



## Validación de la data de precipitación (1981-2016) del producto PISCO v2.1: Estación meteorológica, elevación, estacionalidad del año y región climática en la cuenca Tumbes, Perú

Validation of the precipitation data (1981-2016) of the PISCO v2.1 product: Meteorological station, elevation, seasonality of the year, and climatic region in the Tumbes basin, Peru

Ruben Anthony Roque Quiroz<sup>1</sup>

1. Escuela de Ingeniería Agrícola – Facultad de Ciencias Agrarias, Universidad Nacional de Tumbes, Ciudad Universitaria, Av. Universitaria S/N, Tumbes, Perú.

\* Autor correspondiente: [120164172@untumbes.edu.pe](mailto:120164172@untumbes.edu.pe) (R. A. Roque Quiroz).

ID ORCID de los autores

R. A. Roque Quiroz:  <https://orcid.org/0000-0003-3399-1325>

### RESUMEN

La escasa información que se cuenta y su importancia en los estudios hidrológicos que nos ayuda a comprender, simular, predecir los eventos asociados con el comportamiento del agua. Creo la necesidad de comprobar la data grillada PISCOpm V.2,1 precipitaciones mensuales con las observaciones en nuestra cuenca Tumbes. Para ello se realizó un estudio comparativo entre los valores observados y estimados, con respecto a estación meteorológica, elevación, estacionalidad del año y región climática. Se consideró indicadores como Coeficiente de Determinación(R), Índice de eficiencia de Nash – Sutcliffe (NS) y Sesgo Estadístico Porcentual (PBIAS%). A nivel mensual se obtuvo buenos resultados de correlación, valores mayores a 0,61, en índice de eficiencia de Nash - Sutcliffe valores superiores a 0,60 y una mínima variación en el Sesgo Estadístico Porcentual. En cuanto a elevación de la cuenca se obtuvieron datos con un mínimo ajuste en todos los indicadores. Para el arreglo estacional, verano presenta el mejor ajuste promedio, luego le sigue otoño y primavera. Invierno presenta un ajuste muy bueno para 4 de 9. Por último, de las 3 regiones climáticas, la región D(i)A' presenta el mejor ajuste en los 3 indicadores. Se pudo comprobar el ajuste de los datos teniendo resultados diversos según el criterio evaluado. Con ello se respalda el uso de los datos observados y su utilidad para estudios en la cuenca Tumbes como son inundaciones, sequías, modelamiento hidrológico e hidráulico, completación de datos faltantes, etc. En base a los resultados se puede realizar estudios de disponibilidad hídricas, análisis de eventos extremos, etc.

**Palabras clave:** Precipitación; Validación; Cuenca Tumbes; PISCOpm; Rstudio.

### ABSTRACT

The scarce information that is counted and its importance in hydrological studies helps us to understand, simulate, and predict the events associated with the behavior of water. I think it is necessary to check the gridded data PISCOpm V.2.1 monthly rainfall with the observations in our Tumbes basin. For this, a comparative study was carried out between the observed and estimated values, concerning the weather station, elevation, seasonality of the year, and climatic region. Indicators such as the Coefficient of Determination(R<sup>2</sup>), Nash-Sutcliffe Efficiency Index (NS), and Percentage Statistical Bias (PBIAS%) were considered. At the monthly level, good correlation results were obtained, with values greater than 0,61, in the Nash - Sutcliffe efficiency index values greater than 0,60 and a minimum variation in the Percentage Statistical Bias. Regarding the elevation of the basin, data were obtained with a minimum adjustment in all the indicators. For the seasonal arrangement, summer presents the best average fit, followed by fall and spring. Winter presents a very good adjustment for 4 out of 9. Finally, of the 3 climatic regions, the D(i)A' region presents the best adjustment in the 3 indicators. It was possible to verify the adjustment of the data, having different results according to the evaluated criterion. This supports the use of observed data and its usefulness for studies in the Tumbes basin such as floods, droughts, hydrological and hydraulic modeling, completion of missing data, etc. Based on the results, studies of water availability, analysis of extreme events, etc. can be carried out.

**Keywords:** Rainfall; Validation; Tumbes Basin; PISCOpm; RStudio.

Recibido: 06-02-2023.

Aceptado: 25-03-2023.



Esta obra está publicada bajo la licencia [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)

**INTRODUCCIÓN**

El monitoreo de los eventos hidrológicos presenta una limitación en la data disponible para la modelación computacional de los eventos y comprender su complejidad (Muñoz et al., 2023). Los datos son proporcionados por las estaciones meteorológicas que son de representación puntual, pero en zonas con una baja densidad de estaciones meteorológicas es necesario buscar alternativa que brinden solución a la problemática de escasa data a nivel espacial, surgen métodos de interpolación, reanálisis y validación de datos satelitales (Essou et al., 2017; C.M et al., 2022; Vaheddoost et al., 2023). Las cuencas, sub cuencas y quebradas sin aforo requieren de valores que aproximen a determinar la disponibilidad hídrica para prevenir la escasez de agua, la cual es predominante en regiones áridas y semiáridas, el planteamiento de soluciones para una correcta gestión evitaría conflictos (Marco-Dos Santos et al., 2019). También es necesario para una gestión ante el riesgo de áreas con sensibilidad a las inundaciones (Nieto-Chaupis, 2019). En el Perú se han realizado estudios para construcción de datos de lluvias grillados, de alta resolución para las cuencas hidrográficas, con ello se evalúan balances hídricos, eventos extremos, inundaciones, modelos hidrológicos con el fin de respaldar las políticas y prácticas del agua (Zubieta et al., 2015; Aybar et al., 2019; Fernandez-Palomino

et al., 2022). En Lima se realizó asociación de la variación de almacenamiento de agua entre correlación de dos productos grillado, Tropical Rainfall Measuring Mission (TRMM) y Gravity Recovery and Climate Experiment (GRACE), se obtuvo 0,8677 como mejor resultado (Mendoza et al., 2022). En Áncash se buscó mejorar la correlación entre los productos TRMM y Global Precipitation Measurement (GPM), los resultados son aceptables (Villavicencio et al., 2022). La cuenca del río Piura se realizó una validación de los datos diarios del satélite TRMM para obtener valores extremos y aplícalos a la ingeniería hidráulica (Cabrera et al., 2016). Es necesario realizar validaciones de alternativas de productos grillados de precipitación para respaldar estudios posteriores, ya que en Tumbes no se ha realizado. Es por ello que se busca validar el Producto Pisco v2.1 mensual, por estación meteorológica, elevación de la cuenca y región climática. Cabe destacar que el producto Pisco v2.1 es importante porque contempla un proceso especializado de métodos combinación de datos de estaciones terrenas, reanálisis de productos grillados, un estricto proceso de control de calidad y métodos geoestadísticos.

**MATERIAL Y MÉTODOS**

El tipo de investigación es retrospectivo y transversal para un análisis en un solo corte y de datos pasados (Dagnino, 2014; Calderón &

Alzamora de los Godos, 2018). La metodología realizada se describe en la Figura 1.

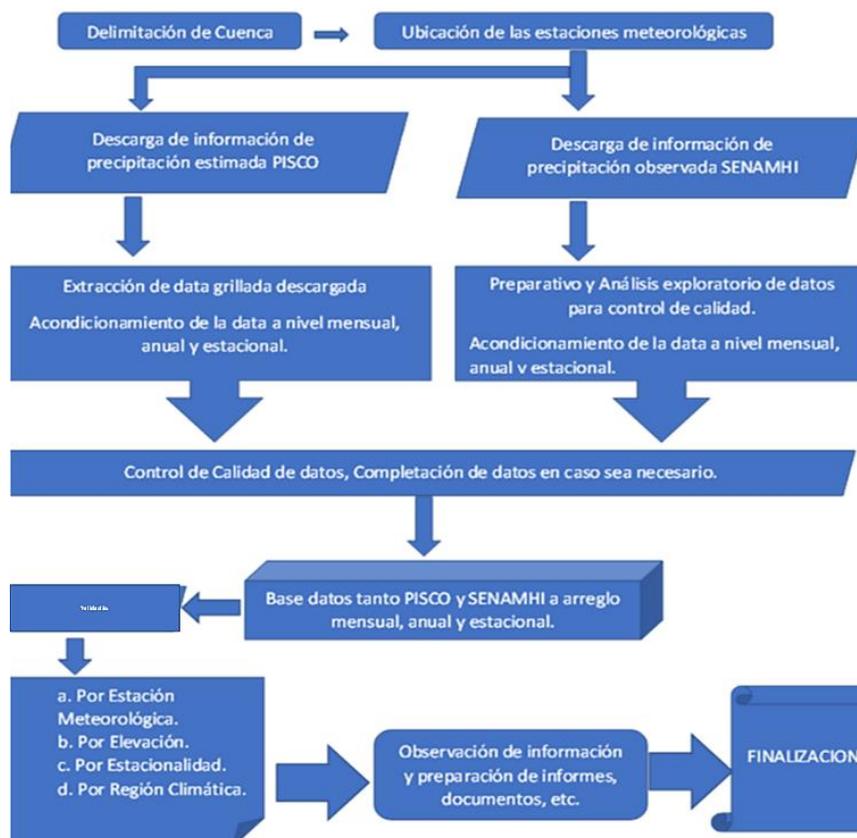


Figura 1. Diagrama del proceso metodológico.

## 1. Obtención de Datos

Los datos de precipitaciones estimadas por el producto grillado PISCO (Peruvian Interpolated data of the SENAMHI's Climatological and hydrological Observations) por sus siglas en inglés (Aybar et al., 2019). La validación fue en el mismo periodo de tiempo (36 años), el producto se encuentra alojado en la siguiente dirección web: <https://iridl.ldeo.columbia.edu/SOURCES/SENAMHI/HSR/PISCO/Prec/v2p1/> donde se puede descargar los datos estimado de precipitaciones mensuales del producto PISCO precipitaciones mensuales (PISCOpm\_v.2,1). Por ser un producto grillado se presenta en cuadrículas para su ello se utiliza software R y R Core Team (2022), en la extracción de los datos se realiza de manera puntual en la Tabla 1 se detalla las estaciones meteorológicas correspondientes.

**Tabla 1**  
Estaciones meteorológicas cuenca Tumbes

N	Nombre	Latitud	Longitud	Altitud
1	Puerto Pizarro	3°30'13.67"	80°23'42.14"	6
2	Los Cedros	3° 38' 1"	80° 32' 1"	74
3	Campamento Sede	3°33'4.14"	80°25'46.64"	17
4	Quebrada	3° 46' 1"	80° 24' 1"	89
5	El Tigre	3°48'8.4"	80°27'24"	45
6	Rica Playa	3°48'42.84"	80°30'19.65"	68
7	Cabo Inga	3°58'33.39"	80°24'6.54"	231
8	Capitán Hoyle	4° 4' 1"	80° 29' 1"	235
9	Condor Flores	3° 53'	80° 8' 22.73"	384

Fuente: adaptado de SENAMHI.

## 2. Preparativos y control de calidad de datos de precipitación

### 2.1. Análisis exploratorio data

Para el análisis se realizó sobre data con ordenamiento multianual y mensual para cada una de las estaciones meteorológica, se realizó una exploración visual, tendencia y saltos (Kundzewicz & Robson, 2000).

En la identificación de datos atípicos se empleó el método de Turkey, para la completación de datos se empleó el método de software HEC- 4 Monthly Streamflow Simulation, el cual cuenta con un respaldoado en variedad de estudios hidrológicos.

## 3. Validación

La validación de las datas de precipitación mensuales del producto PISCO v.2,1 y la data de las estaciones meteorológicas del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI). se empleó métricas estadísticas y la correlación entre ambas data.

### 3.1. Validación por estación meteorológica

El producto PISCO tiene una resolución espacial de 0,1°x 0,1° (10km2x10km2), la información puntual de las estaciones meteorológicas y las observadas por SENAMHI en series de tiempo desde enero de 1981 hasta diciembre de 2016 (dependiendo de la información disponible). La validación se realiza entre las series de tiempos con las medidas de comparación estadísticas como: Coeficiente de determinación( $r^2$ ), error absoluto medio (MAE), raíz error medio cuadrático (RMSE), índice de

eficiencia de nash-sutcliffe (NSE), sesgo estadístico porcentual (PBIAS).

### 3.2. Validación por elevación

Se adecua las series de precipitación en Precipitación Promedio Multianual (PPM), esto para la data estimada y observada. También se considera su altitud de cada una de ella, para ello se debe ordenar en columnas con los siguientes titulados: Estación, Altitud y Precipitación Promedio Multianual Observada (mm) y Precipitación Promedio Multianual Estimada PISCO (mm).

La correlación se realiza entre la altitud de las estaciones y su Precipitación Promedio Multianual (mm) observada, de igual manera la Precipitación Promedio Multianual Estimada PISCO (mm) se emplea el coeficiente de determinación, esto nos indicará el comportamiento y tendencia de los datos de precipitación y su altitud, el ajuste entre observadas y estimadas.

### 3.2. Validación por estacionalidad

Las estaciones del año están dadas por el movimiento de traslación de nuestro planeta, tiene una duración de 3 meses, para el Perú el otoño comienza en marzo, primavera inicia en septiembre, invierno en junio y verano en diciembre (Gob.pe, 2022). Teniendo en consideración lo anterior se genera 4 estaciones de acuerdo con sus meses se agrupará los datos de precipitación mensuales así:

Otoño (marzo-abril-mayo), invierno (junio-Julio-agosto), primavera (septiembre-octubre-noviembre) y verano (diciembre-enero-febrero). En la serie de datos se verá reflejado en las precipitaciones totales mensuales (mm), iniciando desde marzo 1981, abril 1981, mayo 1981 que es otoño de igual manera se agrupará en función del mes y la estación que corresponde hasta llegar al año de 2016, esto se realizará tanto para precipitación estimada y observada. se realizará la validación con los estadísticos de comparación entre las series estacionales de precipitaciones estimadas y observadas.

### 3.3. Validación por región climática

Para la validación se debe tener el mapa de regiones climáticas del Perú que se encuentra en <https://www.senamhi.gob.pe/?p=mapa-climatico-del-peru> el cual fue elaborado por el servicio nacional de meteorología e hidrología del Perú (SENAMHI), con ello se identifica la regiones climáticas teniendo en cuenta la precipitación anual de cada una de las estaciones con la data observada o estimada de precipitación, para ello se realiza series de tiempos de precipitación anual promediadas desde enero de 1981 a diciembre de 2016. luego las series de tiempo según a la región climática que corresponda se realiza la validación entre la data observada y estimada de acuerdo con los estadísticos comparativos.

### 3.4. Medidas estadísticas de comparación.

Para la validación del comportamiento de los datos de precipitación observada y estimada, aplican estadísticos comparativos en la Tabla 2 se detalla

cada una de ellas como: coeficiente de determinación( $r^2$ ), error absoluto medio (MAE), raíz error medio cuadrático (RMSE), índice de eficiencia de nash-sutcliffe (NSE), sesgo estadístico

porcentual (PBIAS). Los resultados de los estadísticos comparativos permiten clasificar y analizar su comportamiento de acuerdo con algunos rangos establecidos.

**Tabla 2**

Métricas estadísticas y sus correspondientes ecuaciones utilizadas para evaluar el desempeño meteorológico

Métrica Estadística	Unidad	Ecuación	Valor óptimo
Error Absoluto Medio	mm/h	$\frac{1}{n} \sum_{i=1}^n  Ei - Oi $	0
Raíz del Error Medio Cuadrático	mm/h	$\sqrt{\frac{1}{n} \sum_{i=1}^n  Ei - Oi ^2}$	0
Coefficiente de Determinación	-----	$r_s^2 = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$	$\pm 1$
Índice de Eficiencia de Nash-Sutcliffe	-----	$1 - \frac{\sum_{i=1}^n  Ei - Oi ^2}{\sum_{i=1}^n  Oi - \bar{O} ^2}$	1
Sesgo Estadístico Porcentual	-----	$100 * \frac{\sum_{i=1}^n Ei - Oi}{\sum_{i=1}^n Oi}$	$< \pm 10$

Nota: n número de muestras, **Ei** Datos Pisco, **Oi** Datos Observados.

**RESULTADOS Y DISCUSIÓN**

**Análisis exploratorio data**

Se realizó el análisis de datos atípicos con el método de box-plot, con ello se pudo determinar los datos que no se ajustan al comportamiento normal de los conjuntos de datos (Figura 2). Se encontraron en promedio 3 valores atípicos, los cuales se asocian a los eventos del Niño costero en los años 1983, 1992, 1997, 1998. Consultando y comparando los datos con estaciones vecinas se decidió conservar esos datos, puesto que eliminarlos sería eliminar los rasgos históricos y negar su existencia Para el caso de las variables de precipitación del producto PISCO no necesita mayor análisis debido a que esta ya tiene una compleja validación y análisis a profundidad.

Para el análisis de consistencia de los datos se realizó con la curva de doble masa, para el análisis estadístico se realizó consistencia en la media y en la desviación estándar. Todo esto se llevó a cabo con el software Hydrognomon 4 y Trend.

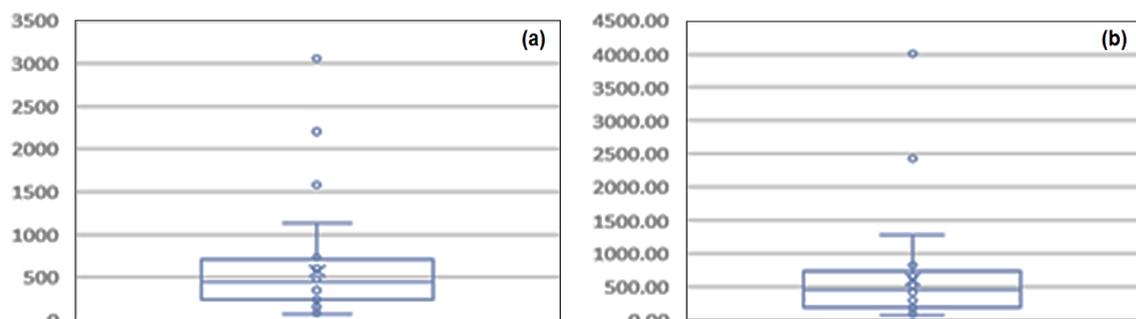
La completación de datos se hizo con Hec- 4 Monthly Streamflow Simulation el cual está basado en un lenguaje fortran. Es sencillo y ofrece buenos resultados, además tienen estudios y experiencia que lo respaldan.

**Validación por estación meteorológica**

El análisis puntual de los datos considerando la variabilidad espacial del producto PISCO se realizó considerando las estaciones meteorológicas y su disponibilidad de data.

Para ello se consideró las métricas estadísticas, se comenzó realizando la lectura de los datos observados y simulados en Rstudio para realizar la comparación con el package “hidroGOF” (Mauricio Zambrano-Bigiarini, 2020), quien cuenta con variedades de métricas. Se observó que en la métrica de Coeficiente de determinación ( $R^2$ ) que las estaciones como Campamento Sede, Tigre y Puerto Pizarro presentaron los máximos valores 0,89, 0,84 y 0,85 respectivamente considerado una correlación Muy Alta según (Mamani, 2019). En contrario la estación Condor Flores tiene el mínimo valor de 0,12 y las estaciones Cedros, Quebrada Angostura y Rica Playa una correlación Alta con valores de 0,61, 0,62 y 0,73. Cabo inga y Capitán Hoyle presentan valores de 0,46 y 0,54 con una clasificación Media.

Entonces se puede inferir a partir de los valores obtenidos que el 66,67% de las estaciones presentan un ajuste bueno para tener en cuenta.



**Figura 2.** a) Precipitaciones Observadas y b) Precipitaciones Estimadas. Estación El Tigre (Total).

Otro análisis de importancia es el Índice de Eficiencia de Nash- Sutcliffe (NSE), nos indica que tan preciso es el modelo generado, se encontraron solo dos estaciones que presenta una precisión Insuficiente que son Condor Flores y Capitán Hoyle, luego la estación Cabo Inga con 0,32 tiene una precisión satisfactoria. Una excelente precisión presenta las estaciones como Campamento Sede y Puerto Pizarro con valores superiores 0,80.

En el análisis el Sesgo Estadístico Porcentual (PBIAS), representa la sobreestimación en caso de valores positivos y con valores contrarios una subestimación (Moriassi, Arnold, Van Liew, & Harmel, 2007). Se tiene un 55,56% de las estaciones totales con sobreestimación, las cuales van desde valores de 5,4 a 49,5 y son Quebrada Angostura, Capitán Hoyle respectivamente.

Con subestimación se tiene los valores extremos de -5,5 a -30,8 que corresponde a Cabo Inga y Rica Playa. Así como se resumen los valores en la Tabla 3.

**Tabla 3**  
Métricas estadísticas validación por estación meteorológica

Estación	R <sup>2</sup>	MAE	RMSE	NSE	PBIAS %
Puerto Pizarro	0,46	38,9	138,89	0,32	-5,5
Los Cedros	0,89	11,52	37,4	0,88	-13,9
Campamento Sede	0,54	41,79	126,51	-0,19	49,5
Quebrada Angostura	0,61	14,81	57,62	0,6	8,6
El Tigre	0,12	78,95	256,36	-0,09	-26
Rica Playa	0,85	12,35	39,42	0,84	9,8
Cabo Inga	0,62	20,46	62,32	0,53	5,4
Capitán Hoyle	0,73	22,77	68,34	0,7	-30,8
Condor Flores	0,84	14,35	51,79	0,75	6,4

### Validación por elevación

Se realizó un clúster por elevación de las estaciones para la comprobación de los datos, como resultado se obtuvo tres clasificaciones, Cuenca Alta (Condor Flores, Cabo Inga, Capitán Hoyle), Cuenca Media (Tigre, Rica Playa, Quebrada Angostura) y Cuenca Baja (Campamento Sede, Cedro, Puerto Pizarro). Para ello se consideró los promedios de las estaciones por cada clasificación.

De acuerdo a la Tabla 4 se encontró los siguientes resultados:

Para el análisis de Coeficiente de Determinación (R<sup>2</sup>), se encontró los mejores resultados en la Cuenca Alta con un valor de 0,57. El menor valor obtuvo la Cuenca Media que es de 0,22, al igual con un valor cercano la Cuenca Baja con 0,27.

Con el Índice de Eficiencia de Nash- Sutcliffe (NSE), que determina que tan preciso es el modelo.

La Cuenca Alta, Cuenca Media y Cuenca Baja se obtuvo resultados insuficientes. Con valores de 0,06, -0,45 y -9,27 respectivamente.

La variación en la precisión de los valores del modelo se puede dar por sobreestimación o subestimación, para ello se empleó el Sesgo Estadístico Porcentual (PBIAS), las Cuenca Baja tiene una sobreestimación con un valor de 7,1. En contrario, las Cuenca Media y Alta tienen una subestimación con valores de -11,9 y -25,9 respectivamente.

Se puede inferir que, en función de la elevación y el agrupamiento de estas estaciones, en la parte baja

presenta una baja relación lineal de los datos, insuficiente precisión y una sobreestimación. En la parte media y alta presenta una subestimación.

**Tabla 4**  
Métricas estadísticas validación por elevación

Elevación	R <sup>2</sup>	MAE	RMSE	NSE	PBIAS %
CUENCA BAJA	0,27	31,29	37,4	-9,27	7,1
CUENCA MEDIA	0,22	91,98	116,62	-0,45	-11,9
CUENCA ALTA	0,57	295,16	461,74	0,06	-25,9

### Validación por estacionalidad

Se consideró el agrupamiento de los datos de precipitación en arreglo de acuerdo los meses de las estaciones de año, para cada estación tanto observada y simulada.

En verano que corresponde a los meses de diciembre-enero-febrero. En la Tabla 5 se tiene el resumen de los resultados obtenidos, pero se detalla que en la métrica del Coeficiente de Determinación (R<sup>2</sup>), las estaciones como Rica Playa con 0,96 alcanzo el máximo valor de ajuste de los datos, siguiéndole Campamento Sede, Puerto Pizarro y Tigre. En contra parte, la estación Condor Flores con 0,26 tiene la peor varianza o ajuste de los datos.

En términos de Índice de Eficiencia de Nash- Sutcliffe (NSE), las estaciones Campamento Sede y Puerto Pizarro presentan el máximo ajuste con 0,9, seguida por Tigre, Rica Playa y Cedros. Las demás estaciones presentan un ajuste de satisfactorio a insuficiente, teniendo en cuenta a la estación Quebrada Angostura con el peor ajuste.

Para términos del Sesgo Estadístico Porcentual (PBIAS), se tiene más subestimación en las estaciones, la estación Rica Playa presenta -42,7 y Puerto Pizarro con -3,3. La sobreestimación se da en las estaciones Capitán Hoyle (39,6), Tigre (9,1) y Quebrada Angostura (3,5).

**Tabla 5**  
Métricas estadísticas validación por estacionalidad (verano)

Estación	R <sup>2</sup>	MAE	RMSE	NSE	PBIAS %
Cabo Inga	0,32	76,68	223,49	0,24	-10,9
Campamento Sede	0,9	18	45,98	0,9	-11,7
Capitán Hoyle	0,43	64,1	156,56	-0,08	39,6
Cedros	0,67	26,84	83,95	0,67	-4,9
Condor Flores	0,26	95,74	156,33	-0,16	-21,8
Puerto Pizarro	0,91	18,62	42,3	0,9	-3,3
Quebrada Angostura	0,37	44,15	99,98	-0,24	3,5
Rica Playa	0,96	46,01	92,07	0,74	-42,7
Tigre	0,89	23,79	56,35	0,82	9,1

El invierno que se desarrolla entre los meses de junio, julio y agosto. Como se muestra en la Tabla 6 se presenta un muy alto ajuste de los datos en 3 estaciones de 9 las cuales son Campamento Sede, Puerto Pizarro y Rica Playa con valores superiores al 0,90, La estación Tigre con 0,61 tiene un ajuste alto y el resto de estaciones que son Cabo Inga, Capitán Hoyle, Cedros, Quebrada Angostura y Condor flores tienen un valor de 0 con respecto a esto (Mamani,2019) lo clasifica como mínima.

Con respecto al Índice de Eficiencia de Nash- Sutcliffe (NSE), se presenta un desempeño Bueno

en las estaciones Puerto Pizarro, Rica Playa y Tigre con un rango de 0,41 – 0,53. En las demás estaciones se evidenció un desempeño insuficiente en la clasificación de (Molnar, 2011).

En desempeño de Sesgo Estadístico Porcentual (PBIAS), las estaciones como Condor Flores, Cabo Inga presentan los menores valores, pero igual junto con las demás estaciones tienen un desempeño insatisfactorio. Pero en cuestiones de sobreestimación tenemos a 4 de 9 estaciones que son: Capitán Hoyle, Cedros, Quebrada Angostura y Condor Flores. Los otros 5 restantes presentan subestimación.

**Tabla 6**  
Métricas estadísticas validación por estacionalidad (invierno)

Estación	R <sup>2</sup>	MAE	RMSE	NSE	PBIAS %
Cabo Inga	0	7,23	34,38	-1,36	-32,9
Campamento Sede	0,98	5,4	37,77	0,29	-73,9
Capitán Hoyle	0	5,1	30,54	-1148,59	2755,5
Cedros	0	1,43	5,35	-10,75	74,1
Condor Flores	0	5,66	18,33	-2,48	10,5
Puerto Pizarro	0,98	3,95	25,21	0,47	-58,5
Quebrada Angostura	0	4,19	21,46	-431,16	923
Rica Playa	0,94	3,78	21,28	0,51	-71,3
Tigre	0,61	5,34	30,12	0,53	-58,1

Los meses de marzo, abril y mayo, periodo donde se da la estación de otoño, las precipitaciones al inicio se presentan en periodo lluvioso y luego se llega al periodo seco. Teniendo en cuenta a la Tabla 7 con respecto a las métricas se detalla que el Coeficiente de Determinación (R<sup>2</sup>), las estaciones Campamento Sede, Puerto Pizarro presenta el mejor ajuste teniendo una clasificación de 0,92 y 0,91 que según (Mamani, 2019) es muy alta. Luego la estación Condor Flores con 0,05 presenta un ajuste con clasificación mínima y Cabo Inga tiene una clasificación de media junto a las demás estaciones.

Con el Sesgo Estadístico Porcentual (PBIAS), se pudo constatar que la estación Campamento Sede tiene una subestimación con una clasificación de buena, como la estación Tigre con una sobreestimación, En subestimación las estaciones tienen un comportamiento insatisfactorio con valores promedio a 25 y cabe mencionar a la estación Capitán Hoyle que presenta el máximo valores de sobreestimación con desempeño insatisfactorio según (Moriasi, Arnold, Van Liew, & Harmel, 2007) con un valor de 62,5.

**Tabla 7**  
Métricas estadísticas validación por estacionalidad (otoño)

Estación	R <sup>2</sup>	MAE	RMSE	NSE	PBIAS %
Cabo Inga	0,43	72,93	187,11	0,9	4,3
Campamento Sede	0,92	20,95	44,74	0,88	-12,2
Capitán Hoyle	0,57	91,41	195,05	-0,64	62,5
Cedros	0,51	29,53	78,72	0,32	25,8
Condor Flores	0,05	178,73	472,11	-0,27	-24,9
Puerto Pizarro	0,91	25,83	61,51	0,75	28,3
Quebrada Angostura	0,74	44,32	88,77	0,71	-29,8
Rica Playa	0,54	39,55	98,55	0,46	-14,7
Tigre	0,78	25,65	80,87	0,61	11,5

La estimación con la métrica de Índice de Eficiencia de Nash- Sutcliffe (NSE), la variación del ajuste en

clasificación excelente con valor de 0,9 y 0,88 son las estaciones Cabo Inga y Campamento Sede. Condor Flores presenta un ajuste insuficiente teniendo en cuenta que cuando los valores son negativos se entiende que la media de los valores observados tiene mejor ajuste que los valores estimados.

De acuerdo la Tabla 8 que corresponde a la estación primavera considerando los meses de setiembre, octubre y noviembre. Se tiene indicando que el Coeficiente de Determinación (R<sup>2</sup>), presenta un excelente ajuste en estaciones como Campamento Sede con 0,96. De igual manera Puerto Pizarro con 0,98. También se tiene las estaciones Cedros y Rica Playa. Existen una proporción de variación mayor que genera un ajuste mínimo en las estaciones de Condor Flores, Quebrada Angostura.

Para el Índice de Eficiencia de Nash- Sutcliffe (NSE), se evidencia un buen desempeño en las estaciones como Campamento Sede, Cedros, Puerto Pizarro con valores desde 0,88 a 0,97 con una clasificación excelente bajo la clasificación de (Molnar, 2011) también se tiene estaciones que tienen un desempeño insuficiente o de pésimo ajuste como son las estaciones Capitán Hoyle, Condor Flores y Quebrada Angostura.

También en la Tabla 8 se encuentran valores para Sesgo Estadístico Porcentual (PBIAS), el desempeño en esta métrica para la estación Puerto Pizarro es el mejor teniendo una mínima sobreestimación con el valor de 4,6. Para las demás estaciones se tiene subestimaciones y sobreestimaciones con un desempeño insatisfactorio.

**Tabla 8**  
Métricas estadísticas validación por estacionalidad (primavera)

Estación	R <sup>2</sup>	MAE	RMSE	NSE	PBIAS %
Cabo Inga	0,77	2,78	9,48	0,6	-41,8
Campamento Sede	0,96	1,74	7,3	0,88	27,7
Capitán Hoyle	0,25	6,54	23,07	0,02	-75,8
Cedros	0,93	1,45	2,8	0,91	-33
Condor Flores	0,07	35,66	123,37	-0,13	-54,1
Puerto Pizarro	0,98	1,01	2,62	0,97	4,6
Quebrada Angostura	0,01	6,04	23,85	-6,02	57,6
Rica Playa	0,94	1,75	6,47	0,68	-60,8
Tigre	0,42	2,62	10,28	0,35	-36,1

### Validación por región climática

Para la zona de estudio se encontraron 3 regiones climáticas con estaciones que nos brindan los datos para realizar la comparación de ajuste entre los valores observados y simulados.

En la Tabla 9 se puede observar que la región climática C(i,p)A' se define como una región de precipitación efectiva semiseco cálido y con invierno, primavera seco que se encuentra a altitud mayor a 600 m.s.n.m. En Tumbes se encuentra en áreas de los cerros de Amotape con temperaturas máximas de 21 °C a 31° C y con lluvias anuales de 700 y 1200 mm. Para el primer análisis de desempeño en la métrica de Coeficiente de Determinación (R<sup>2</sup>), se encuentra un valor de 0,41 considerado un ajuste bueno y un valor de 146,82 para Raíz del Error Medio Cuadrático (RMSE).

Según Índice de Eficiencia de Nash- Sutcliffe (NSE), se tiene un valor de 0,23 se tiene un ajuste

satisfactorio. Con el Sesgo Estadístico Porcentual (PBIAS) se determinó que existe un muy buen ajuste con una mínima subestimación teniendo un valor de -4,2.

La región climática D(i)A' se define como semiárido con invierno seco cálido (Tabla 9). Este se desarrolla en altitudes mayores a los 300 m.s.n.m. en los Cerros de Amotape de Tumbes, caracteriza por tener precipitaciones anuales entre 700 a 1200 mm y temperatura de 17 °C a 29 °C.

Se obtuvo Coeficiente de Determinación ( $R^2$ ) de 0,12 (muy bajo) y una Raíz del Error Medio Cuadrático (RMSE) de 256,36. Para el Índice de Eficiencia de Nash- Sutcliffe (NSE) el modelo presenta un ajuste insatisfactorio con un valor de -0,09 considerando que la media de los observados tiene mejor ajuste a los valores estimados y como final tenemos el Sesgo Estadístico Porcentual (PBIAS) con el cual se encontró una subestimación insatisfactoria con un valor de -26.

En la región climática D(i,p)A' un clima semiárido, cálido con invierno y primavera seco es el clima que se encuentra en gran ocupación en el área de estudio presentando temperaturas de 19 °C a 33 °C y precipitaciones anuales de 200 a 500 mm. Además, se da en altitudes menores a los 300 m.s.n.m. Teniendo en cuenta la métrica de Coeficiente de Determinación ( $R^2$ ) se tiene un ajuste muy alto con un valor 0,92 que nos indica que la variación es la mínima, para la Raíz del Error Medio Cuadrático (RMSE) se tiene un valor de 28,94 un error bajo. Tomado en cuenta el valor de 0,91 para el Índice de Eficiencia de Nash- Sutcliffe (NSE) se tiene una clasificación de excelente Estadístico Porcentual (PBIAS) se una mínima subestimación con una clasificación de muy buena con un valor de -3,6, esta estación presenta un ajuste muy bueno en las métricas.

**Tabla 9**

Métricas estadísticas validación por región climática

Región Climática	$R^2$	MAE	RMSE	NSE	PBIAS %
C(i,p)A'	0,41	39,91	146,82	0,23	-4,2
D(i)A'	0,12	78,95	256,36	-0,09	-26
D(i,p)A'	0,92	8,56	28,94	0,91	-3,6

Para Mamani (2019) que realizó una regionalización de precipitaciones mensuales en la Cuenca de la Vertiente del Lago Titicaca – Lado peruano, se hizo validación de los datos del Producto PISCO teniendo como indicadores el Coeficiente de Determinación ( $R^2$ ), para seis estaciones tuvo cuatro estaciones que presentaba un muy buen ajuste con valores superiores a 0,90 y el resto con valores superiores a 0,70, En contraste, esta validación se asemeja a la comprobación en punto grilla que para nuestras nueve estaciones tenemos un desempeño variable, siendo las estaciones Campamento Sede, Puerto Pizarro y Tigre las que se acercan a valor de 0,90.

El Sesgo Estadístico Porcentual (PBIAS) tuvo un comportamiento de 3 estaciones sobreestimadas y 3 subestimadas con el mínimo valor de 5,236 y máximo de -27,36. En nuestro caso, en este indicador tenemos una un desempeño similar teniendo 4 estaciones subestimadas y 5 sobre-

estimadas con valores extremos de 5,4 y 49,5 en las estaciones Quebrada Angostura y Capitán Hoyle.

Mamani (2019) Obtuvo valores para Índice de Eficiencia de Nash- Sutcliffe (NSE) de bueno con valores superiores a 0,90 y una sola estación de 0,875. En nuestro estudio obtuvimos desempeños diversos con mayor ajuste de 0,88 y peor ajuste de -0,19 para las estaciones de Campamento Sede y Capitán Hoyle.

Zapana (2019) en su estudio intitulado Análisis de la sequía meteorológica de la Intercuenca Alto Apurímac y las cuencas Alto Urubamba y Yavero utilizando el método de L-Momentos, se realizó la validación del Producto PISCO para 24 estaciones en punto grillado, utilizo como métricas el Coeficiente de Determinación ( $R^2$ ) el cual, ha presentado resultado aceptable para su estudio, teniendo como valores extremos 0,543 y 0,946. Con respecto a nuestros resultados tienen mejor desempeño promedio sin considerar a las estaciones Condor Flores con 0,12 y Cabo Inga con 0,46.

Porto (2021) que realizó la validación de las estaciones de la Cuenca de la Vertiente del Lago Titicaca – Lado peruano, considero un total de 34 estaciones con el método del vector regional los agrupo en 5 regiones. Como indicadores utilizo el Coeficiente de Determinación ( $R^2$ ), obteniendo un desempeño muy bueno con valores superiores a 0,90, Para nuestra zona de estudio obtuvimos valores menores, pero con desempeño bueno.

Considerando el Índice de Eficiencia de Nash- Sutcliffe (NSE) se obtuvo un desempeño excelente con valores frecuentes de 0,90, En comparación con los resultados obtenidos en nuestro estudio que son de desempeño diverso teniendo un valore destacable de 0,89.

También considero el Sesgo Estadístico Porcentual (PBIAS) donde obtuvo un diverso desempeño en las estaciones, la región I y región V las estaciones en su totalidad es subestimada. Se encontró valores extremos de subestimación -27,094 y -0,399. Para el desempeño de sobreestimación se tiene 22,119 y 0,837. En contraste con los valores obtenidos, estos presentan mayor sobreestimación y subestimación con valores superiores.

También realizó estudio a nivel estacional obteniendo buenos resultado con algunas estaciones que tiene un desempeño regular, para los meses de marzo, abril y mayo se encontró valores de  $R^2 = 0,392$ . En los meses de junio, julio y agosto un  $R^2 = 0,293$ . En setiembre, octubre y noviembre un  $R^2 = 0,356$  y por último para diciembre, enero y febrero un  $R^2 = 0,273$ . Esto en comparación para los valores obtenidos en nuestra zona de estudio en marzo, abril y junio ( $R^2 = 0,05$ ). Junio, julio y agosto ( $R^2 = 0$ ). En setiembre, octubre y noviembre ( $R^2 = 0,01$ ) y para diciembre, enero y febrero ( $R^2 = 0,26$ ). Lo cual desestima el uso de esas estaciones para dichos periodos.

También se realizó una validación según la región climática donde encontró un desempeño excelente en todas las métricas con una mínima subestimación, en nuestra área de estudio solo una región climática llega a tener un desempeño similar.

## CONCLUSIONES

Se realizó el análisis exploratorio y consistencia de los datos de precipitación. A nivel mensual, en las 9 estaciones que se validó su ajuste en su correlación se obtuvo valores altos en un intervalo de 0,61 a 0,89 donde se puede optar por la utilización de los datos estimados, también se debe considerar el índice de Nash que se obtuvo en un intervalo de 0,60 a 0,88 y para la consideración de que valores de precipitación están sobreestimados o subestimados en un ajuste bueno se tiene un valor más o menos 15 hasta la convergencia a 0, De acuerdo a lo anterior se debe destacar a estaciones como Campamento Sede, Puerto Pizarro, Tigre, Rica Playa y Cedros.

Para la validación en arreglo estacional verano tiene la mayor correlación en las estaciones de acuerdo a las clasificaciones, teniendo para  $R = 4$  estaciones con muy alta, para  $NSE = 3$  estaciones con clasificación excelente y 4 estaciones con el mínimo PBIAS. La siguen las estaciones como otoño y primavera con similar ajuste en las estaciones meteorológicas. En cambio, invierno presenta el peor desempeño tanto en correlación, eficiencia y sesgo.

Se pudo identificar las regiones climáticas donde se ubicaban las estaciones. Se generó un promedio multianual por cada región.

La región climática D (i, p) A' presenta el mejor ajuste tanto en correlación con un valor de  $R=0,92$ . Un  $NSE= 0,91$  y tiene una mínima subestimación en los datos con un  $-3,6 \%$ . Por lo tanto, esa región climática evidencia un desempeño excelente con el Producto PISCO. No como las otras regiones climáticas como lo son C (i, p) A' con un desempeño bueno y D(i)A' con un desempeño mínimo.

En cuanto a elevación se obtuvo correlación media a mínima, teniendo un mayor ajuste a mayor elevación, lo cual se da en la Cuenca Alta, para el índice de eficiencia de Nash-Sutcliffe el desempeño es insatisfactorio, de igual manera para el Sesgo Estadístico Porcentual. En la Cuenca Baja se evidencia una sobreestimación y en Cuenca Medía se tiene un desempeño mínimo en general con una subestimación.

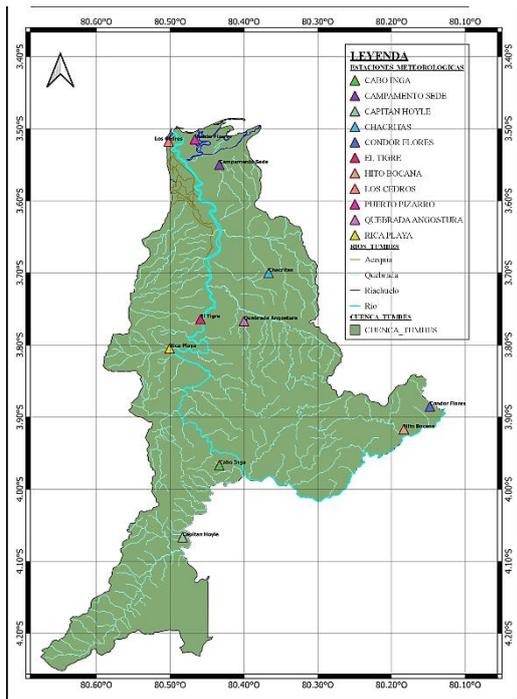
De acuerdo con los resultados obtenidos en la comprobación del Producto Pisco en nuestra área de estudio a nivel estación meteorológica, elevación, estacionalidad y región climática. para ello se puede considerar el uso de los datos de precipitación Para la realización de estudios hidrológicos como son análisis de oferta hídrica, fines de irrigación, balances hídricos, para zonas con escasa información o con información faltante, todo esto teniendo en cuenta su desempeño.

## REFERENCIAS BIBLIOGRÁFICAS

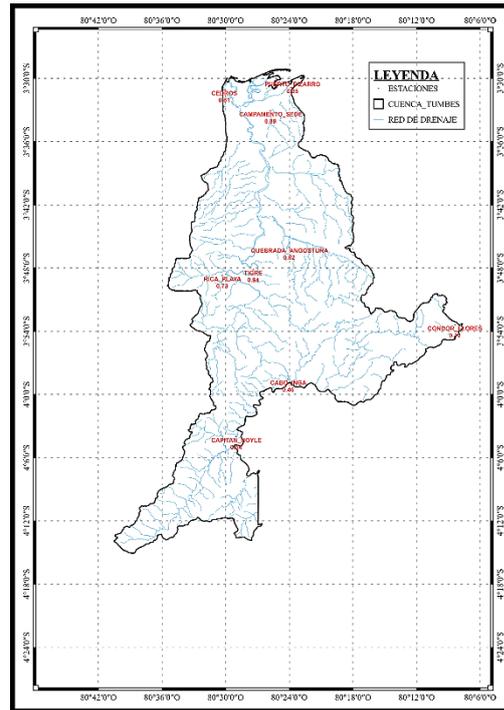
- Achim Zeileis and Gabor Grothendieck (2005). zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software*, 14(6), 1-27. doi:10.18637/jss.v014.i06
- Aybar, C., Fernández, C., Huerta, A., Lavado, W., Vega, F., & Felipe-Obando, O. (2019). Construction of a high-resolution gridded rainfall dataset for Peru from 1981 to the present day. *Hydrological Sciences Journal*, 65(5), 770-785. <https://doi.org/10.1080/02626667.2019.1649411>
- C.M, A. M., Chowdary, V. M., Kesarwani, M., & Neeti, N. (2022). Integrated drought monitoring and assessment using multi-sensor and multi-temporal earth observation datasets: a case study of two agriculture-dominated states of India. *Environmental Monitoring and Assessment*, 195(1). <https://doi.org/10.1007/s10661-022-10550-6>
- Cabrera, J., Yupanqui, R. T., & Rau, P. (2016). Validation of TRMM daily precipitation data for extreme events analysis. the case of piura watershed in peru. *Procedia Engineering*, 154, 154-157. <https://doi.org/10.1016/j.proeng.2016.07.436>
- Calderón Saldaña, J. P., & Alzamora de los Godos, L. (2018). Diseños de Investigación para Tesis de Posgrado. *Revista Peruana de Psicología y Trabajo Social*, 7-N° 2, 71-76.
- Chung-hong Chan, Geoffrey CH Chan, Thomas J. Leeper, and Jason Becker (2021). rio: A Swiss-army knife for data file I/O. R package version 0.5.29.
- Dagnino S., J. (2014). Tipos de estudios. *Revista Chilena de Anestesia*, 43(2), 104-108.
- Essou, G. R. C., Brissette, F., & Lucas-Picher, P. (2017). The Use of Reanalyses and Gridded Observations as Weather Input Data for a Hydrological Model: Comparison of Performances of Simulated River Flows Based on the Density of Weather Stations. *Journal of Hydrometeorology*, 18(2), 497-513. <https://doi.org/10.1175/jhm-d-16-0088.1>
- Fernandez-Palomino, C. A., Hattermann, F. F., Krysanova, V., Lobanova, A., Vega-Jácome, F., Lavado, W., Santini, W., Aybar, C., & Bronstert, A. (2022). A novel high-resolution gridded precipitation dataset for peruvian and ecuadorian watersheds: Development and hydrological evaluation. *Journal of Hydrometeorology*, 23, 309-336. <https://doi.org/10.1175/JHM-D-20-0285.1>
- Gob.pe. (17 de diciembre). Fechas de las estaciones astronómicas en el Perú. Recuperado de <https://www.gob.pe/11000-fechas-de-las-estaciones-astronomicas-en-el-peru>
- Mamani Washualdo, E. L. (2019). Regionalización Espacial de Precipitaciones Mensuales en la Cuenca de la Vertiente del Lago Titicaca-Lado Peruano. (Tesis previa a la obtención del Título de Ingeniero Agrícola, Universidad Nacional del Altiplano). Recuperada de <http://repositorio.unap.edu.pe/handle/UNAP/12489>
- Marco-Dos Santos, G., Melendez-Pastor, I., Navarro-Pedreño, J., & Koch, M. (2019). Assessing Water Availability in Mediterranean Regions Affected by Water Conflicts through MODIS Data Time Series Analysis. *Remote Sensing*, 11(11), 1355. <https://doi.org/10.3390/rs11111355>
- Mauricio Zambrano-Bigiarini. (2020) hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time seriesR package version 0.4-0, URL<https://github.com/hzambbran/hydroGOF>. DOI:10.5281/zenodo.839854.
- Mendoza, J., Rojas, J., Jimenez, C., Leva, A., & Ortega, Y. (2022). Spatio-temporal variation of the change in total water storage and precipitation using GRACE and TRMM data: Lima-Callao in Peru; [Variación espacio-temporal del cambio de agua total almacenada y la precipitación usando datos GRACE y TRMM de Lima-Callao en Perú]. *Proceedings of the LACCEI International Multi-Conference for Engineering, Education and Technology*, 2022-July.0020https://doi.org/10.18687/LACCEI2022.1.1.275
- Muñoz, P., Corzo, G., Solomatine, D., Feyen, J., & Céleri, R. (2023). Near-real-time satellite precipitation data ingestion into peak runoff forecasting models. *Environmental Modelling & Software*, 160, 105582. <https://doi.org/10.1016/j.envsoft.2022.105582>
- Nieto-Chaupis, H. (2019). Volterra convolution as a tool to assess and forestall flooding and aftermaths along the far coast Peruvian. 2018 IEEE Biennial Congress of Argentina, ARGENCON 2018. <https://doi.org/10.1109/ARGENCON.2018.8645990>
- Porto Lopez, A. E. (2021). Validación del producto grillado PISCO de precipitación v2.1 con la data observada de las estaciones pluviométricas del SENAMHI en la cuenca de la vertiente del lago Titicaca - lado peruano. (Tesis previa a la obtención del

- Título de Ingeniero Agrícola, Universidad Nacional del Altiplano). Recuperada de <http://repositorio.unap.edu.pe/handle/UNAP/15498>
- R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Tierney N, Cook D, McBain M, Fay C (2021). `naniar`: Data Structures, Summaries, and Visualisations for Missing Data. R package version 0.6.1, <<https://CRAN.R-project.org/package=naniar>>.
- Vaheddoost, B., Safari, M. J. S., & Yilmaz, M. U. (2023). Rainfall-Runoff Simulation in Ungauged Tributary Streams Using Drainage Area Ratio-Based Multivariate Adaptive Regression Spline and Random Forest Hybrid Models. *Pure and Applied Geophysics*. <https://doi.org/10.1007/s00024-022-03209-3>
- Villavicencio, E. E., Medina, K. D., Loarte, E. A., & León, Hairo A. (2022). Improved rainfall and temperature satellite dataset in areas with scarce weather stations data: case study in Ancash, Peru; [Mejora de los datos satelitales de precipitación y temperatura en áreas con baja disponibilidad de estaciones meteorológicas: caso de estudio en Ancash, Perú]. *Revista de Teledetección*, 2022, 17–28. <https://doi.org/10.4995/raet.2022.16907>
- Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Golemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). "Welcome to the tidyverse." *Journal of Open-Source Software*, 4(43), 1686. doi:10.21105/joss.01686
- Wickham H, Bryan J (2022). `readxl`: Read Excel Files. R package version 1.4.1<<https://CRAN.R-project.org/package=readxl>>.
- Wickham H, François R, Henry L, Müller K (2022). `dplyr`: Grammar of Data Manipulation. R package version 1.0.10, <https://CRAN.R-project.org/package=dplyr>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Golemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., & Takahashi, K. (2019). Welcome to the Tidyverse. *Journal of Open-Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Zapana Quispe, M. M. (2019). Análisis de la sequía meteorológica de la Intercuenca Alto Apurímac y las cuencas Alto Urubamba y Yavero utilizando el método de L-Momentos. Universidad Nacional Agraria la Molina.
- Zeileis, A., & Grothendieck, G. (2005). zoo: S3Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software*, 14(6). <https://doi.org/10.18637/jss.v014.i06>
- Zubieta, R., Getirana, A., Espinoza, J. C., & Lavado, W. (2015). Impacts of satellite-based precipitation datasets on rainfall-runoff modeling of the Western Amazon basin of Peru and Ecuador. *Journal of Hydrology*, 528, 599–612. <https://doi.org/10.1016/j.jhydrol.2015.06.064>
- Rojo-Báez, I., García-Estrada, R. S., León-Félix, J. Sañudo-Barajas, A., & Allende-Molar, R. (2016). Histopatología del proceso de infección de *Colletotrichum truncatum* en hojas de papaya y chicharo. *Revista Mexicana de Fitopatología*, 34(3), 316-325. doi: <https://doi.org/10.18781/r.mex.fit.1604-3>
- Agrios, G. (1995). *Fitopatología*. S.A. de C. V. México, Mexico: Limusa.
- Ministerio de Agricultura & Riego-MINAGRI. (2019). *Encuesta nacional de intenciones de siembra 2019*. Lima, Perú: Dirección general de seguimiento y evaluación de políticas dirección de estadística agraria

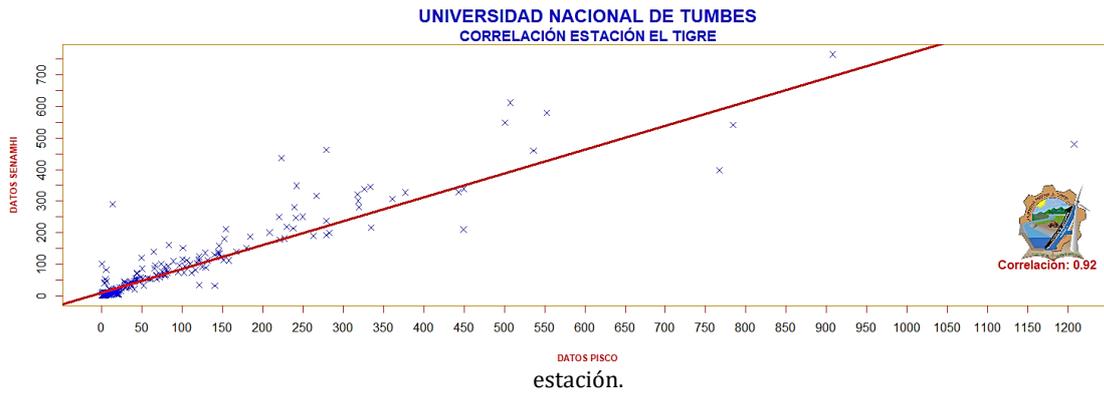
ANEXOS



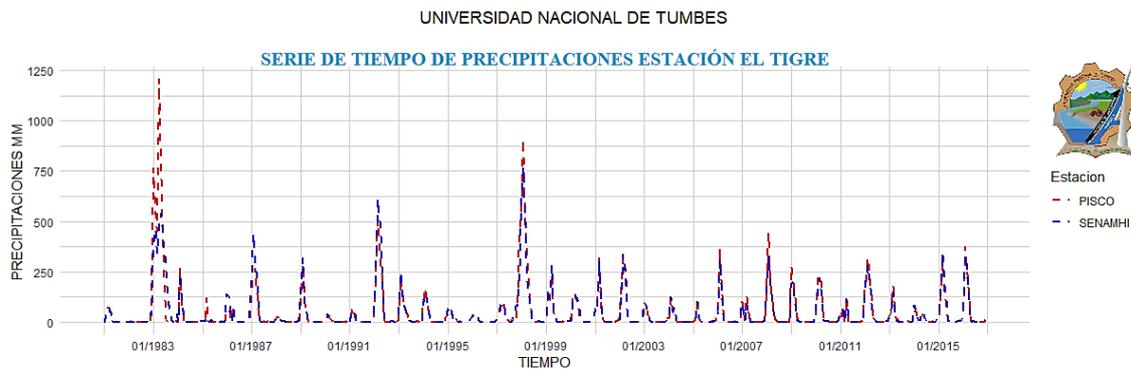
Anexo 1. Mapa de ubicación de las estaciones.



Anexo 2. Coeficiente de determinación a nivel



Anexo 3. Correlación estación el Tigre entre data observada y estimada.



Anexo 4. Gráfico de serie de tiempo estación el Tigre entre data observada y estimada.