

Autores:

Génesis Buitrón,
Doris Cevallos,
Richard Chávez,
Diego Cunalata.

Guía metodológica para la estimación de superficies sembradas



Guía metodológica para la estimación de superficies sembradas

Elaborado por: Génesis Buitrón, Doris Cevallos, Richard Chávez y Diego Cunalata.

Diseño y Diagramación:

Fundación Datalat

Proyecto +Datos +Transparencia

Fundación Datalat

En el marco del **Convenio de Cooperación Interinstitucional entre el Ministerio de Telecomunicaciones y de la Sociedad de la Información y la Fundación Datalat No. 024-2022**, y los compromisos del **Primer Plan de Acción de Gobierno Abierto Ecuador 2019-2021** y el **Segundo Plan de Acción de Gobierno Abierto Ecuador 2022-2024** apoyados por la **Secretaría de Gobierno Abierto**.

Apoyado por:

National Endowment for Democracy

Ecuador, enero del 2024

La presente guía puede ser reproducida y difundida de forma libre siempre y cuando reconozca la autoría y se asegure un uso ético de la información.

Esta guía ha sido producida con el financiamiento de National Endowment for Democracy para equipos seleccionados como ganadores del evento participativo “Datajam 2da. edición: Mide, analiza, actúa” organizado por Fundación Datalat, el Ministerio del Ambiente, Agua y Transición Ecológica (MAATE), el Ministerio de Agricultura y Ganadería (MAG), el Ministerio de Salud Pública (MSP), el Ministerio de Telecomunicaciones y de la Sociedad de la Información (MINTEL), Subsecretaría de Gobierno Abierto de la Presidencia de la República y la Dirección de Innovación y Vinculación de la EPN, en el marco del proyecto “+Datos +Transparencia”, su contenido es responsabilidad exclusiva de sus autores, no necesariamente es un reflejo de los puntos de vista de Fundación Datalat, MAATE, MAG, MSP, MINTEL, Subsecretaría de Gobierno Abierto de la Presidencia de la República, o la Dirección de Innovación y Vinculación de la EPN.

Contenidos

Contenidos	4
Introducción	5
1. Objetivos	6
1.1. Objetivos General	6
1.2. Objetivos Específicos	6
2. Herramientas de trabajo	6
2.1. ArcGis	6
2.2. Minitab	7
2.3. MaxEnt	7
2.4. Excel	8
3. Insumos	8
Datos libres de superficies cultivada a través del Ministerio de Agricultura y Ganadería	8
4. Procedimientos	9
4.1 Estimación de superficies cultivadas	9
4.1.1 Depuración de datos	10
4.1.2 Sistematización de datos	14
4.1.3 Creación del modelo de Regresión Lineal Simple (RLS)	16
4.2 Estimación de las superficies idóneas de mediante MaxEnt Modelos de distribución de especies (SDM)	20
4.2.1 Obtención de datos	21
4.2.2 Preparación y estandarización de datos	24
4.2.3. Análisis de datos en MaxEnt	28
5. Resultados	30
5.1 Resultados de estimación de superficies cultivadas hasta 2025	30
5.2 Resultados de la estimación de superficies idóneas para los cultivos en el 2050	32
6. Recomendaciones técnicas	36
Referencias bibliográficas	36

Introducción

La guía metodológica propuesta tiene como objetivo principal desarrollar un marco integral para la estimación de superficies sembradas de *Oryza sativa* (arroz) y *Zea mays* (maíz) hasta el año 2025. Con un enfoque en la integración de datos abiertos y sistemas de información geográfica (SIG), la metodología se estructura en torno a objetivos específicos que abarcan desde la documentación de la metodología hasta la proyección futura de áreas sembradas bajo escenarios climáticos específicos. Las herramientas de trabajo seleccionadas incluyen ArcGIS, Minitab y MaxEnt, cada una desempeñando un papel crucial en la recopilación, análisis estadístico y modelado de datos geoespaciales. ArcGIS se destaca por su capacidad para visualizar y analizar información geográfica, mientras que Minitab se emplea para análisis estadísticos detallados, y MaxEnt se utiliza para modelar la distribución de especies o idoneidad del hábitat.

La guía utiliza insumos de datos libres provenientes del Ministerio de Agricultura y Ganadería, representando estimaciones de superficie sembrada para arroz y maíz entre los años 2015 y 2022. Estos datos se presentan en formato poligonal, incluyendo información clave como tipo de cultivo, división territorial y superficie cultivada en hectáreas. Los procedimientos detallados se centran en la estimación de superficies cultivadas, utilizando un modelo de Regresión Lineal Simple (RLS). La metodología explora los conceptos fundamentales de la RLS, incluyendo el coeficiente de determinación R^2 como indicador de la calidad del ajuste del modelo.

1. Objetivos

1.1. Objetivos General

Desarrollar una guía metodológica para la estimación de superficies sembradas de *Oryza sativa* (arroz) y *Zea mays* (maíz) hasta 2025, respectivamente, basada en la integración de datos abiertos y sistemas de información geográfica (SIG).

1.2. Objetivos Específicos

- Documentar una metodología para la estimación de superficies sembradas teniendo como base información cartográfica y datos abiertos de la cobertura y uso de tierra en el periodo 2015 - 2022.
- Predecir las superficies sembradas con un modelo de Regresión Lineal Simple y distribución potencial mediante estadística y cartografía.
- Diseñar mapas de proyección futura hasta el año 2050 de las áreas sembradas bajo el escenario ssp585 de las Trayectorias Socioeconómicas Compartidas (SSP) de cambio climático.

2. Herramientas de trabajo

2.1. ArcGis

Dentro de ArcGIS, puedes buscar y acceder a diferentes herramientas utilizando la función de búsqueda (search). ArcGIS es un sistema que permite recopilar, organizar, administrar, analizar, compartir y distribuir información geográfica. A través de la búsqueda, puedes encontrar y utilizar diversas herramientas para compilar información geográfica, resolver problemas de análisis espacial, dar a conocer y compartir información mediante la geografía y la visualización. (ESRI, 2023). Herramientas aplicadas de ArcGis para esta guía:

- Tabla de contenidos: ventana que muestra una lista de todas las capas, tablas y otros elementos que componen un mapa o proyecto en ArcMap.
- Visualización de mapas: ArcMap permite la visualización de datos geospaciales en forma de mapas. Proporciona una amplia gama de símbolos y estilos para representar diferentes tipos de datos
- Selección por atributos permite seleccionar elementos en una capa o tabla basándote en ciertos criterios o condiciones definidas por los atributos de los datos.
- Summarize" se refiere a una herramienta de análisis que permite resumir datos geospaciales en función de un campo o atributo específico
- Clip herramienta de análisis espacial que permiten recortar o limitar una capa o entidad a un área definido.
- Extract by mask: permite extraer los datos de una capa raster o imagen basada en una máscara definida por una capa de polígonos, recortando los datos raster a una región específica.

2.2. Minitab

Software estadístico utilizado para el análisis de datos y la generación de informes. Es ampliamente utilizado en la industria, la educación y la investigación para realizar análisis estadísticos y explorar patrones y tendencias en los datos. MINITAB permite hacer un uso adecuado de la información, es por ello por lo que se utiliza muy habitualmente en el control de Calidad (Minitab, 2023). Herramientas aplicadas de Minitab para esta guía:

- Gráfico de probabilidad normal en Minitab es una herramienta visual para evaluar la normalidad de los datos.
- La matriz de correlación en Minitab muestra los coeficientes de correlación de Pearson, que varían entre -1 y 1

2.3. MaxEnt

Maximum Entropy Modeling (MaxEnt) es un método estadístico y algoritmo de modelado utilizado para modelar la distribución de especies o la idoneidad del hábitat. Este software está programado en Java y tiene una interfaz gráfica de usuario. El enfoque se basa en el principio de entropía máxima, que busca encontrar el modelo más simple y generalizado que sea consistente con los datos y las restricciones conocidas (Phillips, Duvick, & Schapire, 2017).

2.4. Excel

La depuración y sistematización de datos en Excel es un proceso fundamental para garantizar la calidad y fiabilidad de la información. La depuración implica identificar y corregir errores o inconsistencia en los datos como valores faltantes. La sistematización facilita la interpretación y el análisis de los datos para la toma de decisiones. Herramientas aplicadas de Excel para esta guía:

- Libro de Excel: archivo que almacena múltiples hojas de cálculo y ofrece una amplia gama de funciones y herramientas para organizar, analizar y visualizar datos de manera efectiva
- Fórmulas y funciones: Excel cuenta con una amplia variedad de fórmulas y funciones predefinidas que permiten realizar cálculos y operaciones en los datos. Desde sumas y promedios hasta funciones más complejas como buscar, contar o condicionar valores, estas herramientas son fundamentales para realizar cálculos y análisis avanzados.
- Gráficos: Excel ofrece una amplia variedad de tipos de gráficos, como barras, líneas, áreas y pastel, que permiten visualizar datos de manera efectiva.
- Herramientas de análisis: Excel cuenta con herramientas de análisis que permiten realizar tareas más avanzadas, como análisis de regresión, análisis de hipótesis y pronósticos.

3. Insumos

Datos libres de superficies cultivada a través del Ministerio de Agricultura y Ganadería

Indican la estimación de superficie sembrada para los cultivos de arroz y maíz duro durante los periodos del 2015-2022 representado en hectáreas para las diferentes provincias del Ecuador.

Shape: Representan a los distintos cultivos en formato polígono para el cálculo de la superficie dentro del software.

Tipo de cultivo: este campo muestra si el cultivo pertenece a arroz o maíz duro.

Despro y descan: Corresponde a la división territorial tanto a nivel provincial como cantonal para los 2 cultivos a trabajar.

Tcu: Muestra el tipo de cultivo, siendo el maíz y arroz cereales

Sce: Aquí se encuentra calculado cada polígono o shape que representan a los dos cultivos de preferencia en hectáreas ya que es la unidad mayormente usada para estos casos.

Simbología: Para el cultivo de arroz la simbología es “CAcz: Arroz” mientras que para el maíz duro es “CACm: Maiz amarillo duro”.

Conjunto de datos utilizados: 2015, 2016, 2017, 2018, 2019, 2020, 2021 y 2022, disponible en el Sistema de Información Pública Agropecuaria ([SIPA](#)) del Ministerio de Agricultura y Ganadería (MAG) y en el mapa interactivo del [Geo portal del Agro Ecuatoriano](#).

4. Procedimientos

4.1 Estimación de superficies cultivadas

En esta sección se detalla el procedimiento a seguir para la proyección de la superficie cultivada de *Oryza sativa* (arroz) y *Zea mays* (maíz) hasta el año 2025 mediante un modelo de Regresión Lineal Simple (RLS).

La Regresión Lineal Simple (RLS)

Es un modelo matemático, en el cual la variable dependiente (y) se relaciona con una sola variable independiente (x). El estudio de la relación funcional matemática de estas variables se ajusta a un conjunto de datos mediante el método de los mínimos cuadrados (Palacios-cruz et al., 2013). Estas estimaciones se usan para formar la recta de mejor ajuste para un conjunto de datos determinado y representa la siguiente notación (Mendenhall et al., 2010):

$$y = \alpha + \beta x + \varepsilon$$

Donde:

α es el valor de la ordenada donde la línea de la regresión se intercepta con el eje y .

β es el coeficiente de regresión poblacional (pendiente de la línea recta) e indica el efecto de x en y .

ε es el error, es decir, es lo que no está explicada por la regresión lineal de y por parte de x .

y es la variable dependiente o predicha.

x es la variable independiente o predictora.

Coeficiente de determinación R^2

Es la proporción de la varianza total de la variable explicada (y) por la regresión, es decir, el porcentaje de explicación del modelo a la variable dependiente, por ende, indica la calidad del ajuste del modelo y permite su validación (Palacios-cruz et al., 2013). Dicho valor oscila entre 0 y 1, mientras más cercano a 1, la regresión tiene mayor ajuste y confiabilidad, no obstante, si se obtiene un valor de mayor de 0,7 es posible validar la RLS (Aragón, 2016).

Proyección año 2050

Dentro de los 4 periodos climáticos disponibles, se consideró el más cercano que corresponde al periodo hasta el año 2050 de los datos climáticos futuros de CMIP6 (Proyecto de Intercomparación de Modelos Acoplados) que incluyen los desarrollos más recientes en la modelización del clima global, estos consideran cuatro escenarios denominados trayectorias socioeconómicas compartidas (SSP, por sus siglas en inglés), que estudian el futuro en condiciones enfocándose en desafíos de mitigación y adaptación (ssp126, ssp245, ssp370 y ssp585) (Escoto, Sánchez, & Gachuz, 2017). Las variables climáticas que estudian

son temperatura, precipitación y las bioclimáticas que, dentro de la proyección de cultivos estos dos factores están estrechamente relacionados, pues así se asegura un crecimiento o pérdida de estos.

4.1.1 Depuración de datos

Consiste en obtener los datos de interés de un conjunto de bases de datos, es decir, en este caso, clasificar y extraer los datos de superficies cultivadas de los años 2015 a 2022 de *Oryza sativa* (arroz) y *Zea mays* (maíz) en formato *shapefile* (shp.) mediante herramientas del software ArcGIS 10.8., para luego obtener la superficie cultivada de cada especie vegetal por cantón.

a) Agregar archivos *shapefile*

Cargar los archivos *shapefile* (shp.) de las superficies cultivadas obtenidos del Ministerio de Agricultura al software ArcGIS 10.8, cabe mencionar que no es necesario configurar el sistema de referencia de coordenadas a WGS84 UTM zona 17S, puesto que, por defecto el software configura el sistema de referencia de acuerdo con el primer archivo vectorial agregado, en este sentido los *shapefile* cargados tienen la proyección WGS84 UTM zona 17S. Los archivos que se agregan son los siguientes:

CULTIVO_PERIODO1_2015_A

CULTIVO_PERIODO1_2016_A

CULTIVO_PERIODO1_2017_A

CULTIVO_PERIODO1_2018_A

CULTIVO_PERIODO1_2019_A

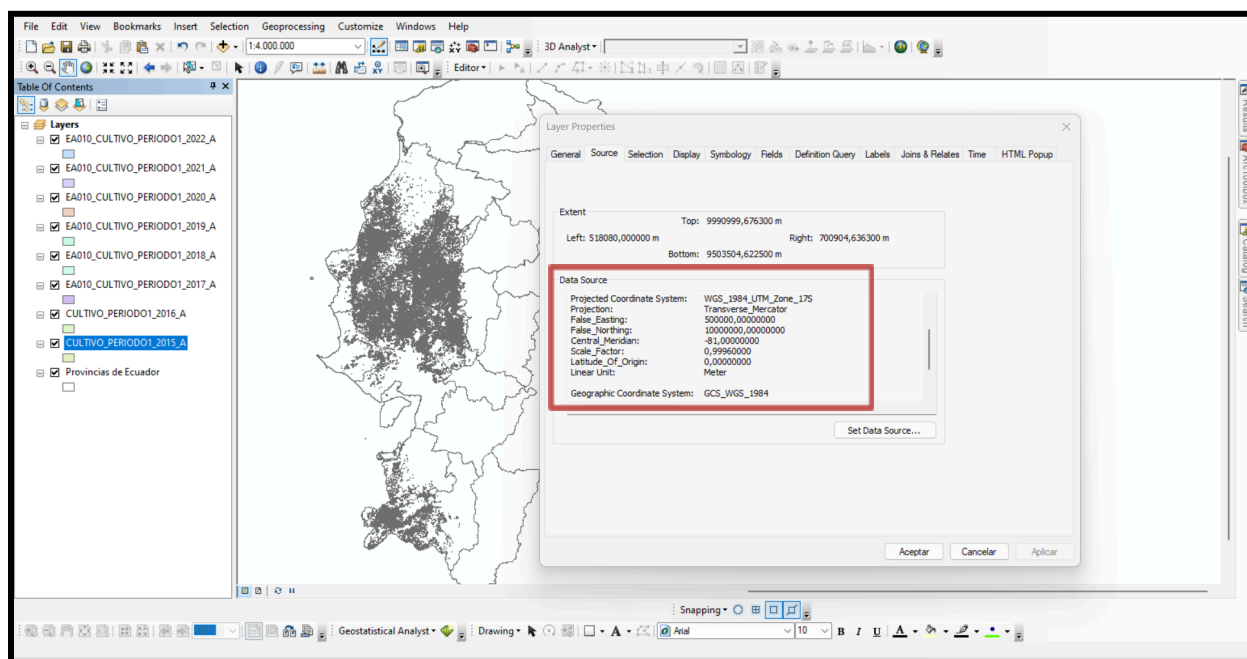
CULTIVO_PERIODO1_2020_A

CULTIVO_PERIODO1_2021_A

CULTIVO_PERIODO1_2022_A

Figura 1

Agregación de los archivos shapefile a Arcmap



b) Clasificación de datos de *Oryza sativa* (arroz) y *Zea mays* (maíz)

Abrir la tabla de atributos (1 en Figura 2) de la capa “CULTIVO_PERIODO1_2015_A”, dentro del cual hacemos uso de la herramienta “Select by Attributes” (2 en Figura 2) para seleccionar los polígonos de cada cultivo que están clasificados en el campo “ncu” (3 en Figura 2) o “simbología” (4 en Figura 2) de la tabla de atributos. Posteriormente, damos clic derecho a la capa y vamos a “Export Data” (1 en Figura 3), esta herramienta va a permitir exportar y guardar la selección de cada cultivo como un nuevo archivo shp., con el nombre y en la dirección de carpeta requerida (2 en Figura 3). Se debe realizar el mismo proceso para cada una de las capas *shapefile* de los años 2015 a 2022.

Figura 2

Clasificación de cultivos mediante la herramienta “Select by attributes” de la tabla de atributos

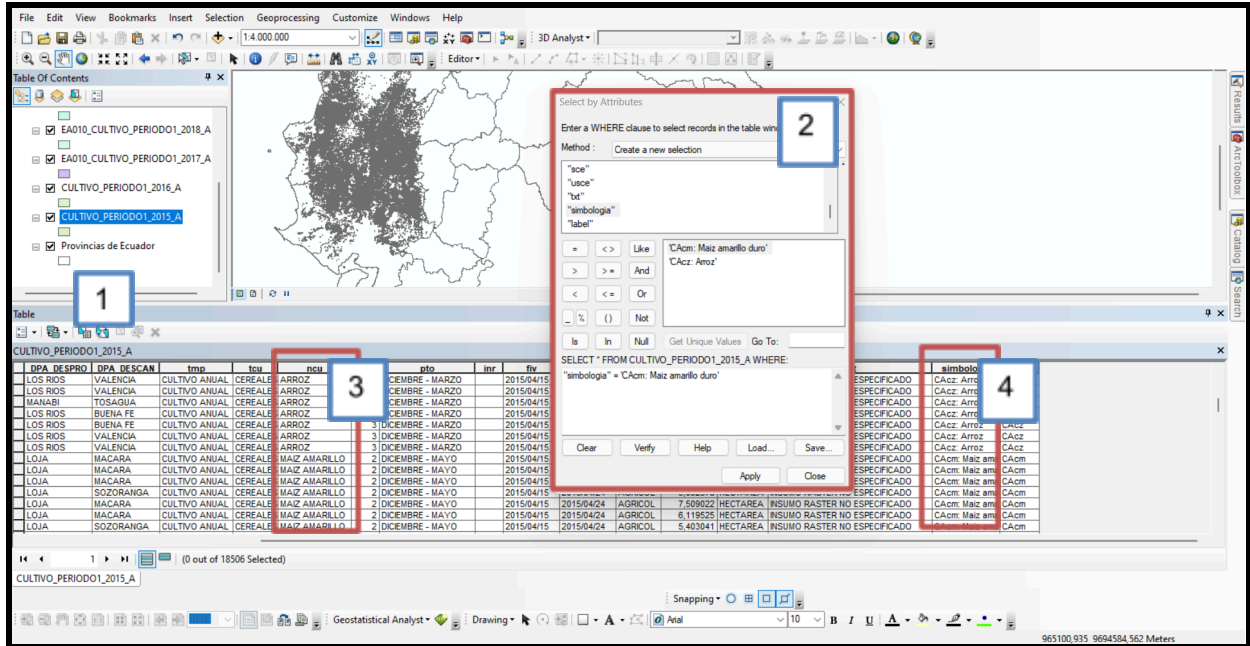
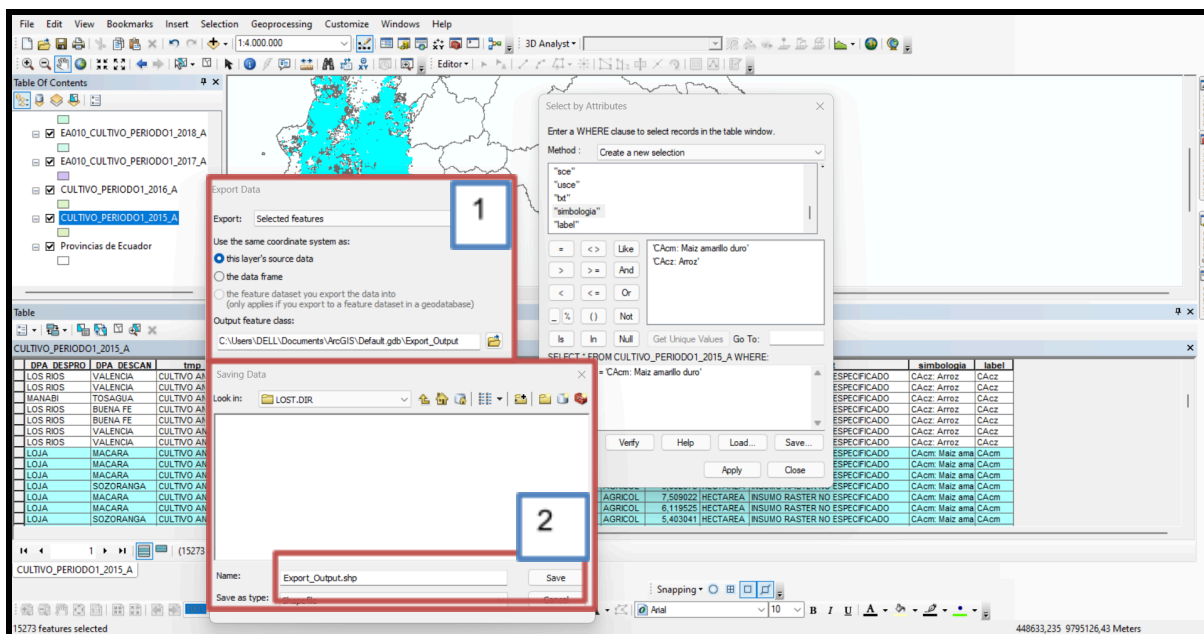


Figura 3

Exportación de los cultivos clasificados en nuevos shapefiles



c) Obtención de las superficies de cada cultivo por cantones

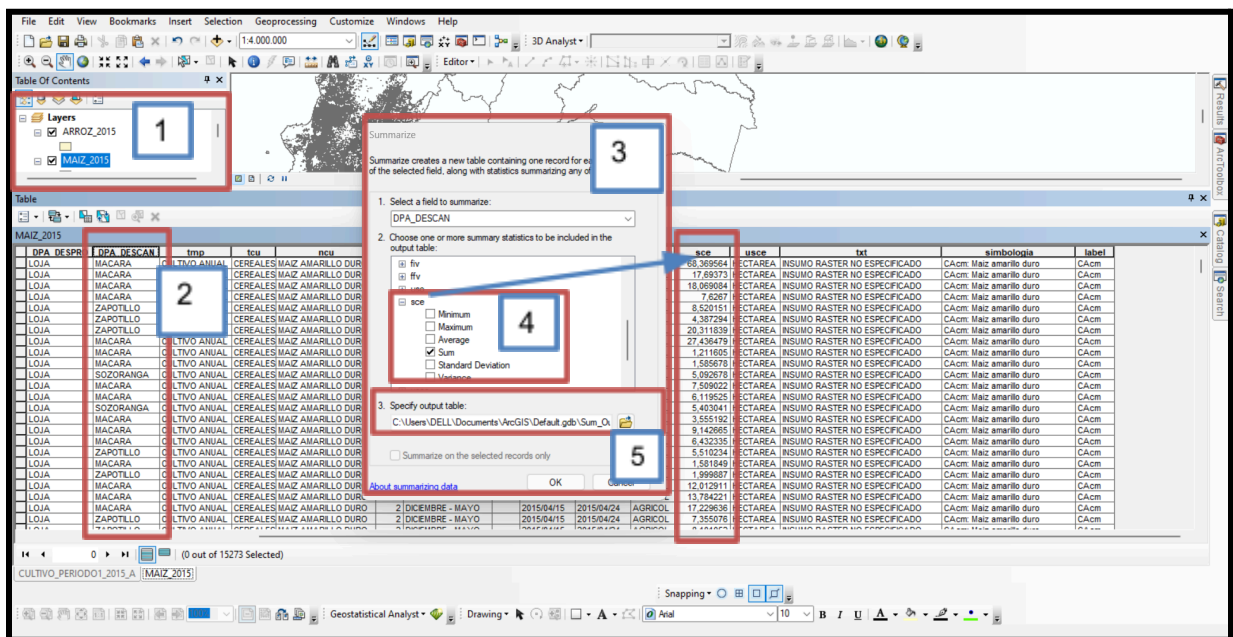
Calcular la superficie total cultivada de cada especie vegetal para cada cantón mediante la herramienta “summarize” en la tabla de atributos de los archivos shp., de las clasificaciones generadas anteriormente (2015 – 2022, por cada cultivo) (1 en Figura 4), esta herramienta permite calcular eficientemente la suma de las áreas de los polígonos de cultivos existentes en cada cantón. Para ello, dar clic derecho en la columna “DPA_DESCAN” (2 en Figura 4), posteriormente ir a “summarize” (3 en Figura 4) y seleccionar “Sum” en el campo “sce” (4 en Figura 4), esta última pertenece a la columna de superficie en hectáreas del cual se va a calcular la suma. Finalmente, guardar la tabla de la operación (5 en Figura 4).

Para la presente guía se seleccionó los siguientes cantones: Simón Bolívar, para el análisis del cultivo de arroz y; Jipijapa, para el cultivo de maíz duro, debido a que estas contaban con datos completos de las superficies

cultivadas para el periodo entre los años 2015-2022, además, el valor de r^2 de los modelos de regresión lineal fue más cercano a 1.

Figura 4

Obtención de las superficies cultivadas totales por cantón



4.1.2 Sistematización de datos

Este proceso consiste en crear y ajustar una matriz de datos con las superficies cultivadas de cada cantón ordenadas desde los años 2015 a 2022 para cada cultivo. De manera que permita crear una Regresión Lineal Simple.

a) Estructuración de las superficies cultivadas por cantón a través del tiempo

Las diferentes tablas con las superficies totales por cantones (1 en Figura 5) de cada año se deben reorganizar, para ello, se debe copiar dichas tablas a dos hojas de un archivo Excel (.xlsx), una hoja por cada cultivo de manera

ordenada (2 en Figura 5). Posteriormente, en una diferente hoja de Excel crear una matriz para los datos de superficie cultivada por cantón (y) y el periodo de tiempo en años (x) (1 en Figura 6), de cada cultivo. Para completar dicha matriz se va a utilizar la función “Buscarv” (2 en Figura 6) que nos permite buscar el valor de la superficie cultivada para el mismo cantón en diferentes periodos de tiempo de un conjunto de base de datos.

Figura 5

Ordenación por cantón y año de los datos de las tablas de superficies totales en Excel

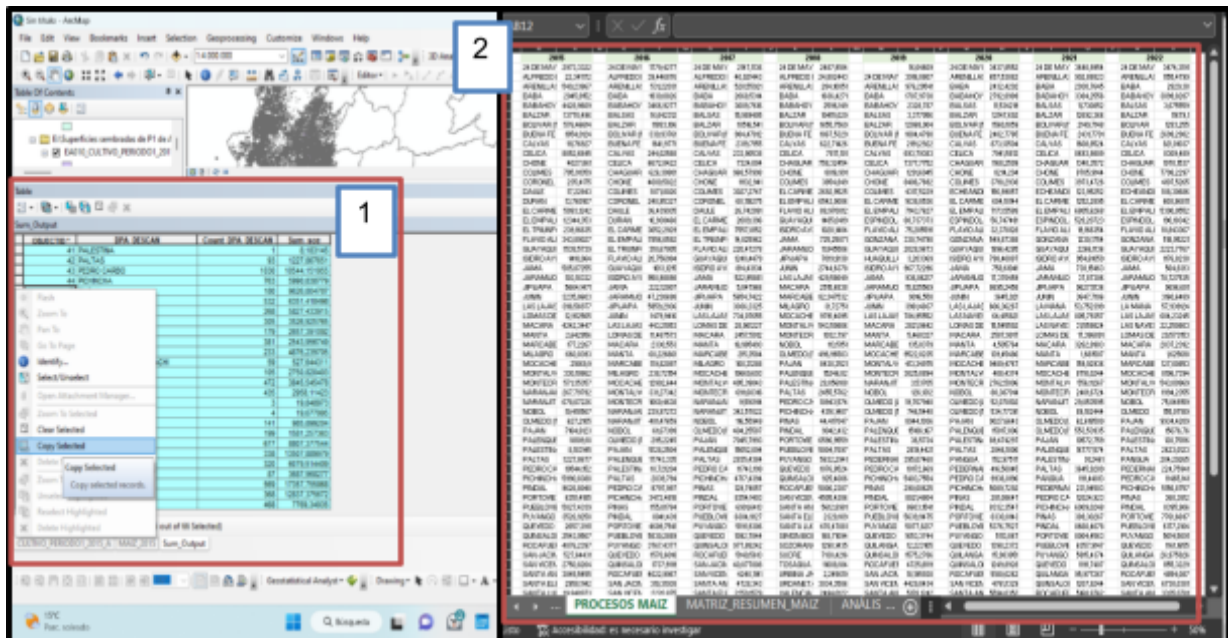


Figura 6

Creación y ajuste de la matriz de datos con las superficies (y) y años (x) por cada cantón

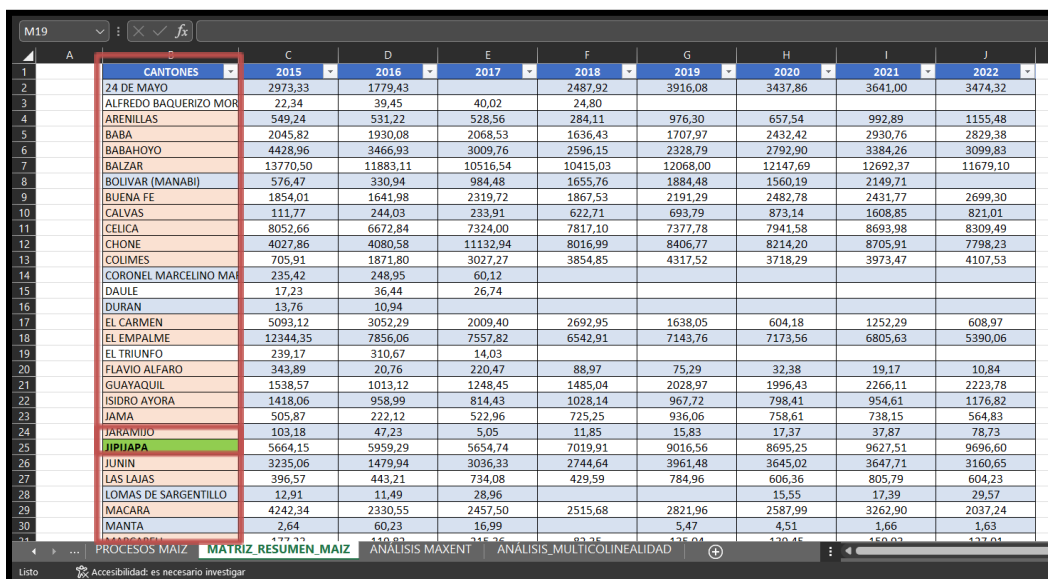
CANTONES	2015	2016	2017	2018	2019	2020	2021	2022
DECE MAYO	1779.83	1779.83	2487.83	2487.83	3918.08	3917.86	3641.80	3874.52
ALFREDO BACQUERIZO MORA	21.34	26.45	48.02	34.80				
ARTIMILLAS	549.14	531.22	528.56	354.11	335.58	697.34	591.89	1355.48
BAGUA	1949.92	1993.98	2088.93	1838.43	1720.92	2317.47	2920.76	1919.38
BAGUACHICO	4448.08	2466.52	3089.76	2586.15	2328.78	2782.80	2184.26	2698.82
BALLEN	13770.50	11883.11	10518.54	10415.83	12068.30	12347.68	13592.37	13670.30
BOLIVAR (MAMARU)	518.67	595.84	584.46	1855.76	1884.48	3385.39	2149.71	
BOLIVAR FE	1854.02	1641.98	2109.72	1867.53	2396.38	2482.38	2431.77	2498.36
COCHAS	111.77	244.82	233.91	622.71	693.79	873.34	658.85	815.02
COLETA	2023.08	4677.84	7524.82	7817.10	1717.78	7941.58	8051.58	8378.45
COYAMA	8823.86	43863.84	11132.04	8818.89	8888.72	8238.32	8135.81	1798.22
COYANES	205.92	1871.80	3027.27	2854.85	4217.52	2728.39	2671.47	4002.52
CORONEL MARCELO MORA	235.62	248.85	46.12					
EL CARMEN	59.36	16.94						
EL EMPALME	3251.12	3251.29	3080.40	2631.92	1838.05	684.38	1311.29	808.07
EL EMPALME	12344.80	3399.06	7337.82	3941.91	7343.78	7133.58	8895.82	3798.08
EL EMPALME	238.17	333.67	14.02					
ELVALEFARO	343.88	26.76	320.47	84.97	75.39	52.38	18.17	30.84
GUAYAZUL	1518.57	3713.12	5148.45	1492.04	3318.97	3295.43	2186.51	2213.78
GUAYAZUL	1418.08	998.89	814.45	1618.14	967.12	788.41	854.81	1178.82
GUAYAZUL	505.87	227.12	521.86	325.25	436.06	758.41	736.15	564.83
GUAYAZUL	505.18	41.22	3.08	11.85	16.82	13.87	12.82	38.73
GUAYAZUL	5864.13	3293.29	5654.34	7815.31	5913.58	8995.25	8621.51	5876.05
GUAYAZUL	1218.08	3474.84	3036.33	2744.64	3863.48	3045.82	3641.71	2368.65
LAS LAJAS	396.57	443.21	734.86	419.58	784.95	696.35	805.75	804.22
LAS LAJAS	51.81	13.49	38.98					
LAS LAJAS	6242.39	2383.55	2457.50	2515.88	2811.96	2387.99	3361.80	2817.24
LAS LAJAS	3.84	36.22	18.98	5.47				
LAS LAJAS	111.11	129.85	129.85	82.38	144.68	146.45	144.45	144.45

b) Identificación de cantones con las superficies cultivadas completas en el periodo 2015 a 2022

Se debe identificar aquellos cantones con datos completos y de mayor superficie cultivada del periodo 2015 (año 1) - 2022 (año 8), para cada cultivo (Figura 7). Para de este modo lograr desarrollar la Regresión Lineal Simple (RLS).

Figura 7

Identificación de cantones con datos completos en el periodo de tiempo considerado



CANTONES	2015	2016	2017	2018	2019	2020	2021	2022
24 DE MAYO	2973,33	1779,43		2487,92	3916,08	3437,86	3641,00	3474,32
ALFREDO BAQUERIZO MOR	22,34	39,45	40,02	24,80				
ARENILLAS	549,24	531,22	528,56	284,11	976,30	657,54	992,89	1155,48
BABA	2045,82	1930,08	2068,53	1636,43	1707,97	2432,42	2930,76	2829,38
BABAHOYO	4428,96	3466,93	3009,76	2596,15	2328,79	2792,90	3384,26	3099,83
BALZAR	13770,50	11883,11	10516,54	10415,03	12068,00	12147,69	12692,37	11679,10
BOLIVAR (MANABI)	576,47	330,94	984,48	1655,76	1884,48	1560,19	2149,71	
BUENA FE	1854,01	1641,98	2319,72	1867,53	2191,29	2482,78	2431,77	2699,30
CALVAS	111,77	244,03	233,91	622,71	693,79	873,14	1608,85	821,01
CELICA	8052,66	6672,84	7324,00	7817,10	7377,78	7941,58	8693,98	8309,49
CHONE	4027,86	4080,58	11132,94	8016,99	8406,77	8214,20	8705,91	7798,23
COLIMES	705,91	1871,80	3027,27	3854,85	4317,52	3718,29	3973,47	4107,53
CORONEL MARCELINO MA	235,42	248,95	60,12					
DAULE	17,23	36,44	26,74					
DURAN	13,76	10,94						
EL CARMEN	5093,12	3052,29	2009,40	2692,95	1638,05	604,18	1252,29	608,97
EL EMPALME	12344,35	7856,06	7557,82	6542,91	7143,76	7173,56	6805,63	5390,06
EL TRIUNFO	239,17	310,67	14,03					
FLAVIO ALFARO	343,89	20,76	220,47	88,97	75,29	32,38	19,17	10,84
GUAYACUIL	1538,57	1013,12	1248,45	1485,04	2028,97	1996,43	2266,11	2223,78
ISIDRO AYORA	1418,06	958,99	814,43	1028,14	967,72	798,41	954,61	1176,82
JAMA	505,87	222,12	522,95	725,25	936,06	758,61	738,15	564,83
JARAMIJO	103,38	47,23	5,05	11,85	15,83	17,37	37,87	78,73
JIPIJAPA	5664,15	5959,29	5654,74	7019,91	9016,56	8695,25	9627,51	9696,60
JUNIN	3235,06	1479,94	3036,33	2744,64	3961,48	3645,02	3647,71	3160,65
LAS LAJAS	396,57	443,21	734,08	429,59	784,96	606,36	805,79	604,23
LOMAS DE SARGENTILLO	12,91	11,49	28,96			15,55	17,39	29,57
MACARA	4242,34	2330,55	2457,50	2515,68	2821,96	2587,99	3262,90	2037,24
MANTA	2,64	60,23	16,99		5,47	4,51	1,66	1,63

4.1.3 Creación del modelo de Regresión Lineal Simple (RLS)

Este proceso consiste en crear una Regresión Lineal Simple con la matriz anterior. Es decir, se va a generar un diagrama de dispersión para obtener de manera eficiente la ecuación de la recta y, de esta manera, estimar la superficie cultivada de *Oryza sativa* (arroz) y *Zea mays* (maíz), para los periodos sugeridos, en este caso hasta 2025.

a. Normalidad de los datos

La distribución normal de datos constituye un postulado de la regresión Lineal Simple. La distribución normal de una variable está determinada por dos parámetros: la media y la desviación estándar, con esta se determina la densidad normal de datos, que generalmente se representa dentro del área bajo la curva delimitada por dos ejes perpendiculares al eje x, conocida como la campana de Gauss (Pértegas & Pita, 2001).

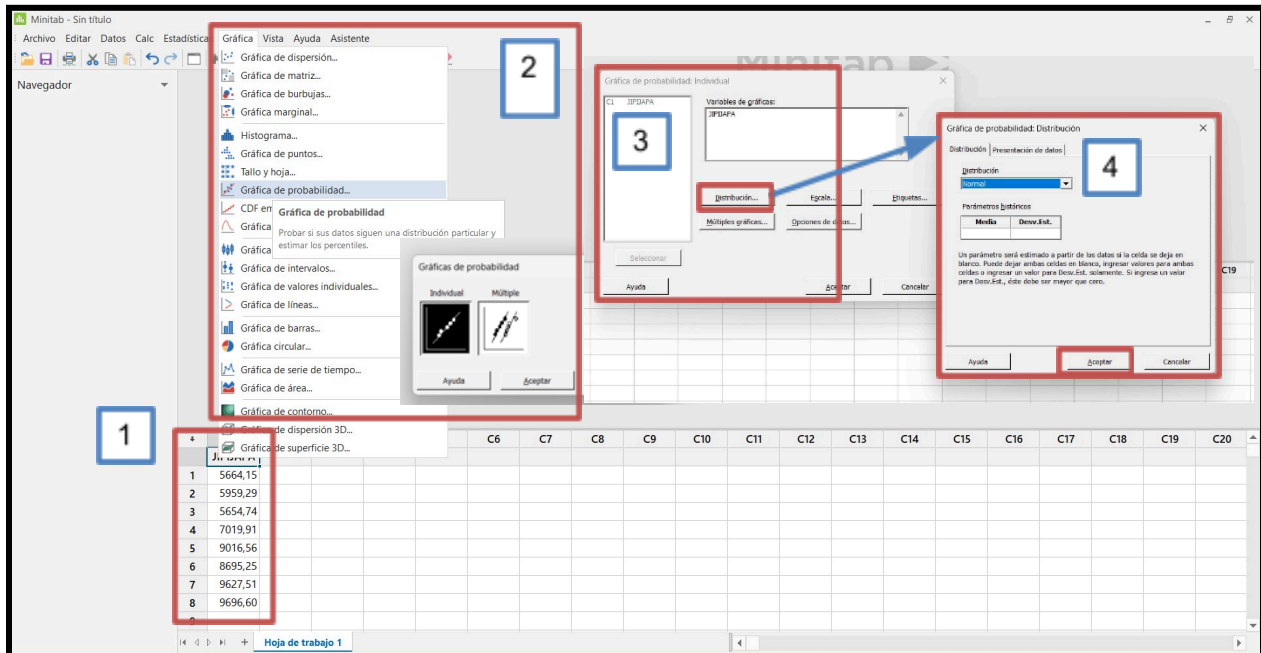
Según Dagnino (2014) “la curva alcanza su mayor altura en torno a la media, mientras que sus “ramas” se extienden asintóticamente hacia los ejes”, de este modo, si una variable tiene una distribución normal, es más probable observar un dato cercano al valor medio que uno que se encuentre en los extremos.

Existen diferentes pruebas para determinar la normalidad, entre estas se encuentran: las gráficas de probabilidad normal Q-Q plot o también llamado la prueba del lápiz grueso + p valor; prueba de normalidad de Anderson-Darling; prueba de Ryan-Joiner; prueba de normalidad de Kolmogorov Smirnov, entre otros (Flores & Flores, 2021). No obstante, es importante considerar el tamaño de los datos a analizar, puesto que ciertas pruebas tienen restricciones.

Para determinar si el conjunto de datos tiene una distribución normal mediante el uso del software Minitab 19.1. Para ello, transcribir los datos de las superficies cultivadas (y) y sus respectivos años (x) de un cantón específico (1 en Figura 8). Luego, ir a la pestaña **Gráfica** ⇒ **Gráfica de probabilidad** ⇒ **Individual** (2 en Figura 8). Posteriormente, seleccionar la columna de los datos y dar clic a **Distribución de probabilidad** (3 en Figura 8) ⇒ **Distribución** ⇒ **Normal** (4 en Figura 8) y, finalmente, aceptar.

Figura 8

Cálculo de normalidad de datos en Minitab 19.1



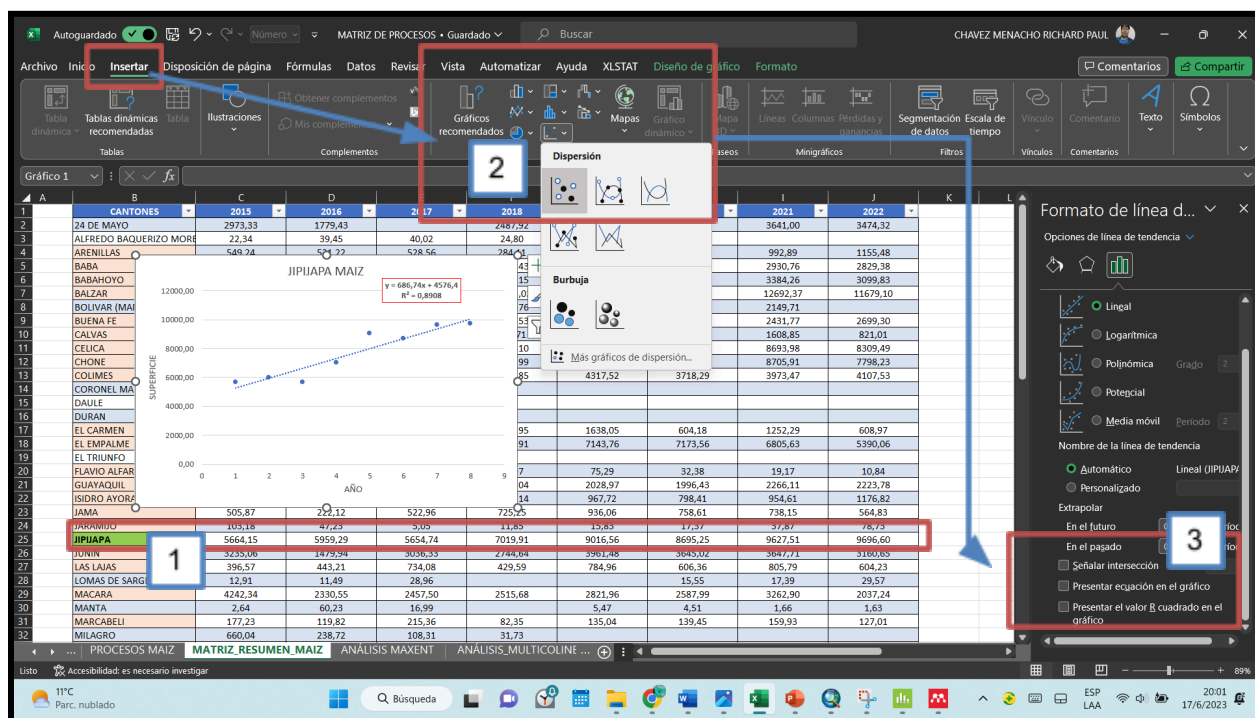
b. Diagrama de dispersión y ecuación de la recta

Los diagramas de dispersión o gráficos de correlación lineal representan la relación entre dos variables, ya que cada par de valores están representados como coordenadas de un punto (X_i, Y_i) , de este modo, permiten determinar la existencia de una relación entre las variables y obtener de manera eficiente la ecuación de la recta que se ajusta a un conjunto de datos mediante el método de los mínimos cuadrados. Además, permite conocer el coeficiente de determinación R^2 que indica la bondad del ajuste del modelo a la variable respuesta (y: superficie cultivada), es decir, si se valida o no el modelo de RLS.

Para generar este diagrama se debe seleccionar el periodo de tiempo (x) y los datos de superficie (y) de un determinado cantón (1 en Figura 9), e insertar un diagrama desde la pestaña **insertar** ⇒ **Gráficos** ⇒ **Dispersión** (2 en Figura 9). Así como también, indicar que se presenten la ecuación de la recta y el valor de R² en el gráfico dando clic en la **opción más (+)** de la **gráfica** ⇒ **Línea de tendencia** ⇒ **Más opciones** (3 en Figura 9).

Figura 9

Obtención de la ecuación de la recta y el R² a través de la gráfica de dispersión



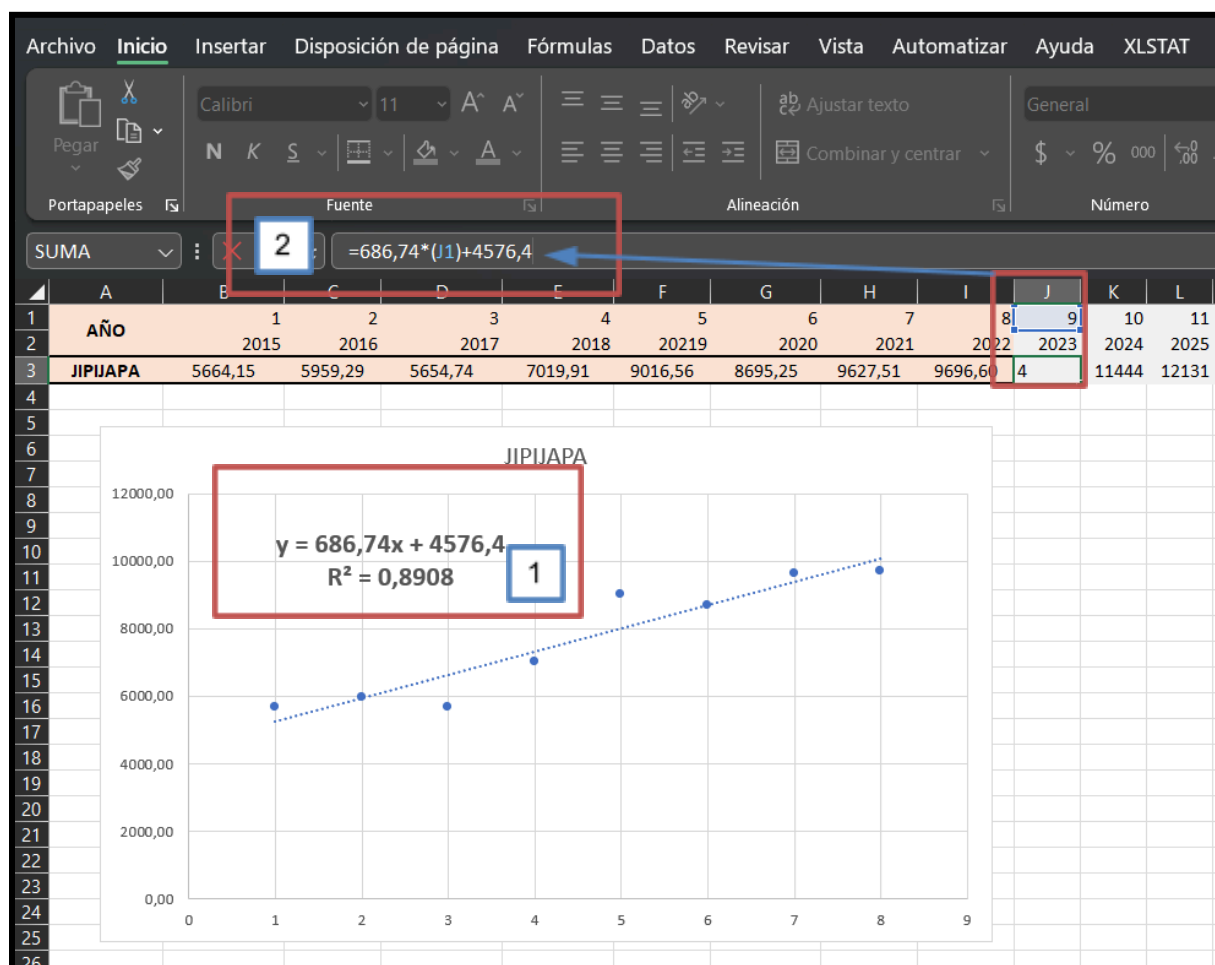
c. Estimación de la superficie cultivada hasta 2025

Estimar la superficie cultivada para los años sugeridos (2023 - 2025) mediante el uso de la ecuación de la recta ($y = a + \Delta x$) obtenido del gráfico de dispersión (1 en Figura 10). En el que, y es la variable dependiente (superficie

cultivada), α es la constante de la ecuación, Δ es el coeficiente del modelo que indica el efecto de x en y , por último, x es la variable independiente (tiempo en años). Por ende, para estimar la superficie cultivada se debe reemplazar los datos de los años en la variable x de la ecuación (2 en Figura 10).

Figura 10

Estimación de las superficies cultivadas hasta el año 2025



4.2 Estimación de las superficies idóneas de mediante MaxEnt Modelos de distribución de especies (SDM)

El modelo MaxEnt, también conocido como modelo de entropía máxima, es ampliamente utilizado en estadística y aprendizaje automático. Se basa en el principio de máxima entropía para seleccionar la distribución de probabilidad más equitativa o imparcial, dado un conjunto de restricciones o información disponible. Es una herramienta poderosa en aplicaciones como el procesamiento del lenguaje natural y la clasificación de textos.

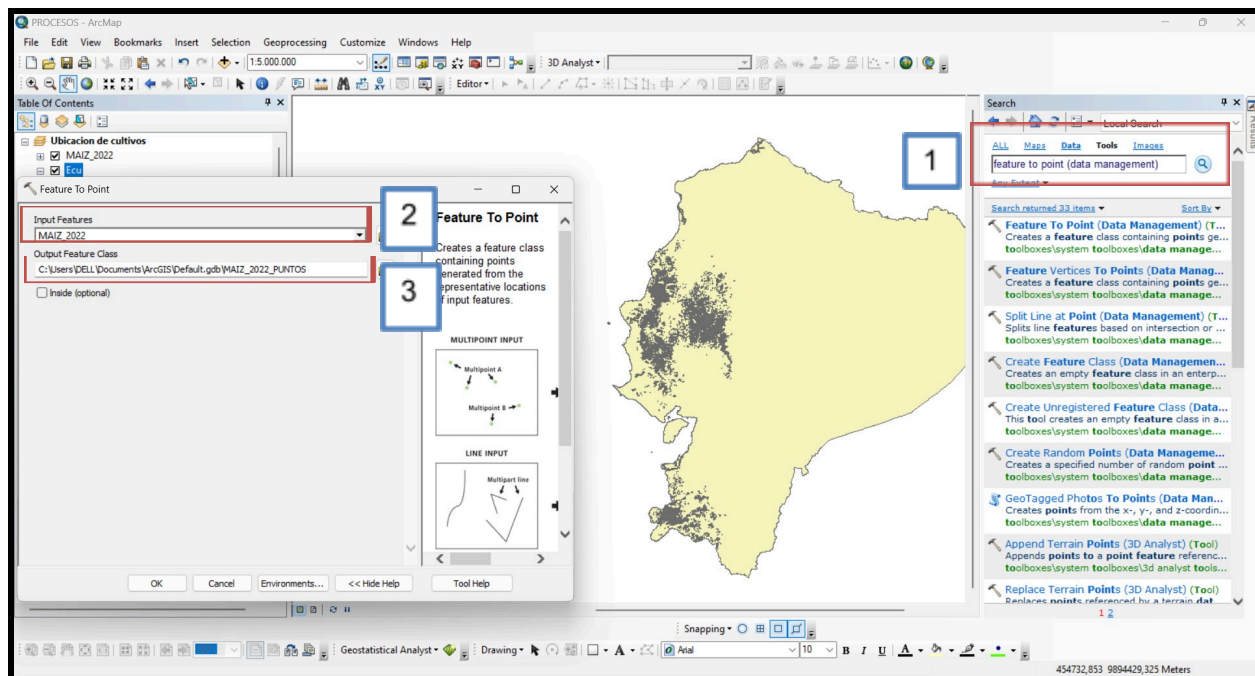
4.2.1 Obtención de datos

a) Puntos de presencia

Para la estimación de superficies idóneas hasta el periodo 2050 es necesario conocer la ubicación de los cultivos de arroz y maíz que está distribuidas a nivel cantonal, para ello en la barra de búsqueda de ArcMap se debe escribir “feature to point” (1 en figura 11), y colocar en la parte de “input features” a los shapes de cultivos (2 en figura 11) y en “output feature class” la dirección en la que se guardará (3 en figura 11). Esto permitirá más adelante tomar una muestra de todos los puntos que simbolizan a un tipo de cultivo y proyectar las superficies.

Figura 11

Transformación de polígonos de los cultivos a puntos



b) Variables bioclimáticas y biofísicas

Las variables como temperatura o precipitación influyen dentro del crecimiento de los cultivos de arroz y maíz duro, además de que la proyección requiere trabajar con estos datos para un análisis más completo y confiable. De este modo, del portal de [worldclim](http://worldclim.org) se debe descargar las 19 variables bioclimáticas (Tabla 1) y biofísica (elevación) (figura 12), dentro del portal existen 4 diferentes resoluciones disponibles para el modelo, no obstante, es recomendable optar por la de mayor resolución para que el mapa a futuro se pueda comprender de mejor manera. Por ende, se opta por descargar la de 30s.

Las variables climáticas futuras para realizar el mapa de proyección se obtienen del mismo portal, para ello se debe escoger la opción de “future climate data” (1 en Figura 13), el periodo (2021-2050), Simulador del sistema climático y terrestre de la comunidad australiana (ACCES-CM2) y el escenario climático ssp (trayectorias socioeconómicas compartida) (2 en Figura 13) que indican lo

siguiente:

ssp126: modelo positivista

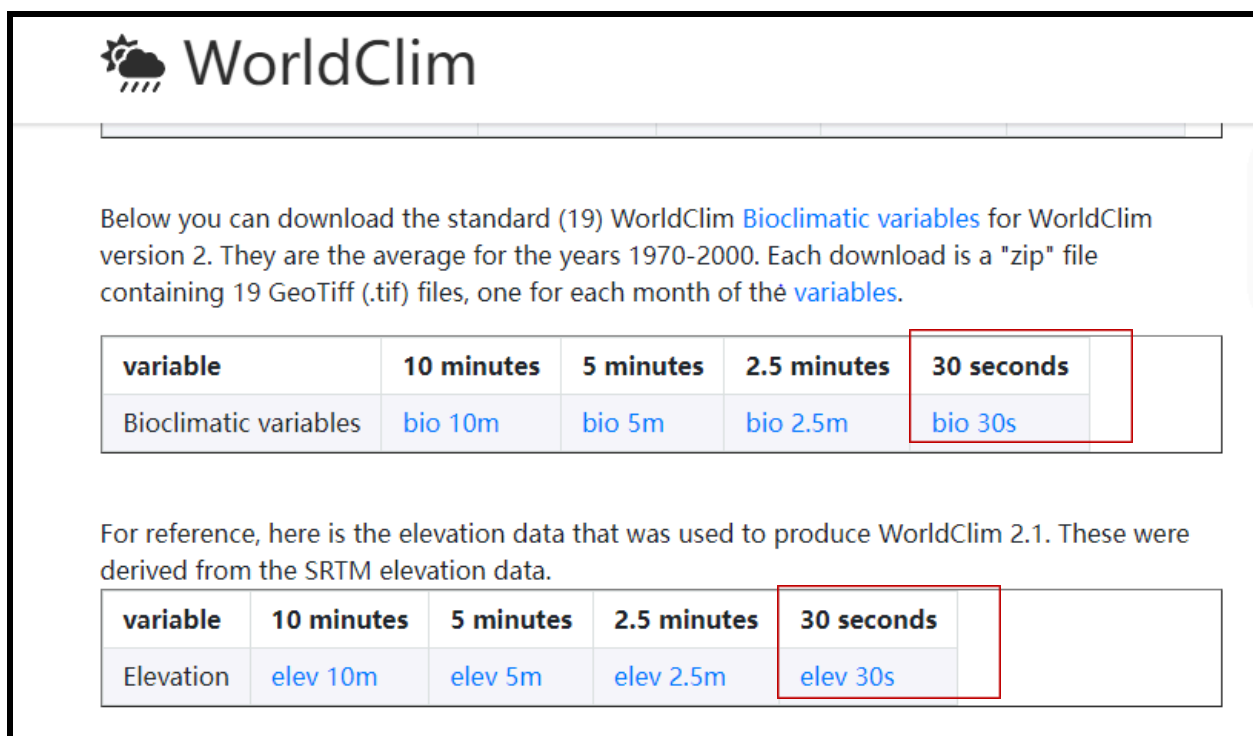
ssp245: modelo medianamente positivo


ssp370: modelo medianamente negativo

ssp585: modelo pesimista

Figura 12

Descarga de variables del portal worldclim



 WorldClim

Below you can download the standard (19) WorldClim [Bioclimatic variables](#) for WorldClim version 2. They are the average for the years 1970-2000. Each download is a "zip" file containing 19 GeoTiff (.tif) files, one for each month of the [variables](#).

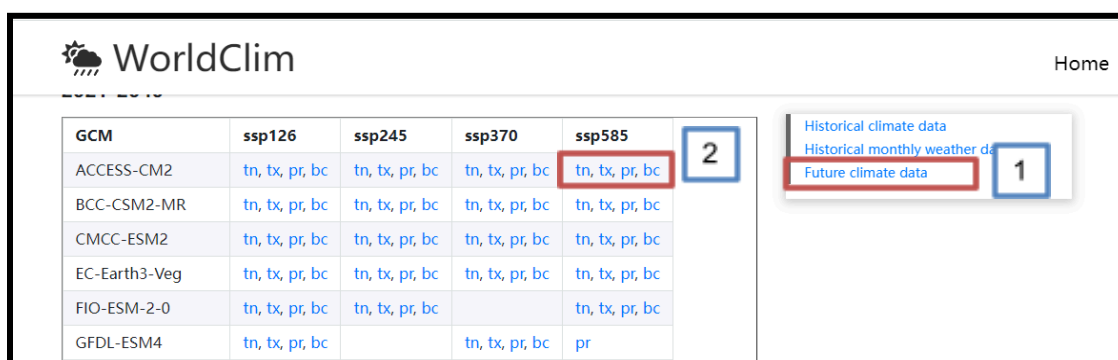
variable	10 minutes	5 minutes	2.5 minutes	30 seconds
Bioclimatic variables	bio 10m	bio 5m	bio 2.5m	bio 30s

For reference, here is the elevation data that was used to produce WorldClim 2.1. These were derived from the SRTM elevation data.

variable	10 minutes	5 minutes	2.5 minutes	30 seconds
Elevation	elev 10m	elev 5m	elev 2.5m	elev 30s

Figura 13

Obtención de variables climáticas futuras



The screenshot shows the WorldClim website interface. At the top left is the WorldClim logo and at the top right is a 'Home' link. Below the header is a table with columns for GCM, ssp126, ssp245, ssp370, and ssp585. The ssp585 column for the first row is highlighted with a red box and a blue box containing the number '2'. To the right of the table is a dropdown menu with three options: 'Historical climate data', 'Historical monthly weather data', and 'Future climate data'. The 'Future climate data' option is highlighted with a red box and a blue box containing the number '1'.

GCM	ssp126	ssp245	ssp370	ssp585
ACCESS-CM2	tn, tx, pr, bc	tn, tx, pr, bc	tn, tx, pr, bc	tn, tx, pr, bc
BCC-CSM2-MR	tn, tx, pr, bc	tn, tx, pr, bc	tn, tx, pr, bc	tn, tx, pr, bc
CMCC-ESM2	tn, tx, pr, bc	tn, tx, pr, bc	tn, tx, pr, bc	tn, tx, pr, bc
EC-Earth3-Veg	tn, tx, pr, bc	tn, tx, pr, bc	tn, tx, pr, bc	tn, tx, pr, bc
FIO-ESM-2-0	tn, tx, pr, bc	tn, tx, pr, bc		tn, tx, pr, bc
GFDL-ESM4	tn, tx, pr, bc		tn, tx, pr, bc	pr

Tabla 1

Variables bioclimáticas

BIO1	Temperatura Media Anual
BIO2	Intervalo medio diurno (media mensual (temperatura máxima - temperatura mínima))
BIO3	Isotermalidad (BIO2/BIO7) (×100)
BIO4	Estacionalidad de la temperatura (desviación estándar ×100)
BIO5	Temperatura máxima del mes más cálido
BIO6	Temperatura mínima del mes más frío
BIO7	Rango Anual de Temperatura (BIO5-BIO6)

BIO8	Temperatura media del trimestre más húmedo
BIO9	Temperatura Media del Cuarto Más Seco
BIO10	Temperatura media del trimestre más cálido
BIO11	Temperatura media del trimestre más frío
BIO12	Precipitación Anual
BIO13	Precipitación del mes más lluvioso
BIO14	Precipitación del mes más seco
BIO15	Estacionalidad de la Precipitación (Coeficiente de Variación)
BIO16	Precipitación del trimestre más húmedo
BIO17	Precipitación del Trimestre Más Seco
BIO18	Precipitación del trimestre más cálido
BIO19	Precipitación del trimestre más frío

(WorldClim, 2023)

4.2.2 Preparación y estandarización de datos

a) Muestra de puntos de presencia de cultivos

Los puntos de ubicación de las superficies cultivadas de maíz y arroz se deben analizar de manera que de estas se extraiga una muestra representativa, para ello se calcula el tamaño de la muestra mediante la siguiente ecuación para una población finita:

$$n = \frac{N \cdot Z^2 \cdot p \cdot q}{e^2(N-1) + Z^2 \cdot p \cdot q}$$

Donde:

n = Tamaño de la muestra

N = Población total

Z = Nivel de confianza

e = Margen de error

p = Probabilidad de éxito o proporción esperada

q = Probabilidad de fracaso (1 - p)

(Aguilar-Barojas,

2005)

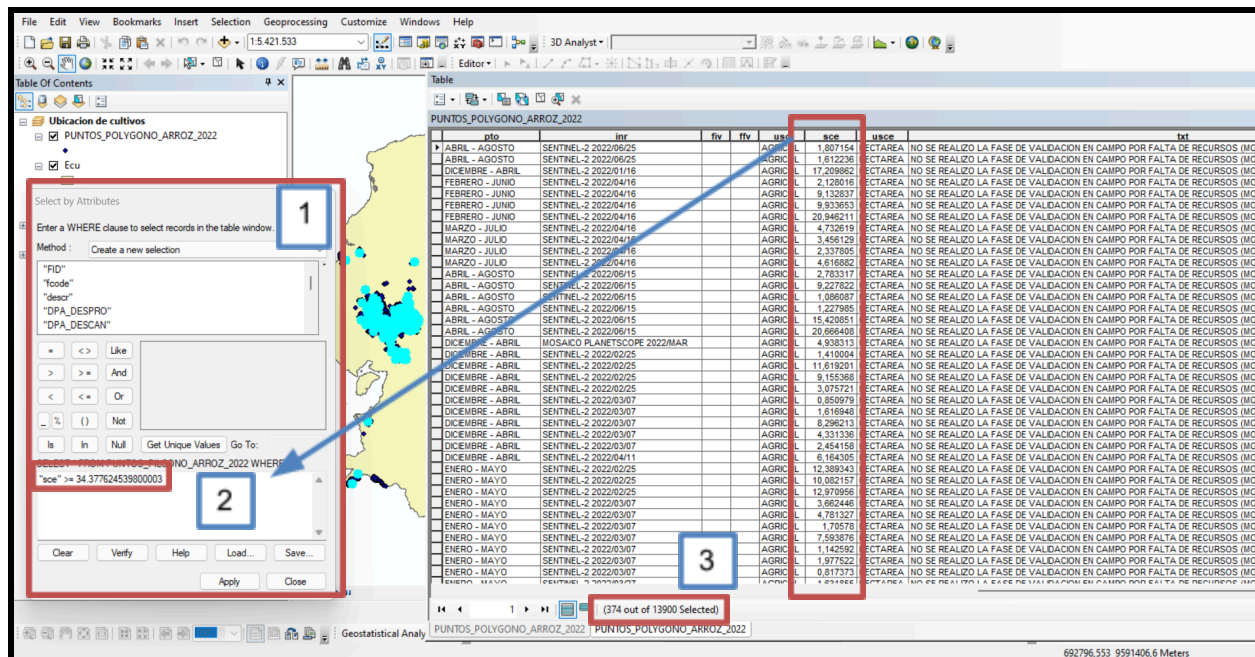
A continuación, se detallan los coeficientes de la ecuación anterior y el valor de cada una de ellas mediante el cual se calcula el tamaño de la muestra. En este sentido, la población total (N) hace referencia a la cantidad de puntos de ubicación de las superficies de los cultivos (13 900 puntos para el arroz y 36 198 para el maíz), nivel de confianza (Z) es el grado de certeza con el que se requiere trabajar en la muestra y para obtener una muestra confiable se debe optar por considerar un valor de 1.96 (95%), margen de error (e) es el error de estimación máximo esperado y se recomienda trabajar con un margen de error de 5%, por otro lado, p y q conjuntamente indican la probabilidad de que exista o no éxito del evento estudiado a través del muestreo representativo, en este caso para ambos generalmente se considera el 0,5 (5%) (Aguilar-Barojas, 2005).

Se recomienda que la muestra puede ser mensurada de manera eficiente mediante el uso de la plataforma [SurveyMonkey](https://www.surveymonkey.com), en el cual se debe colocar el tamaño de la población, nivel de confianza (%) y el margen de error (%).

Por otro lado, es importante mencionar que para la muestra se debe considerar los polígonos de mayor superficie, de este modo, mediante la herramienta “Select by attributes” (1 en Figura 14), seleccionar los de mayor superficie aplicando la fórmula pertinente y considerando la columna “sca” de superficie (2 en Figura 14) hasta llegar al tamaño de la muestra (3 en figura 14). Cabe mencionar que, este proceso se lo debe realizar para ambos cultivo.

Figura 14

Obtención de polígonos de mayor superficie de acuerdo con el tamaño de la muestra



b) Análisis de multicolinealidad

Las 19 variables bioclimáticas obtenidas de wordclim deben ser analizadas de modo que no exista correlación entre estas, es decir, evitar la multicolinealidad. Cuando se presenta multicolinealidad, puede ser difícil determinar la contribución individual de cada variable y puede afectar la precisión del modelo. La multicolinealidad puede ser detectada utilizando el método de los valores absolutos de los coeficientes de correlación por pares y tomando en cuenta un umbral de 0,7. Al seleccionar variables con valores bajos de 0,7, se busca promover la independencia de las variables y asegurar que cada una aporte información única al análisis, permitiendo así una interpretación más precisa y confiable de los datos.

En este caso se selección las variables con baja correlación con respecto a las demás como: “bio2” “bio3” “bio4” “bio5” “bio6” “bio7” “bio12” “bio13” “bio14” “bio15” “bio16” “bio17” “bio18” “bio19” y “elev”.

Figura 15

Matriz de correlación en Excel

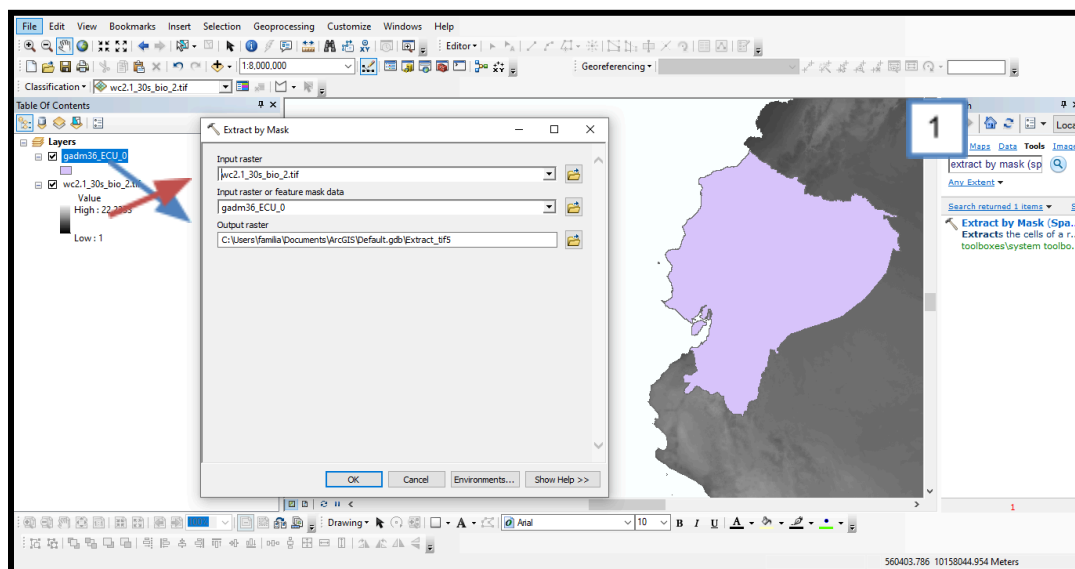
Variables	bio1	bio2	bio3	bio4	bio5	bio6	bio7	bio8	bio9	bio10	bio11	bio12	bio13	bio14	bio15	bio16	bio17	bio18	bio19	elev
bio1	1,00	0,60	0,26	0,37	1,00	1,00	0,41	1,00	1,00	1,00	1,00	0,53	0,67	0,33	0,29	0,66	0,33	0,64	0,32	0,99
bio2	0,60	1,00	0,41	0,36	0,55	0,64	0,72	0,62	0,57	0,61	0,59	0,26	0,48	0,03	0,37	0,49	0,04	0,59	0,01	0,66
bio3	0,26	0,41	1,00	0,94	0,31	0,22	0,33	0,34	0,19	0,32	0,21	0,45	0,17	0,51	0,67	0,19	0,53	0,07	0,52	0,34
bio4	0,37	0,36	0,94	1,00	0,41	0,32	0,35	0,44	0,30	0,42	0,32	0,32	0,08	0,36	0,63	0,11	0,39	0,00	0,37	0,45
bio5	1,00	0,55	0,31	0,41	1,00	0,99	0,33	0,99	0,99	1,00	0,99	0,48	0,63	0,31	0,31	0,62	0,30	0,60	0,29	0,99
bio6	1,00	0,64	0,22	0,32	0,99	1,00	0,49	0,99	0,99	0,99	1,00	0,56	0,70	0,35	0,27	0,69	0,35	0,67	0,34	0,99
bio7	0,41	0,72	0,33	0,35	0,33	0,49	1,00	0,38	0,45	0,38	0,45	0,61	0,65	0,40	0,10	0,67	0,43	0,68	0,39	0,41
bio8	1,00	0,62	0,34	0,44	0,99	0,99	0,38	1,00	0,98	1,00	0,99	0,46	0,63	0,26	0,36	0,62	0,25	0,63	0,24	1,00
bio9	1,00	0,57	0,19	0,30	0,99	0,99	0,45	0,98	1,00	0,99	1,00	0,59	0,70	0,41	0,21	0,69	0,40	0,65	0,40	0,98
bio10	1,00	0,61	0,32	0,42	1,00	0,99	0,38	1,00	0,99	1,00	0,99	0,49	0,65	0,30	0,33	0,64	0,29	0,63	0,28	1,00
bio11	1,00	0,59	0,21	0,32	0,99	1,00	0,45	0,99	1,00	0,99	1,00	0,55	0,69	0,35	0,27	0,68	0,35	0,66	0,34	0,99
bio12	0,53	0,26	0,45	0,32	0,48	0,56	0,61	0,46	0,59	0,49	0,55	1,00	0,84	0,90	0,44	0,84	0,91	0,68	0,90	0,49
bio13	0,67	0,48	0,17	0,08	0,63	0,70	0,65	0,63	0,70	0,65	0,69	0,84	1,00	0,54	0,06	1,00	0,54	0,93	0,53	0,66
bio14	0,33	0,03	0,51	0,36	0,31	0,35	0,40	0,26	0,41	0,30	0,35	0,90	0,54	1,00	0,70	0,55	1,00	0,34	0,98	0,28
bio15	0,29	0,37	0,67	0,63	0,31	0,27	0,10	0,36	0,21	0,33	0,27	0,44	0,06	0,70	1,00	0,04	0,71	0,24	0,70	0,36
bio16	0,66	0,49	0,19	0,11	0,62	0,69	0,67	0,62	0,69	0,64	0,68	0,84	1,00	0,55	0,04	1,00	0,55	0,93	0,54	0,65
bio17	0,33	0,04	0,53	0,39	0,30	0,35	0,43	0,25	0,40	0,29	0,35	0,91	0,54	1,00	0,71	0,55	1,00	0,35	0,98	0,27
bio18	0,64	0,59	0,07	0,00	0,60	0,67	0,68	0,63	0,65	0,63	0,66	0,68	0,93	0,34	0,24	0,93	0,35	1,00	0,31	0,65
bio19	0,32	0,01	0,52	0,37	0,29	0,34	0,39	0,24	0,40	0,28	0,34	0,90	0,53	0,98	0,70	0,54	0,98	0,31	1,00	0,26
elev	0,99	0,66	0,34	0,45	0,99	0,99	0,41	1,00	0,98	1,00	0,99	0,49	0,66	0,28	0,36	0,65	0,27	0,65	0,26	1,00

c) Estandarización de formatos

Área de estudio: Cuando se obtiene las variables a utilizar se debe delimitar el área de estudio, para este caso el Ecuador continental, para ello se utiliza la herramienta “extract by mask” (1 en Figura 16).

Figura 16

Cortes de las variables al área de estudio

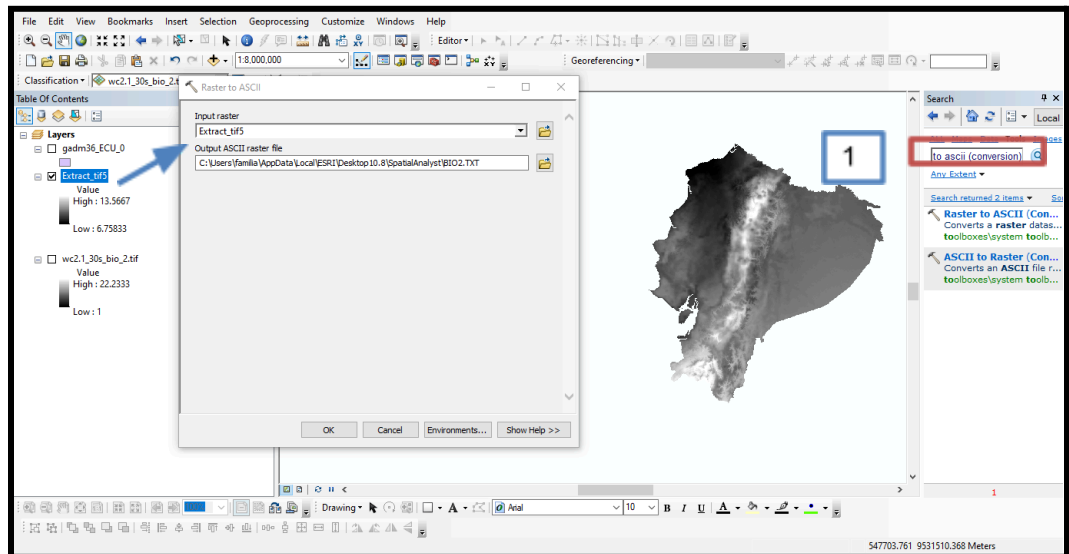


Los modelos predictivos de Maxent son generados en formato ASCII por lo que resulta necesario transformar las capas de las variables recortadas al formato ASCII, con la herramienta “raster to ascii (conversion)” (1 en Figura 17).

El formato ASCII (American Standard Code for Information Interchange) es un sistema de codificación desarrollado como un estándar para su procesamiento y almacenamiento en sistemas informáticos.

Figura 17

Transformación de raster a formato ASCII



Otro insumo necesario para el modelo MaxEnt es un archivo en formato .CSV donde conste el nombre de la especie y las coordenadas x e y. Para ello se utiliza la fórmula de Excel “=Concat” (Figura 18).

Figura 18

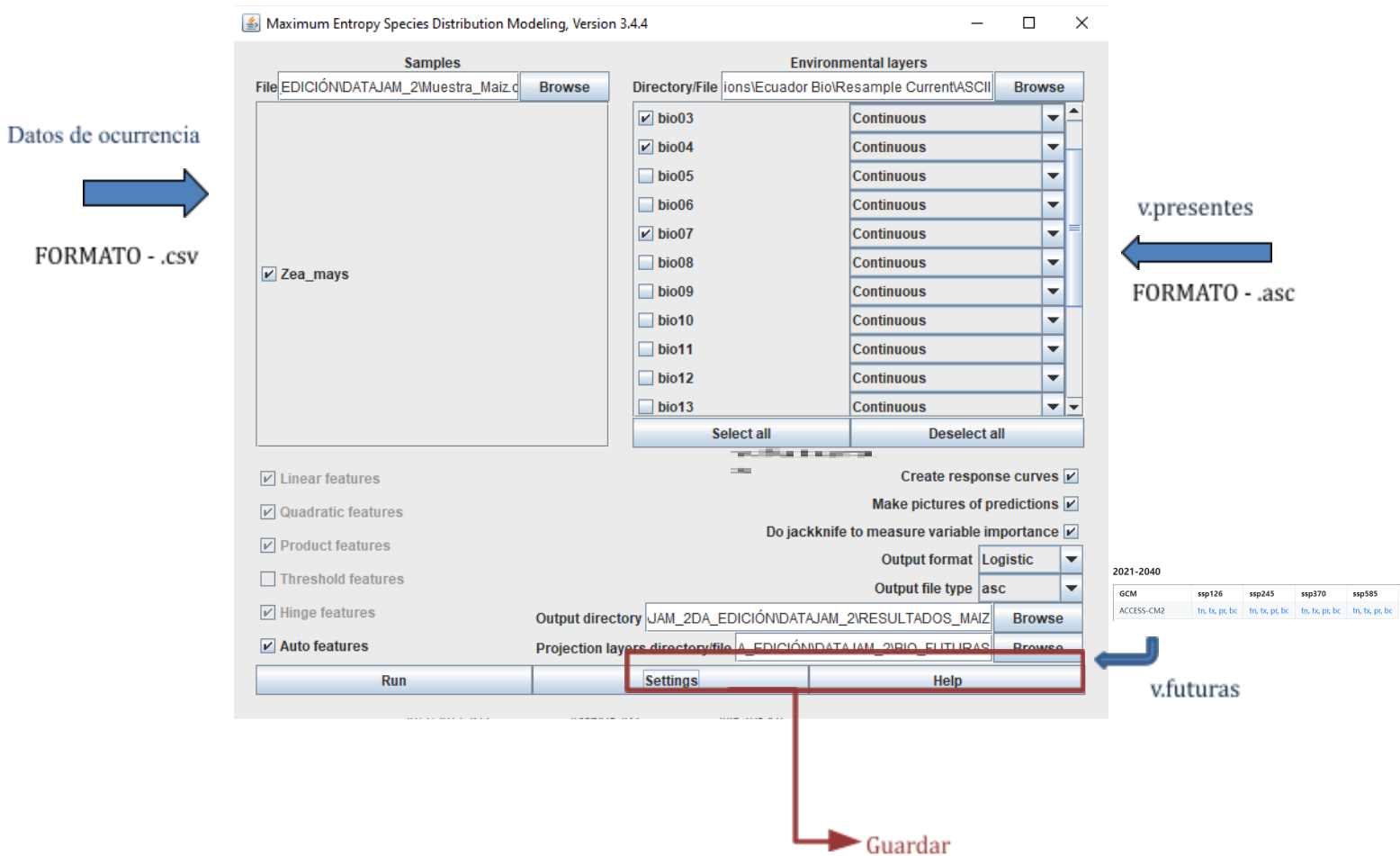
Fórmula concatenar en Excel

=CONCAT(D4;E4;F4)				
CONCAT(texto1; [texto2]; [texto3]; ...)			F	G
Especie	X	Y		
Oryza sativa	-80.0899	-3.46299		
	D4;E4;F4)			

4.2.3. Análisis de datos en MaxEnt

A continuación, se presenta la interfaz y se indica cómo debe ser completada.

Figura 19
Interfaz de MaxEnt



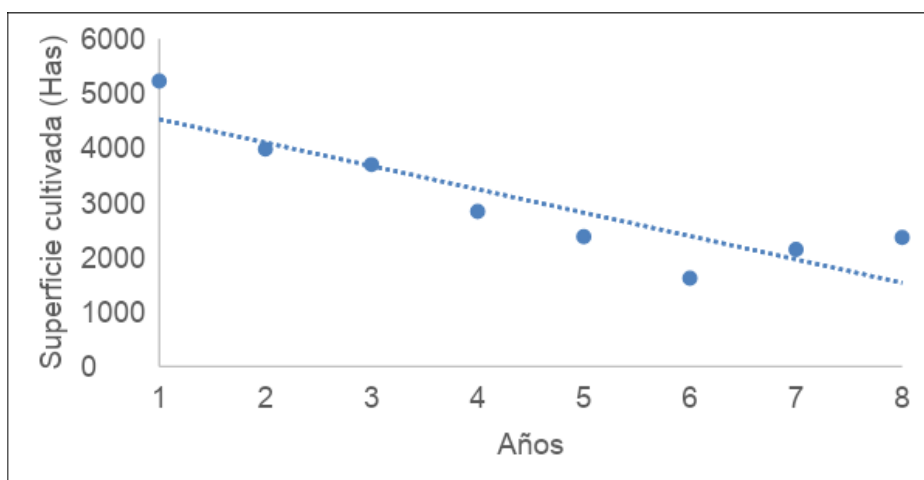
5. Resultados

5.1 Resultados de estimación de superficies cultivadas hasta 2025

A continuación, se presentan los resultados obtenidos de la estimación de superficies cultivadas de arroz y maíz duro obtenidos mediante la Regresión Lineal Simple (RLS).

Figura 20

Diagrama de dispersión de áreas cultivadas (arroz) en el cantón Simón Bolívar



De acuerdo con diagrama de dispersión el valor de r cuadrado para el cantón Simón Bolívar fue de (0.77) lo cual sugiere que el 77% de la variable dependiente como es la superficie cultivada en hectáreas de arroz es predicha por la variable independiente,

en este caso el tiempo, desde 2015 a 2022, asegurando así la fiabilidad del modelo para estimaciones a futuro.

Tabla 2

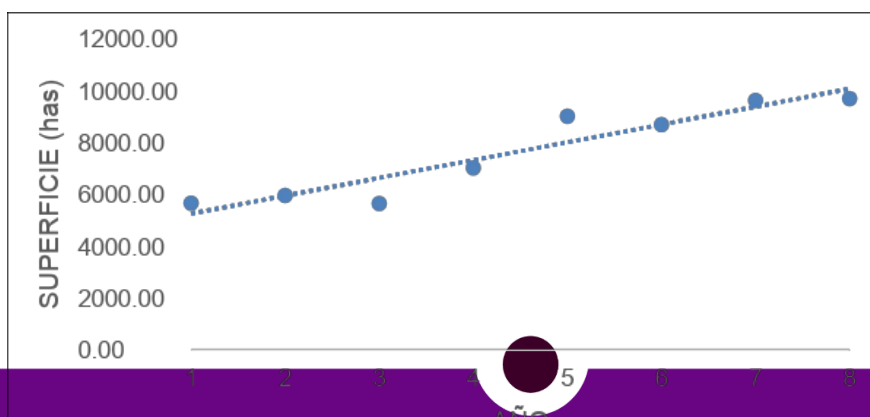
Estimación de las superficies cultivadas (arroz) hasta el año 2025 en el cantón Simón Bolívar

AÑO	SUPERFICIE (HAS)
2015	5224.5
2016	3976.9
2017	3694.1
2018	2838.2
2019	2373.1
2020	1614.0
2021	2138.9
2022	2361.7
2023	1102.5
2024	674.7
2025	246.9

De acuerdo con el diagrama de dispersión para el cantón Simón Bolívar el cultivo de arroz se observa una disminución del cultivo durante el periodo 2015 y 2025. La superficie cultivada inicialmente era de 5224.5 ha. Sin embargo, para el año 2025 esta cifra disminuyó significativamente a solo 246.9 ha. Es importante mencionar y considerar diversas razones que contribuyen esta disminución como la variación climática, la demanda del mercado e incluso la preferencia de los agricultores.

Figura 21

Diagrama de dispersión de áreas cultivadas (maíz) en el cantón Jipijapa



De acuerdo con diagrama de dispersión el valor de r cuadrado para el cantón Jipijpa fue de (0.89) lo cual sugiere que el 89% de la variable dependiente como es la superficie cultivada en hectáreas de maíz es predicha por la variable independiente, en este caso el tiempo, desde 2015 a 2022, asegurando así la fiabilidad del modelo para estimaciones a futuro.

Tabla 3

Estimación de las superficies cultivadas (maíz) hasta el año 2025 en el cantón Jipijapa

AÑO	SUPERFICIE (HAS)
2015	5664.15
2016	5959.29
2017	5654.74
2018	7019.91
2019	9016.56
2020	8695.25
2021	9627.51
2022	9696.60
2023	10757.06
2024	11443.8
2025	12130.54

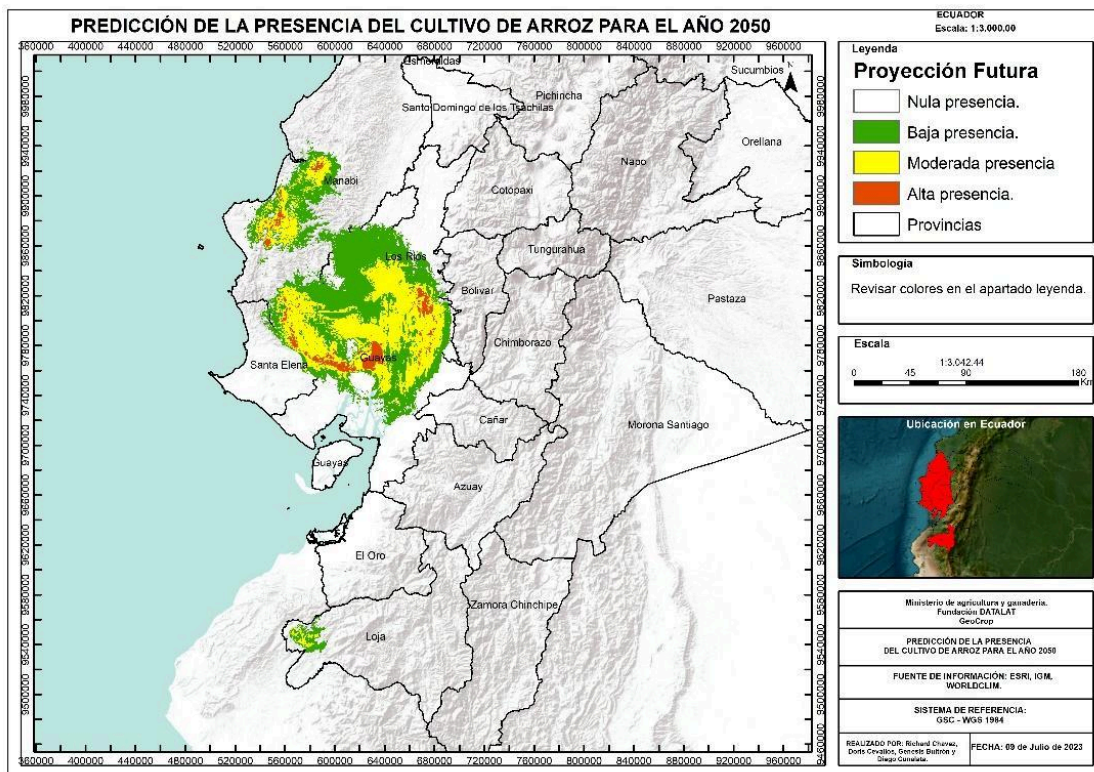
De acuerdo con el diagrama de dispersión para el año 2025 en el cantón Jipijapa el cultivo de maíz será de 12 130.54 has, lo cual indica que habrá un aumento considerable y se diversificará en el cantón.

5.2 Resultados de la estimación de superficies idóneas para los cultivos en el 2050

A continuación, se presentan los resultados obtenidos a partir de los modelos de distribución potencial mediante MaxEnt y la representación cartográfica con mapas predictivos con proyección hacia el año 2050 en Ecuador.

Figura 22

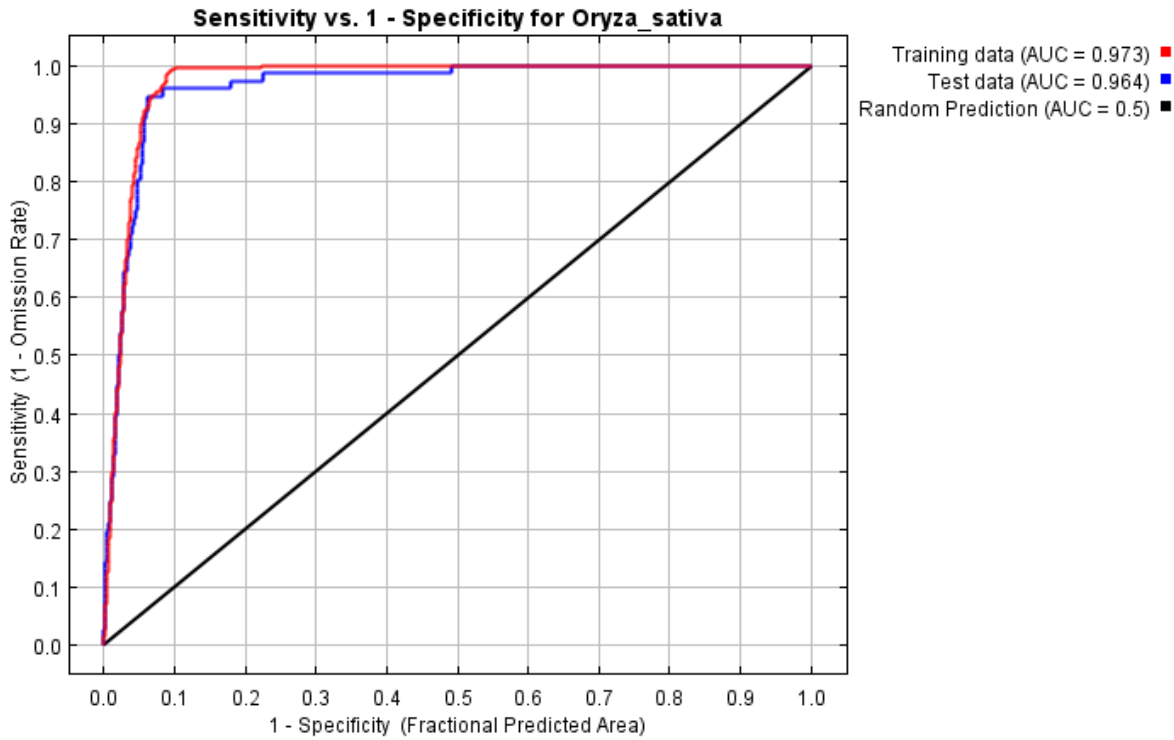
Predicción de la presencia de arroz



La distribución potencial del cultivo de arroz para el año 2050 basado en escenarios SSP5 8.5 indica que los cultivos de arroz tienen una alta y moderada probabilidad de presencia en las provincias de Los Ríos, Santa Elena, Guayas, mientras que, en las provincias de Loja y Manabí, la probabilidad de presencia va desde baja a nula, en las otras provincias no se evidencia presencia de estos cultivos debido a que el nicho ecológico generado por las condiciones bioclimáticas y geomorfológicas no es idóneo para que la especie pueda adaptarse.

Figura 23

Área bajo la curva de arroz



Los valores de AUC (ROC) del test y entrenamiento están en el rango excelente (0,9 – 1), lo que permite validar nuestro modelo como confiable (Araújo & Guisan, 2006; Mateo et al., 2011).

Figura 24

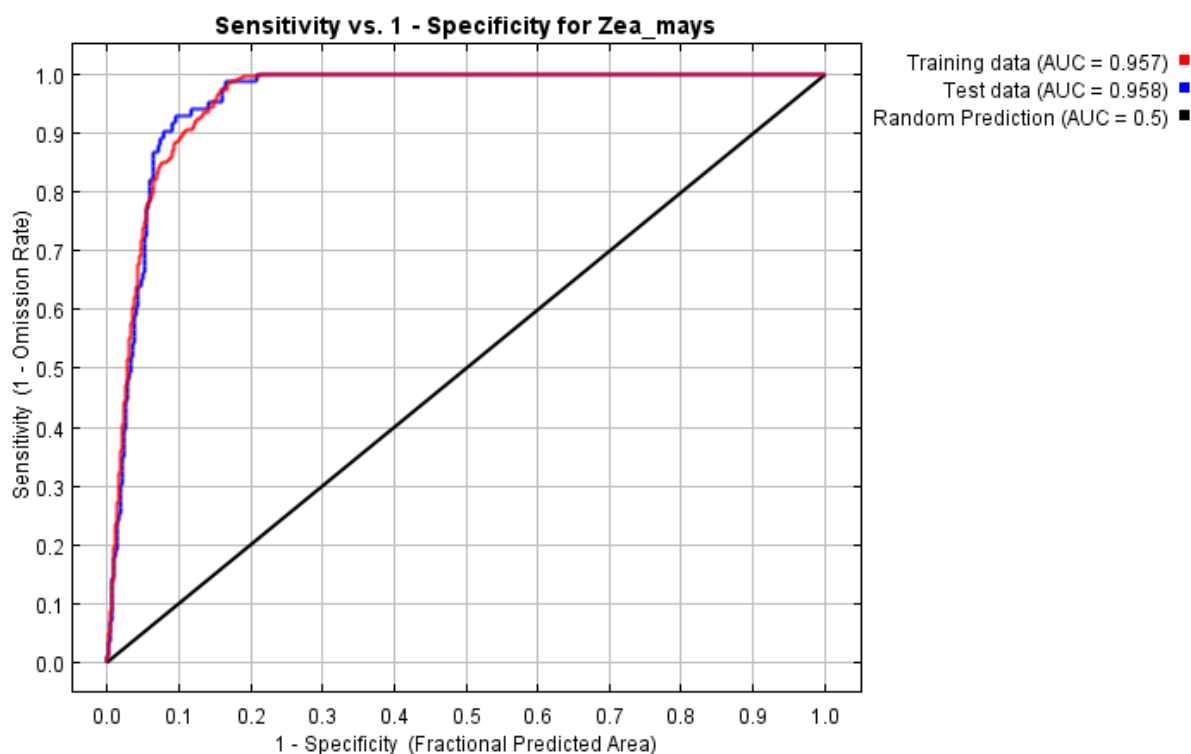
Predicción de la presencia de maíz



La distribución potencial para el año 2050 basado en escenarios SSP5 8.5 indica que los cultivos de maíz tienen una alta y moderada probabilidad de presencia en las provincias de Los Ríos y Manabí, mientras que, en las provincias de Santa Elena y Loja, la probabilidad de presencia va desde baja y moderada a nula, en las otras provincias no se evidencia presencia de estos cultivos debido a que el nicho ecológico generado por las condiciones bioclimáticas y geomorfológicas no es idóneo para que la especie pueda adaptarse.

Figura 25

Área bajo la curva de maíz



Los valores de AUC (ROC) del test y entrenamiento están en el rango excelente (0,9 - 1), lo que permite validar nuestro modelo como confiable (Araújo & Guisan, 2006; Mateo et al., 2011).

6. Recomendaciones técnicas

Cuando se trabaja con superficies cultivadas, dentro de la tabla de atributos del Software Arc Gis se debe verificar que la unidad de medida sea la misma, por ejemplo, en hectáreas que es lo más recomendable y esto se puede realizar haciendo clic derecho en la tabla de atributos y con ayuda de “calculate geometry” calcular la superficie en hectáreas.

Al obtener las variables bioclimáticas del portal worldclim se debe asegurar que la resolución sea la misma para el presente y futuro, ya que eso garantizará que el software Maxent pueda generar un mapa predictivo sin inconvenientes.

Referencias bibliográficas

Agro Bayer Ecuador. (2023). *Solución para arroz*. Obtenido de <https://www.agro.bayer.ec/es-ec/cultivos/arroz.html>

- Aguilar-Barojas, S. (2005). Fórmulas para el cálculo de la muestra en investigaciones de salud. *Salud En Tabasco*, 11(1-2), 333-338. <https://www.redalyc.org/pdf/487/48711206.pdf>
- Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33(10), 1677-1688. <https://doi.org/10.1111/j.1365-2699.2006.01584.x>
- Dagnino, J. (2014). La distribución normal. *Revista Chilena Anestesia*, 43. Escoto, A., Sánchez, L., & Gachuz, S. (2017). Trayectorias Socioeconómicas Compartidas (SSP): nuevas maneras de comprender el cambio climático y social. *Estudios demográficos y urbanos*, 32, 669-693. doi:<https://doi.org/10.24201/edu.v32i3.1684>
- ESRI. (2023). *¿Qué es ArcGIS™*. New York Street, Estados Unidos: Redlands.
- Flores, C., & Flores, K. (2021). Tests To Verify the Normality of Data in Production Processes : Anderson- Darling , Ryan-Joiner , Shapiro-Wilk and. *Societas. Revista de Ciencias Sociales y Humanísticas*, 23(2). <http://portal.amelica.org/ameli/jatsRepo/341/3412237018/index.html>
- Mateo, R. G., Felicísimo, Á. M., & Muñoz, J. (2011). Modelos de distribución de especies: Una revisión sintética. *Revista Chilena de Historia Natural*, 84(2), 217-240. <https://doi.org/10.4067/S0716-078X2011000200008>
- Mendenhall, W., Beaver, R., & Beaver, B. (2010). *Introducción a la probabilidad y estadística* (Cengage Learning). <http://latinoamerica.cengage.com>
- Minitab. (2023). *Introducción a Minitab Statistical Software*. Estados Unidos.
- Palacios-cruz, L., Pérez, M., Rivas-ruiz, R., & Talavera, J. (2013). Investigación clínica XVIII Del juicio clínico al modelo de regresión lineal. *Revista Médica Del Instituto de Mexico Seguro Social*, 51(6), 656-661. <http://www.medigraphic.com/pdfs/imss/im-2013/im136l.pdf>
- Pértegas, S., & Pita, S. (2001). *La distribución normal*.
- Phillips, S., Duvick, M., & Schapire, R. (18 de 06 de 2017). *biodiversityinformatics*. Obtenido de biodiversityinformatics: https://biodiversityinformatics.amnh.org/open_source/maxent/