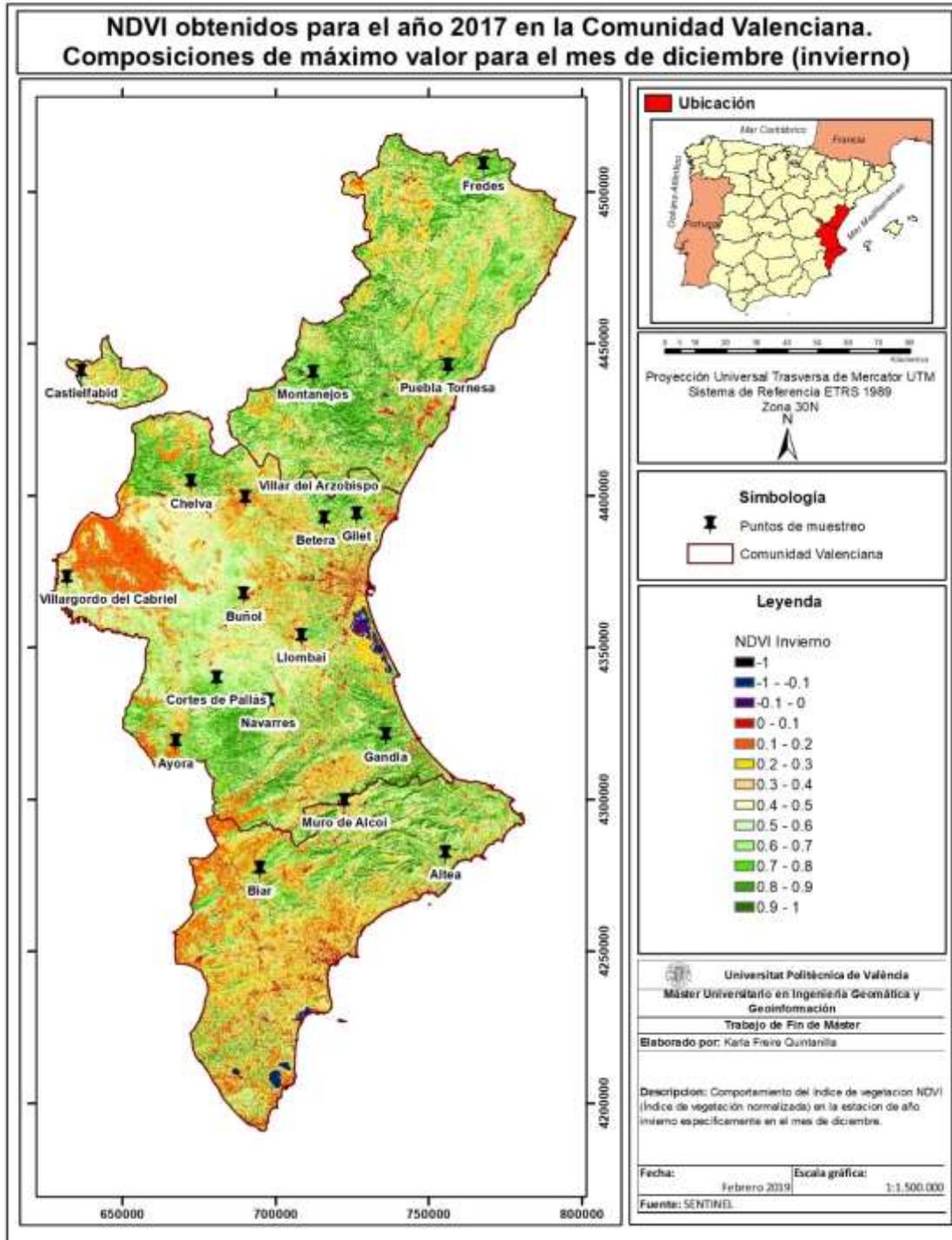
Los costos de contratación de personal técnico de nivel medio son aproximadamente de 1200 euros por mes considerando que se trabaja 25 horas semanales durante 20 semanas; en este caso se han invertido varias horas más, pero para una aproximación se considera lo que se visualiza a continuación.

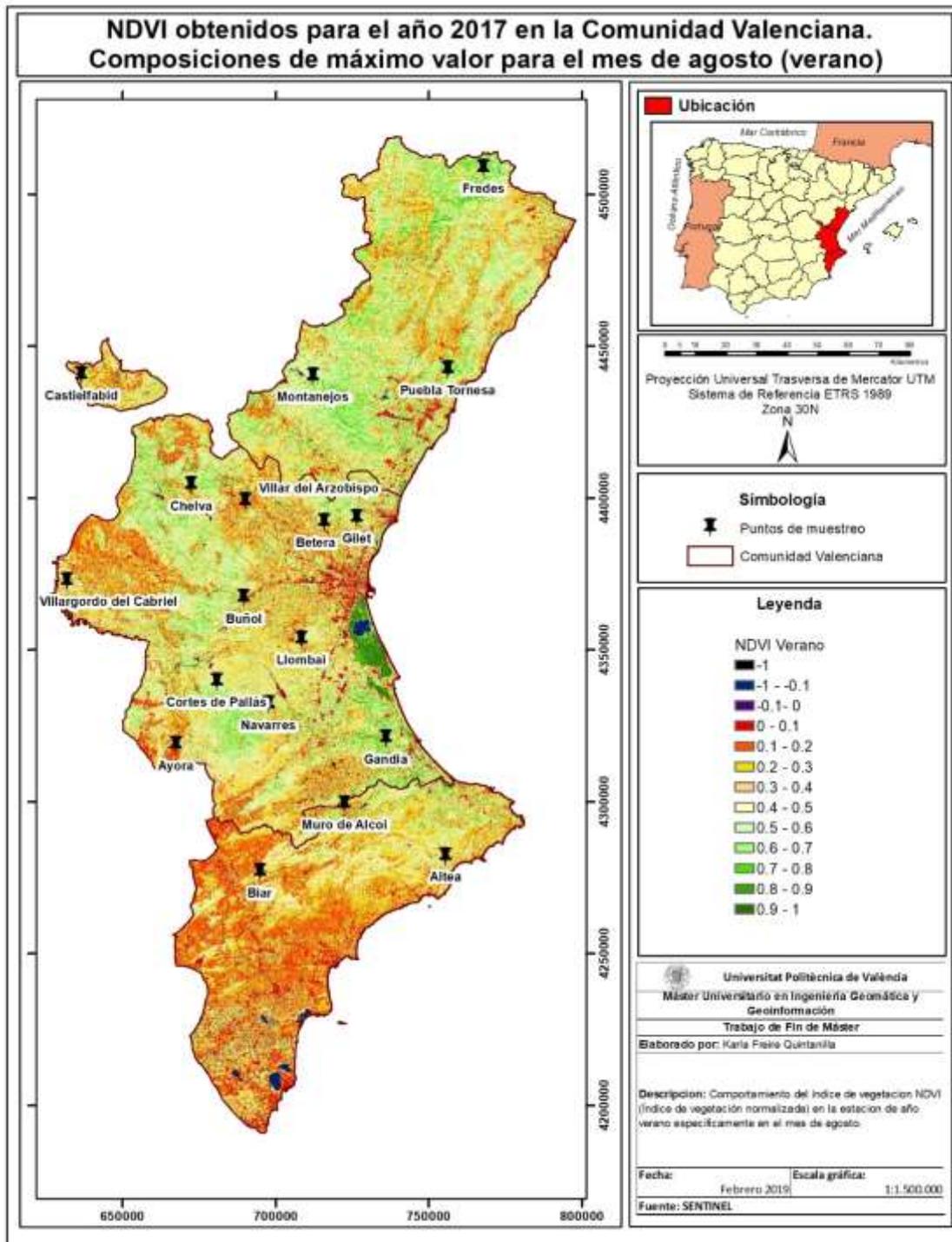
Tabla 25. Presupuesto TFM.

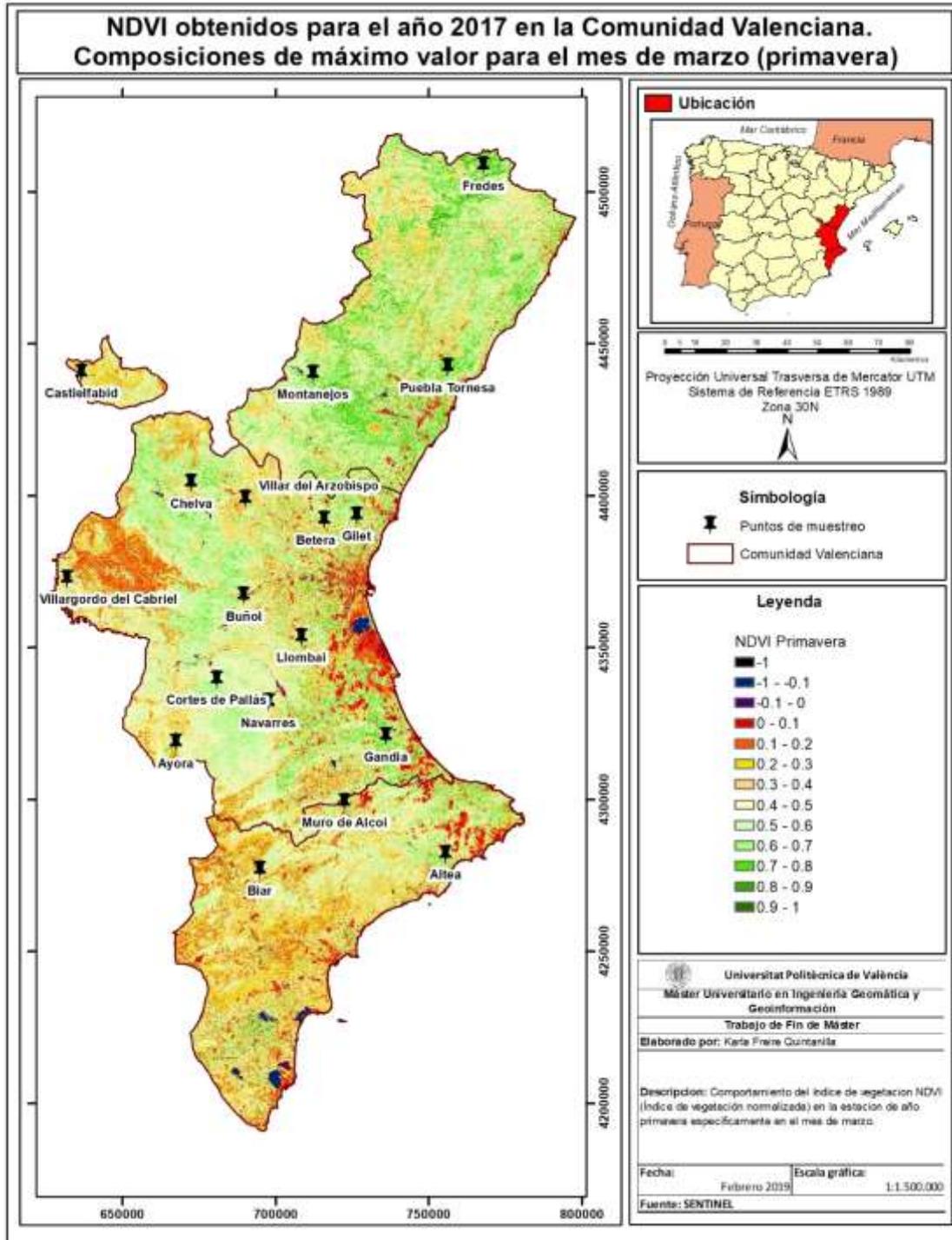
	Mes Insumos	Descripción		Costo (Observatorio/ Variables/Mes)	Septiembre	Octubre	Noviembre	Diciembre	Enero	Total €
Personal Técnico	Personal Técnico Nivel Medio	Contratación			1200	1200	1200	1200	1200	6000
	Seguridad Social y Gastos Adicionales				703,44	703,44	703,44	703,44	703,44	3517,2
Datos	Datos Meteorológicos *	Número de estaciones	18	2,81						12139,2
		Numero de variables (Precipitación, Temperatura, Humedad Relativa y Viento)	4							
		Tiempo (meses)	60							
	Datos de HCV	Provisos por VAERSA de forma gratuita			0		0	0	0	0
	Imágenes Satelitales	Provisas por Copernicus de manera gratuita			0		0	0	0	0
Software	Software Estadístico R	Software Libre			0	0	0	0	0	0
	Python	Software Libre			0	0	0	0	0	0
	QGIS	Software Libre			0	0	0	0	0	0
	StatGraphics XVII - X64**	Licencia de Estudiante (1año)		105,66						105,66
									Total	21762,06
* https://www.boe.es/buscar/doc.php?id=BOE-A-2016-111 ** http://statvision.com/pricing.htm#Academic_Licenses0										

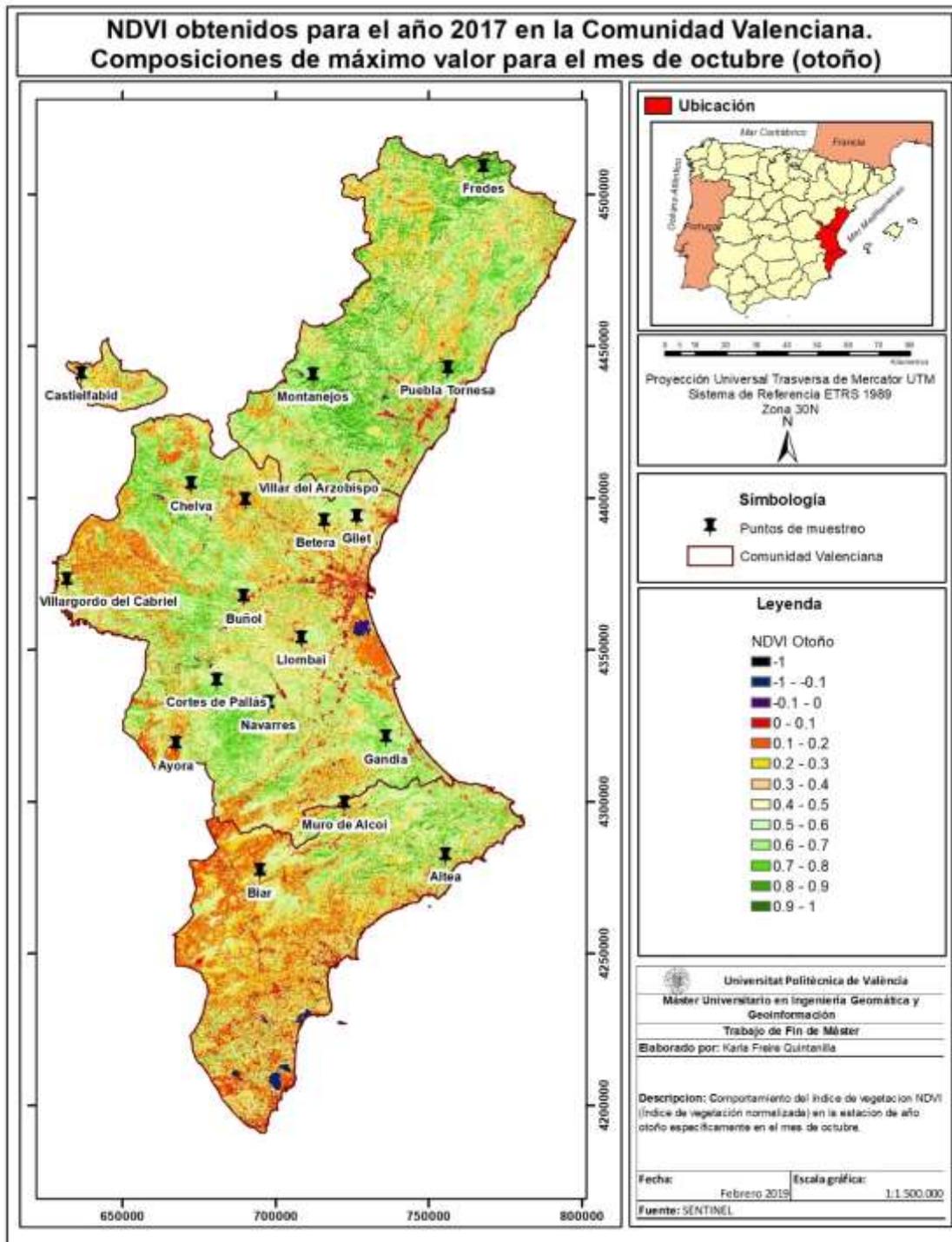
10. CARTOGRAFÍA

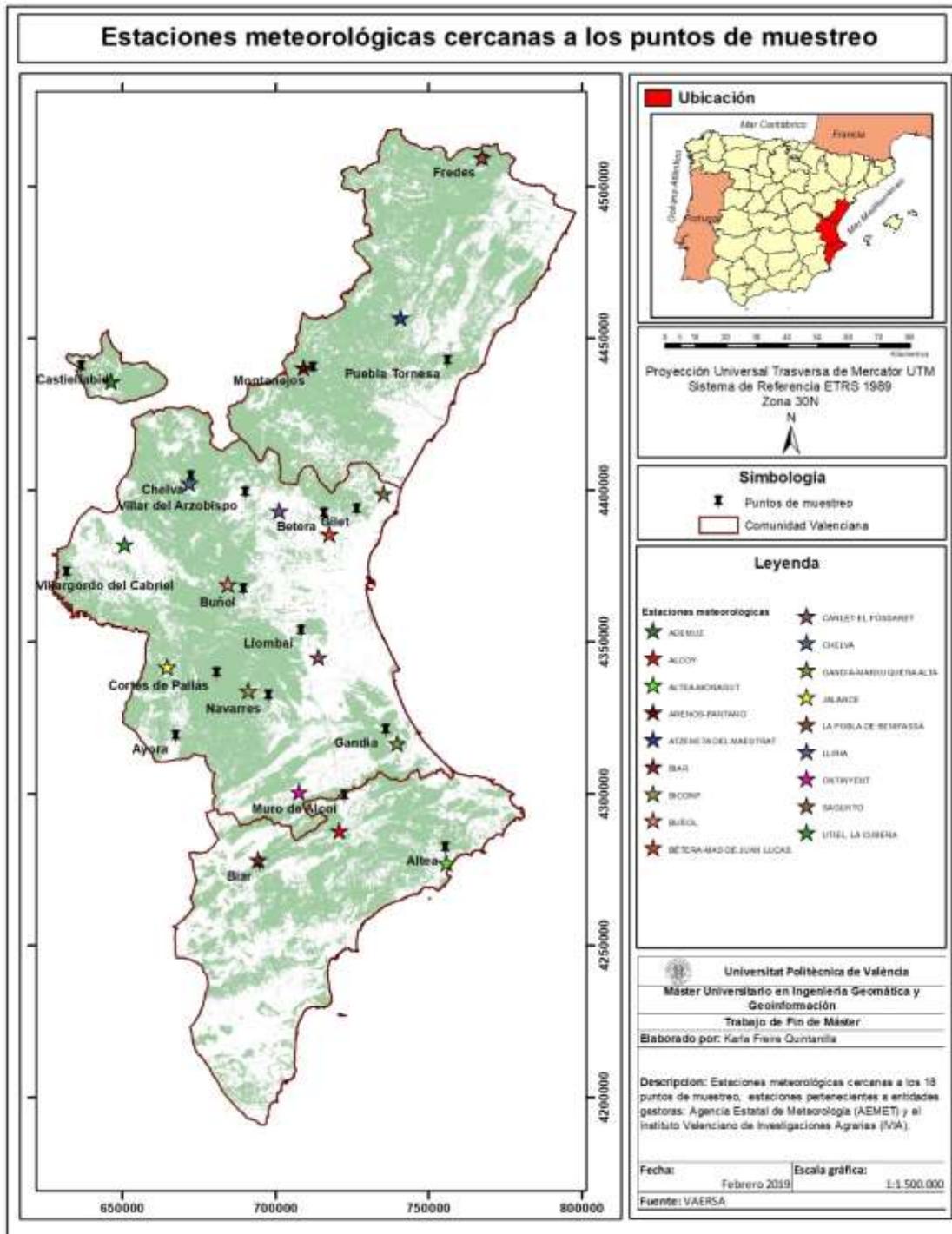












11. ANEXOS

ANEXO 1

```
#Importacion de librerias
import arcpy
from arcpy.sa import *

#Acceso a las variable de entorno
entorno = arcpy.env

#Permitir la escritura sobre un archivo de salida existente
entorno.overwriteOutput = True

#Asignacion del directorio de trabajo (workspace)
entorno.workspace = r'E:\RESPALDO\VALENCIA\MAESTRIA INGENIERIA
GEOMATICA Y GEOINFORMACION\TFM\Imagenes\Procesos\TBF_2016'

ras = arcpy.ListRasters("*", "jp2")

for i in range(0, len(ras), 4):

    b2= ras[i]
    b3= ras[i+1]
    b4= ras[i+2]
    b8 = ras[i+3]
    nom = b4[11:19]#25-33    #11-19

    ndvi_nom = str(nom) +'ndvi.tif'
    evi_nom = str(nom) +'evi.tif'
    savi_nom = str(nom) +'savi.tif'
    ndwi_nom = str(nom) +'ndwi.tif'
    vig_nom = str(nom) +'vig.tif'
    vari_nom = str(nom) +'vari.tif'

    #Ingreso al area de la parcela que se va a extraer la informacion
del raster
    area = r'E:\RESPALDO\VALENCIA\MAESTRIA INGENIERIA GEOMATICA Y
GEOINFORMACION\TFM\Calculos\_SHP\bf_15_TBF.shp'
    campo = "FID"

    #Tabla de salida donde se guarda los datos
    tabla_salida_ndvi = str(nom) +"mean_ndvi.dbf"
    tabla_salida_evi = str(nom) +"mean_evi.dbf"
    tabla_salida_savi = str(nom) +"mean_savi.dbf"
    tabla_salida_ndwi = str(nom) +"mean_ndwi.dbf"
    tabla_salida_vig = str(nom) +"mean_vig.dbf"
    tabla_salida_vari = str(nom) +"mean_vari.dbf"

    try:
```

```
#Comprobacion de la disponibilidad de la licencia de la
extension de Spatial Analysis
if arcpy.CheckExtension('Spatial') == 'Available':
    #Petición de la licencia
    arcpy.CheckOutExtension('Spatial')

#Raster carga en memoria el archivo raster y Float lo
convierte en numeros de coma flotante
r = Float(Raster(b4)) #banda del rojo en sentinel B4
irc = Float(Raster(b8)) #banda del infrarrojo cercano
sentinel B8
a = Float(Raster(b2)) #banda del azul en sentinel B2
v = Float(Raster(b3)) #banda del verde en sentinel B3

#Expresion para el calculo del NDVI
ndvi = (irc-r)/(irc+r)
ndvi.save(ndvi_nom)

#Expresion para el calculo del EVI
evi = (2.5*(irc-r))/(irc+6*r-7.5*a+1)
evi.save(evi_nom)

#Expresion para el calculo del SAVI
savi = (1.5*(irc-r))/(irc-r+0.5)
savi.save(savi_nom)

#Expresion para el calculo del NDWI
ndwi = (v-irc)/(v+irc)
ndwi.save(ndwi_nom)

#Expresion para el calculo del VGreen
vigreen = (v-r)/(v+r)
vigreen.save(vig_nom)

#Expresion para el calculo del VARI
vari = (v-r)/(v+r-a)
vari.save(vari_nom)

#Se calcula las estadísticas del raster obtenido
arcpy.CalculateStatistics_management (ndvi)
arcpy.CalculateStatistics_management (evi)
arcpy.CalculateStatistics_management (savi)
arcpy.CalculateStatistics_management (ndwi)
arcpy.CalculateStatistics_management (vigreen)
arcpy.CalculateStatistics_management (vari)

#Se calcula los estadísticos en la tabla de salida para el
area según el índice de vegetación
ZonalStatisticsAsTable (area, campo, ndvi,
tabla_salida_ndvi, "DATA", "MEAN")
ZonalStatisticsAsTable (area, campo, evi,
tabla_salida_evi, "DATA", "MEAN")
```

```
ZonalStatisticsAsTable (area, campo, savi,
tabla_salida_savi, "DATA", "MEAN")
ZonalStatisticsAsTable (area, campo, ndwi,
tabla_salida_ndwi, "DATA", "MEAN")
ZonalStatisticsAsTable (area, campo, vigreen,
tabla_salida_vig, "DATA", "MEAN")
ZonalStatisticsAsTable (area, campo, vari,
tabla_salida_vari, "DATA", "MEAN")

#Añade filas a cada una de las estadísticas
arcpy.AddField_management(tabla_salida_ndvi, "MEAN_NDVI",
"FLOAT")
arcpy.AddField_management(tabla_salida_evi, "MEAN_EVI",
"FLOAT")
arcpy.AddField_management(tabla_salida_savi, "MEAN_SAVI",
"FLOAT")
arcpy.AddField_management(tabla_salida_ndwi, "MEAN_NDWI",
"FLOAT")
arcpy.AddField_management(tabla_salida_vig, "MEAN_VIG",
"FLOAT")
arcpy.AddField_management(tabla_salida_vari, "MEAN_VARI",
"FLOAT")

#Asigna los estadísticos a las filas
arcpy.CalculateField_management (tabla_salida_ndvi,
"MEAN_NDVI", "[MEAN]")
arcpy.CalculateField_management (tabla_salida_evi,
"MEAN_EVI", "[MEAN]")
arcpy.CalculateField_management (tabla_salida_savi,
"MEAN_SAVI", "[MEAN]")
arcpy.CalculateField_management (tabla_salida_ndwi,
"MEAN_NDWI", "[MEAN]")
arcpy.CalculateField_management (tabla_salida_vig,
"MEAN_VIG", "[MEAN]")
arcpy.CalculateField_management (tabla_salida_vari,
"MEAN_VARI", "[MEAN]")

#Devolucion de la licencia
arcpy.CheckInExtension('Spatial')
else:
    print ('Licencia no disponible')

except:
    arcpy.AddMessage(arcpy.GetMessages())
print 'terminado'
```

ANEXO 2

```
library(readxl)
library(GGally)
library(corpcor)
library(mctest)
library(ppcor)

setwd("E:/RESPALDO/VALENCIA/MAESTRIA INGENIERIA GEOMATICA Y
GEOINFORMACION/TFM/Calculos/DATOS/IV_Ponderado")
input_file<-read_xlsx("CASTIELFABID_16_18.xlsx")

#Modelo calculado
equation<- (HCV_P~NDVI+EVI+NDWI+VARI+ NDWI2)

#Calculo regresion
reg <- lm(as.formula(equation),data=input_file)
summary(reg)

#Plot de la regresion
par(mfrow=c(2,2))
plot(reg)

#Plot correlaciones
X<-input_file[,c("NDVI","EVI","NDWI","VARI","NDWI2")]
ggpairs(X)

#Calculo de la covarianza
cor2pcor(cov(X))
y<-input_file["HCV_P"]
omcdiag(X,y)
imcdiag(X,y)
pcor(X, method = "pearson")
```

ANEXO 3

```
library(faraway)
library(readxl)
library(pls)

setwd("E:/RESPALDO/VALENCIA/MAESTRIA INGENIERIA GEOMATICA Y
GEOINFORMACION/TFM/Calculos/DATOS/IV_PONDERADO")
input_file<-read_xlsx("CASTIELFABID_16_18.xlsx")

training <- input_file[c(1:12,14:15,17:18,20:22,24:30),]
test <- input_file[c(13,16,19,23),]

equation<-(HCV_P~NDVI+EVI+NDWI+VARI+I(NDWI^2))

#Calculo regresion

modelo <- lm(as.formula(equation),data=training)
summary(modelo)

# MSE empleando las observaciones de entrenamiento

training_mse <- mean((modelo$fitted.values - training$HCV_P)^2)
training_mse

# MSE empleando nuevas observaciones

predicciones <- predict(modelo, newdata = test)
test_mse <- mean((predicciones - test$HCV_P)^2)
test_mse

modelo_step_selection <- step(object = modelo, trace = FALSE)

# Número de predictores del modelo resultante

length(modelo_step_selection$coefficients)

# Training-MSE

training_mse <- mean((modelo_step_selection$fitted.values -
training$HCV_P)^2)
training_mse

# Test-MSE

predicciones_st <- predict(modelo_step_selection, newdata = test)
test_mse <- mean((predicciones_st - test$HCV_P)^2)
test_mse

# Cálculo de componentes principales. Se excluye la columna con la
variable respuesta *HCV*

pca <- prcomp(training[,c("NDVI", "EVI", "NDWI", "VARI", "NDWI2")], scale.
= TRUE)
```

```
# Se muestra la proporción de varianza explicada y acumulada de las 14  
primeras componentes
```

```
summary(pca)$importance[, 1:5]
```

```
modelo_pcr <- pcr(formula = equation, data = training, scale. = TRUE,  
ncomp = 4)
```

```
training_mse <- mean((modelo_pcr$fitted.values - training$HCV_P)^2)  
training_mse
```

```
# Test-MSE
```

```
predicciones_pcr <- predict(modelo_pcr, newdata = test, ncomp = 4)  
test_mse <- mean((predicciones_pcr - test$HCV_P)^2)  
test_mse
```

```
set.seed(123)
```

```
modelo_pcr <- pcr(formula = equation, data = training, scale. = TRUE,  
validation = "CV")  
modelo_pcr_CV <- MSEP(modelo_pcr, estimate = "CV")  
which.min(modelo_pcr_CV$val)
```

```
par(mfrow = c(1, 2))
```

```
plot(modelo_pcr_CV$val, main = "MSE vs n° componentes", type = "l",  
ylab = "MSE",  
col = "blue", xlab = "Componentes")
```

```
plot(modelo_pcr_CV$val, main = "zoom", type = "l", ylab = "MSE",  
xlab = "Componentes", col = "blue", ylim = c(0, 20))
```

```
# Test-MSE
```

```
predicciones_pcr_CV <- predict(modelo_pcr, newdata = test, ncomp = 4)  
test_mse <- mean((predicciones_pcr_CV - test$HCV_P)^2)  
test_mse
```

ANEXO 4

```
library(readxl)
library(signal)
library(ggplot2)

setwd("E:/RESPALDO/VALENCIA/MAESTRIA INGENIERIA GEOMATICA Y
GEOINFORMACION/TFM/Calculos/DATOS/IV_PONDERADO")
input_file<-read_xlsx("MUELA_CORTEES_17-18.xlsx")
attach(input_file)

x1<-Vigreen
x2<- (filter(filt = sgolay(p = 1, n = 3), x = x1))
y<-FECHA_FIN

x1
x2

plot(data.frame( y,x1), type='l', lwd=2, lty=1,ylab='')
lines(y,x2, lwd=2,lty=1, col="red",ylab='')

legend('top', legend=c('original', 'Savitzky-Golay'), col=1:2, lty=1,
lwd=2)
title(xlab="FECHA", col.lab="black")
title(ylab="HCV_P", col.lab="black")
```