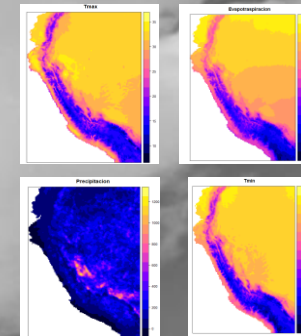


MANEJO DE DATOS HIDROMETEOROLOGICOS CON R



Presentado por:

Garcia Tito, Miguel Angel

Contacto: garciatitomiguel@gmail.com



UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA

CONTENIDO



1

¿Qué es R?

2

Ciclo Hidrológico

3

Datos Grillados

4

Productos Climáticos Globales

5

Librerías de R aplicado a la hidrología

6

Preguntas comunes

7

Conclusiones y Recomendaciones

¿QUÉ ES R?

R es un entorno de software libre (licencia GNU GLP) y lenguaje de programación interpretado, es decir, ejecuta las instrucciones directamente, sin una previa compilación del programa a instrucciones en lenguaje máquina.



Desarrollado por Robert Gentleman y Ross Ihaka en Univ. Auckland (NZ) en 1993, sin embargo sus orígenes se remontan al lenguaje S (1976).

Características que debo saber de R

- Es gratuito
- Amplia comunidad científica
- Variedad de paquetes especializados para diferentes ámbitos
- Gráficos y reportes estadísticos de alta calidad y de manera rápida
- Disponible para los sistemas Windows, Linux y Mac
- El entorno IDE Rstudio es sumamente amigable



GNU: Sistema operativo UNIX

GPL: Licencia publica general

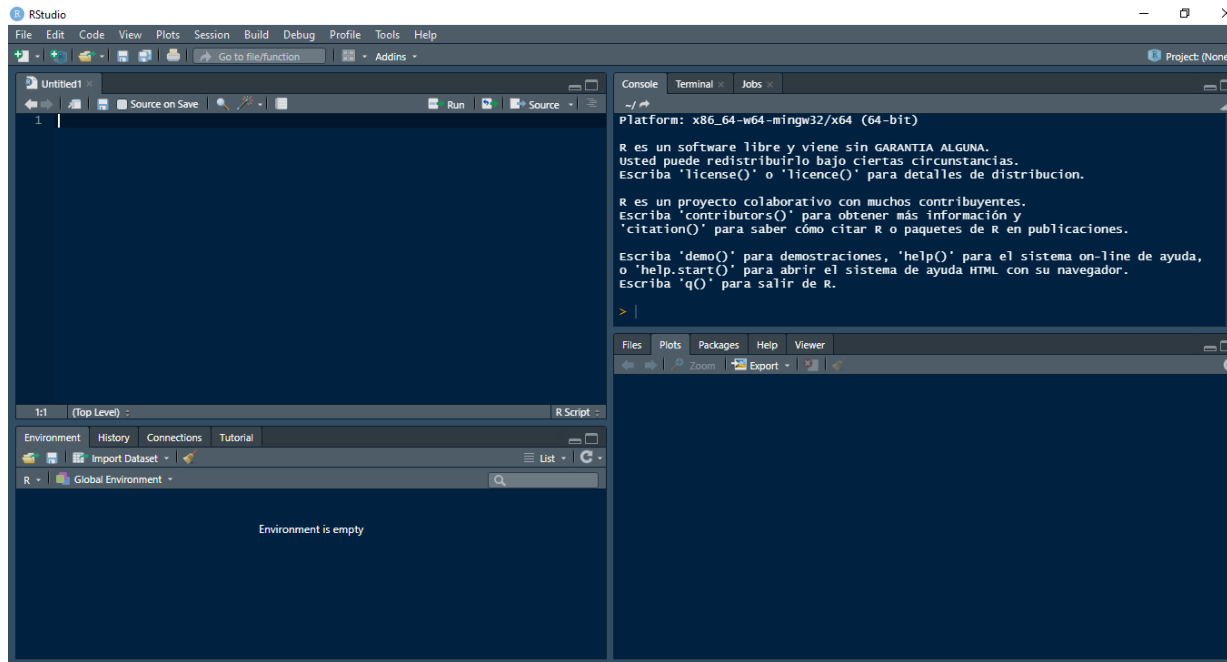
<https://www.r-project.org/>



R es un lenguaje de programación y entorno computacional dedicado a la estadística.

RSTUDIO

Si ya hemos instalado R en nuestro equipo, RStudio lo detectará automáticamente y podremos utilizarlo desde este entorno. Si no instalamos RStudio antes que R, no hay problema, cada vez que iniciamos este programa, verificará la instalación de R.



<https://www.rstudio.com/products/rstudio/download/>

Hidrología

Medicina

Teledetección

Gráficos

Análisis
Espacial

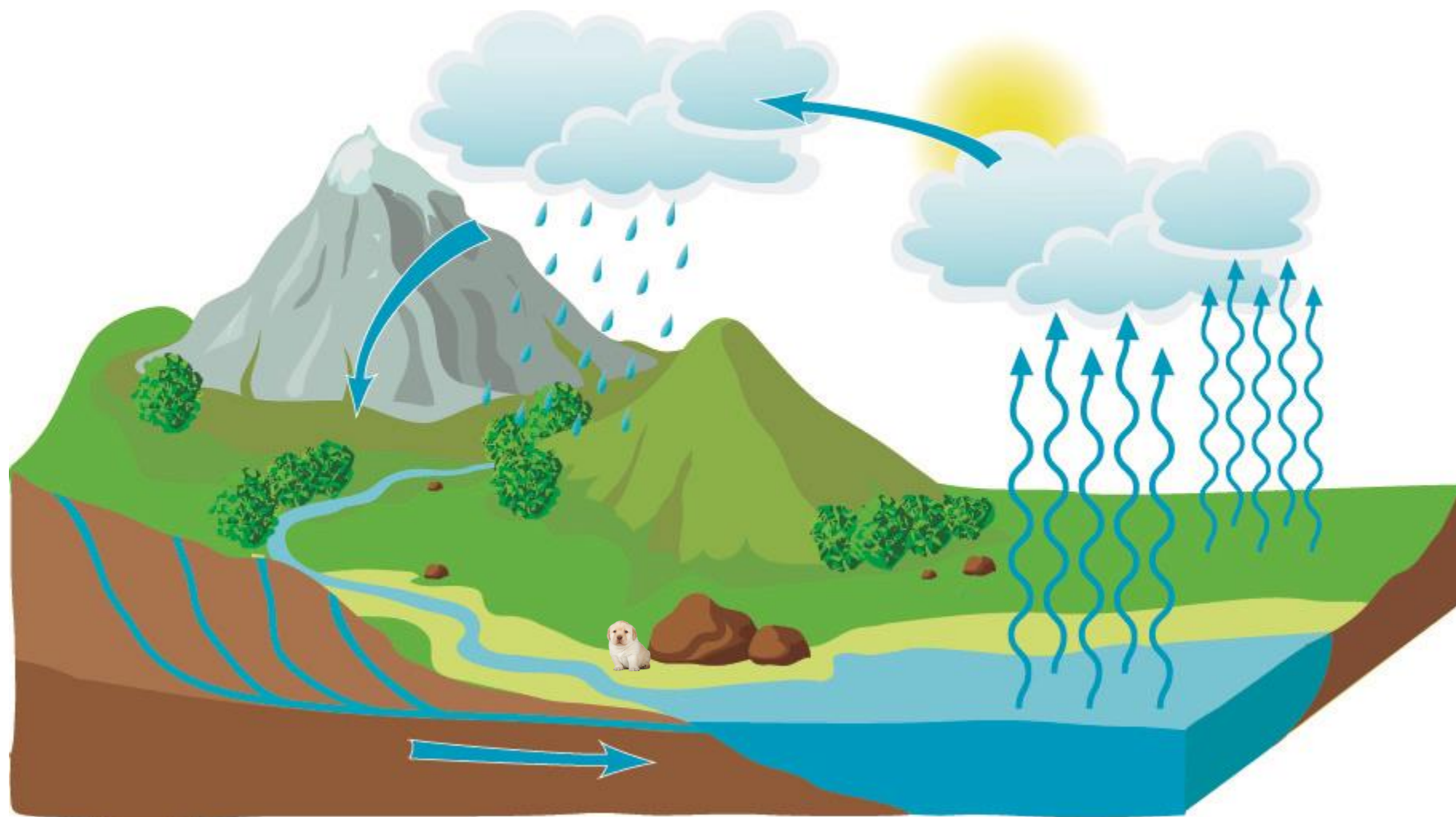
Minería de
datos

Biología

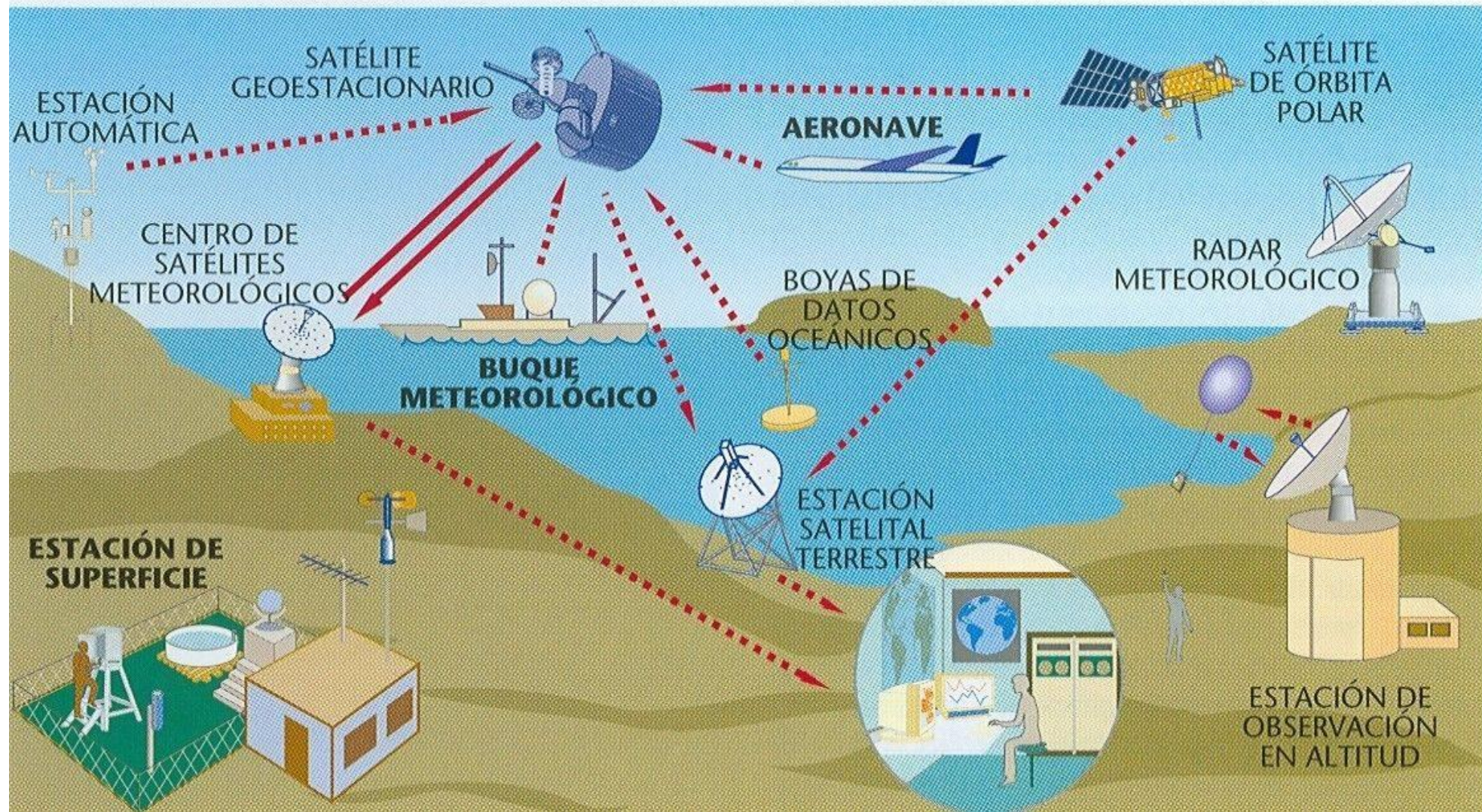
Climatología

Hay varias opciones de IDE para R, y entre ellas mi preferido es RStudio.





CICLO HIDROLOGICO



SISTEMA DE OBSERVACIÓN GLOBAL

ABSTRACT

PISCOp se ha desarrollado para el período 1981 hasta la actualidad, promedio de ocho semanas a una resolución espacial de $0,1^\circ$. El algoritmo de fusión se basa en estadísticas geoestadísticas y deterministas. métodos de interpolación que incluyen tres fuentes de lluvia diferentes: (i) el control nacional de calidad y el conjunto de datos de pluviómetros rellenos, (ii) climatologías de precipitación combinadas con indicadores de radar y (iii) Estimaciones de Precipitación Infrarroja del Grupo de Peligros Climáticos (CHIRP).

Los resultados de la validación sugieren que las estimaciones de precipitación son aceptables y muestran el desempeño más alto para la costa del Pacífico y el flanco occidental de los Andes.



Hydrological Sciences Journal



ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: <https://www.tandfonline.com/loi/thsj20>

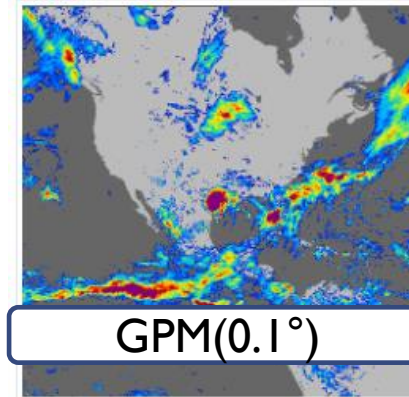
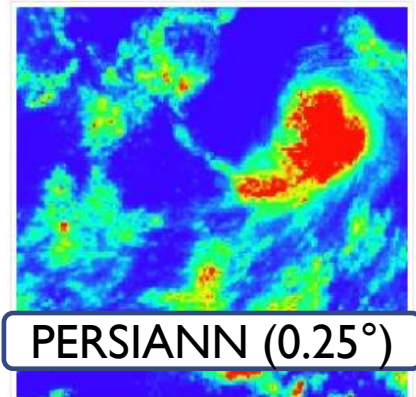
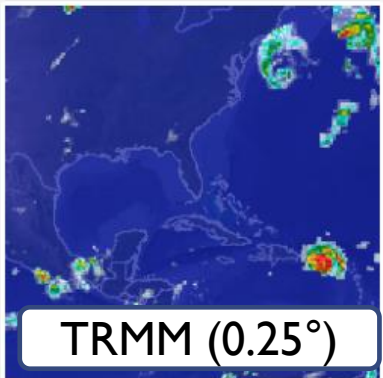
Construction of a high-resolution gridded rainfall dataset for Peru from 1981 to the present day

Cesar Aybar, Carlos Fernández, Adrian Huerta, Waldo Lavado, Fiorella Vega & Oscar Felipe-Obando

FINALIDAD

Los pluviómetros ofrecen una acertada información acerca de la cantidad y la frecuencia de las lluvias. Sin embargo, no posibilitan una adecuada cobertura espacial (Almazroui, 2011; Chappell et al., 2013). La estimación de la distribución espacial de las precipitaciones en el Perú representan un gran desafío, considerando su compleja fisiografía y la baja densidad de estaciones pluviométricas sobre todo en la región amazónica (Mantas et al., 2015).

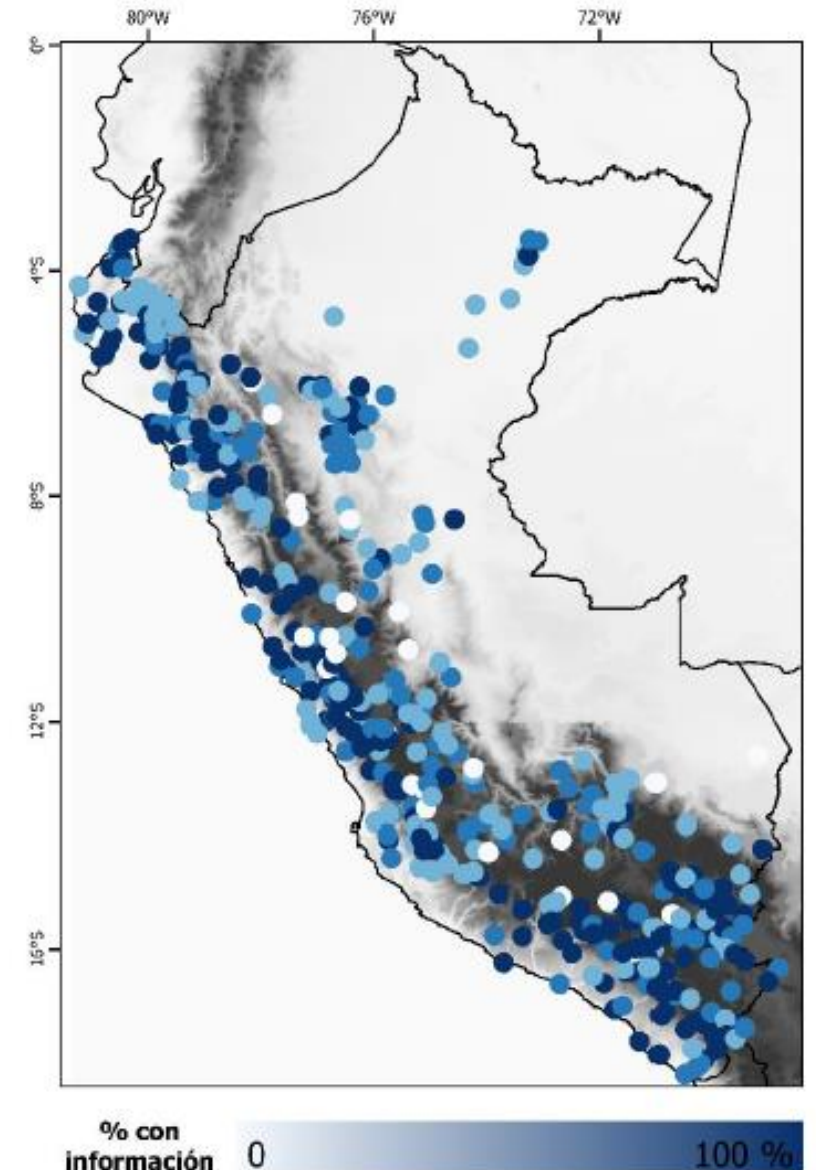
En la actualidad, la libre disponibilidad de productos satelitales de estimación de lluvias como el "Tropical Rainfall Measuring Mission" (TRMM), el "National Oceanic and Atmospheric Administration's Climate Prediction Center morphing technique" (CMORPH), el "Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks" (PERSIANN) y últimamente el "Global Precipitation Measurement" (GPM) son una alternativa valiosa para mejorar los vacíos de información pluviométrica en muchas partes del mundo.



CONSTRUCCIÓN

■ Para la generación de PISCO – precipitación se utilizaron datos que consistieron en : i) una base de precipitación de todo el Perú y ii) datos satelitales del producto CHIRPS a escala mensual. Los datos observados fueron obtenidos por el Servicio Nacional de Meteorología e Hidrología (SENAMHI). Estos datos fueron sometidos a un proceso de control de calidad visual enfocada en reportar valores atípicos y quiebres temporales de precipitación.

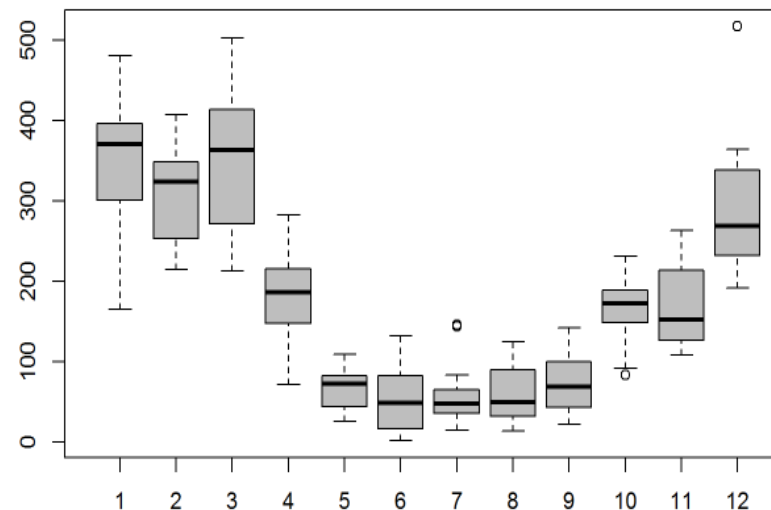
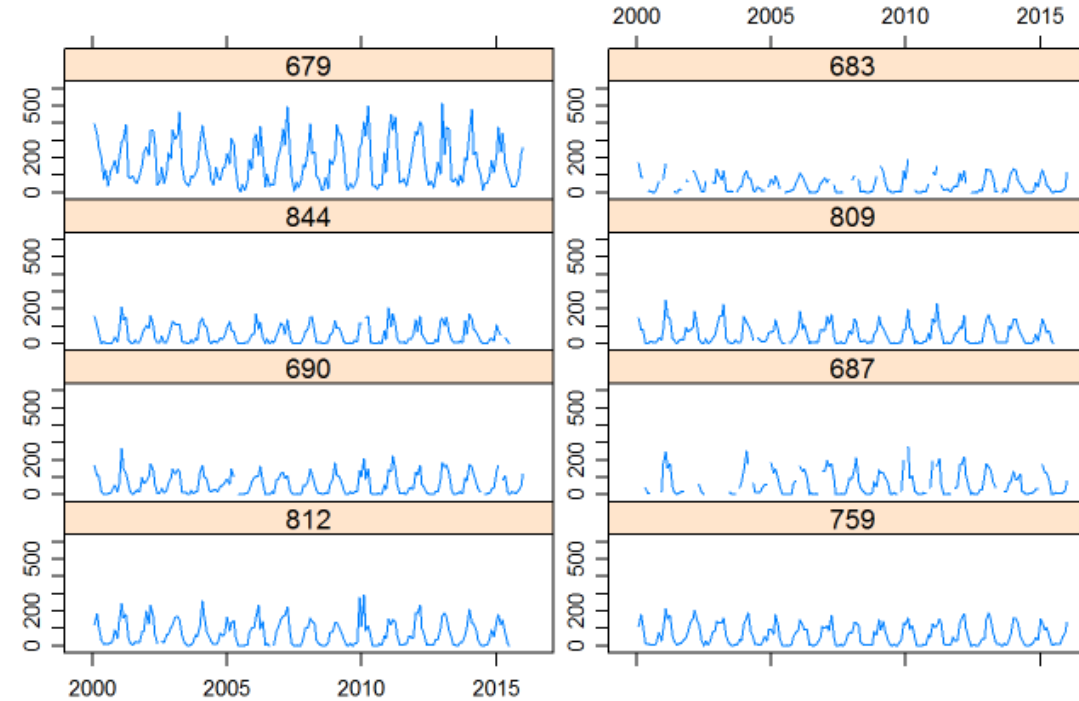
Este procedimiento el más tedioso de todos, fue realizado a 681 estaciones meteorológicas gestionadas por el Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI). El proceso de control de calidad se dividió en dos partes: automático y visual.



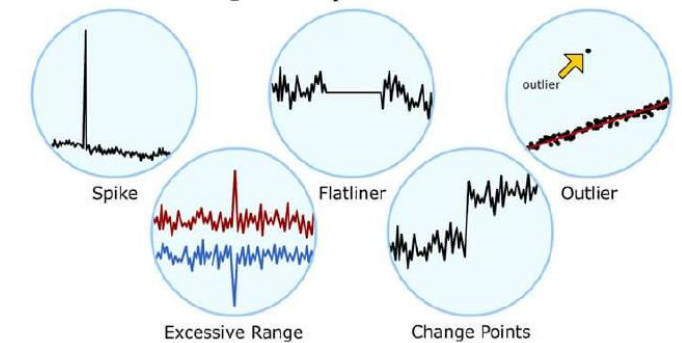
CONTROL DE CALIDAD

El control automático, consistió en la evaluación general de los datos de precipitación para cada estación (fechas duplicadas, comas decimales, valores -999 y/o -888), rangos fijos, coherencia temporal y espacial. Los valores que superaron los límites establecidos fueron eliminados.

Este procedimiento el más tedioso de todos, fue realizado a 681 estaciones meteorológicas gestionadas por el Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI). El proceso de control de calidad se dividió en dos partes: automático y visual.



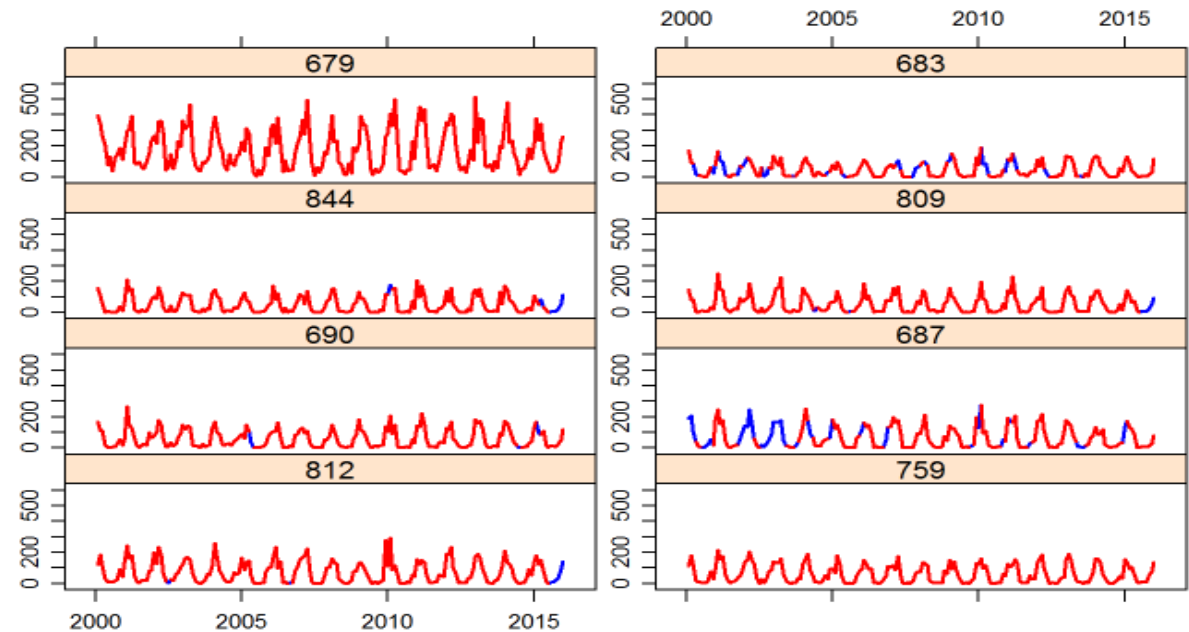
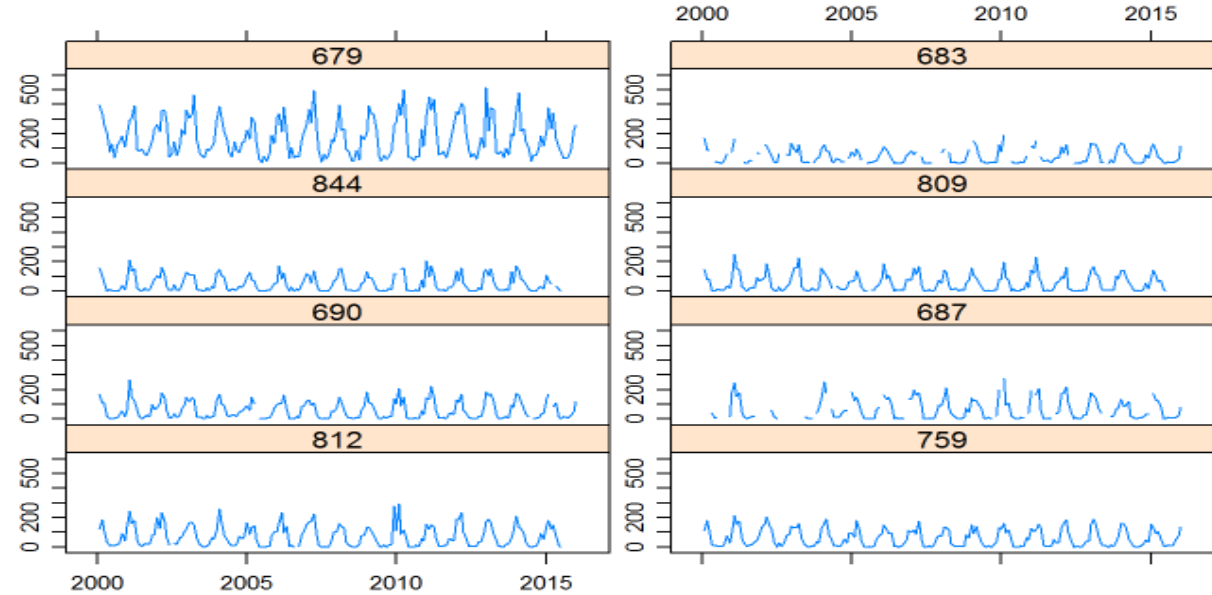
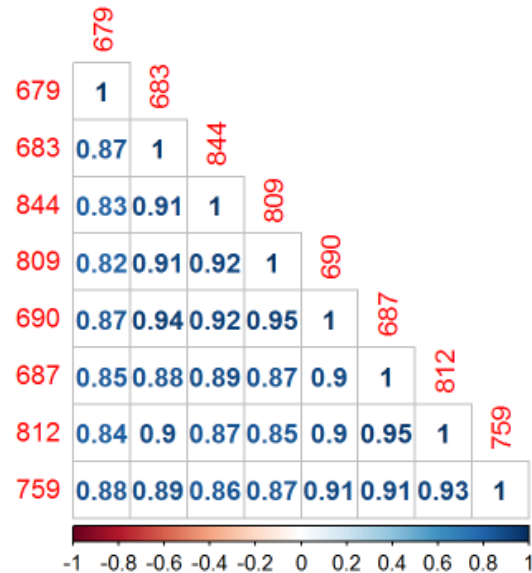
Quality Control



Meteorologists use computers to perform a basic check of all incoming data. Are there spikes, flatliners, outliers, excessive range, or change points in the data?

CONTROL DE CALIDAD

Además, la completación de datos de las series mensuales y diarias fue llevado a cabo basado en una modificación del algoritmo CUTOFF ([Feng et al., 2014](#)) y mapeos de cuantiles ("quantile mapping") respectivamente.



TRMM Y CHIRPS

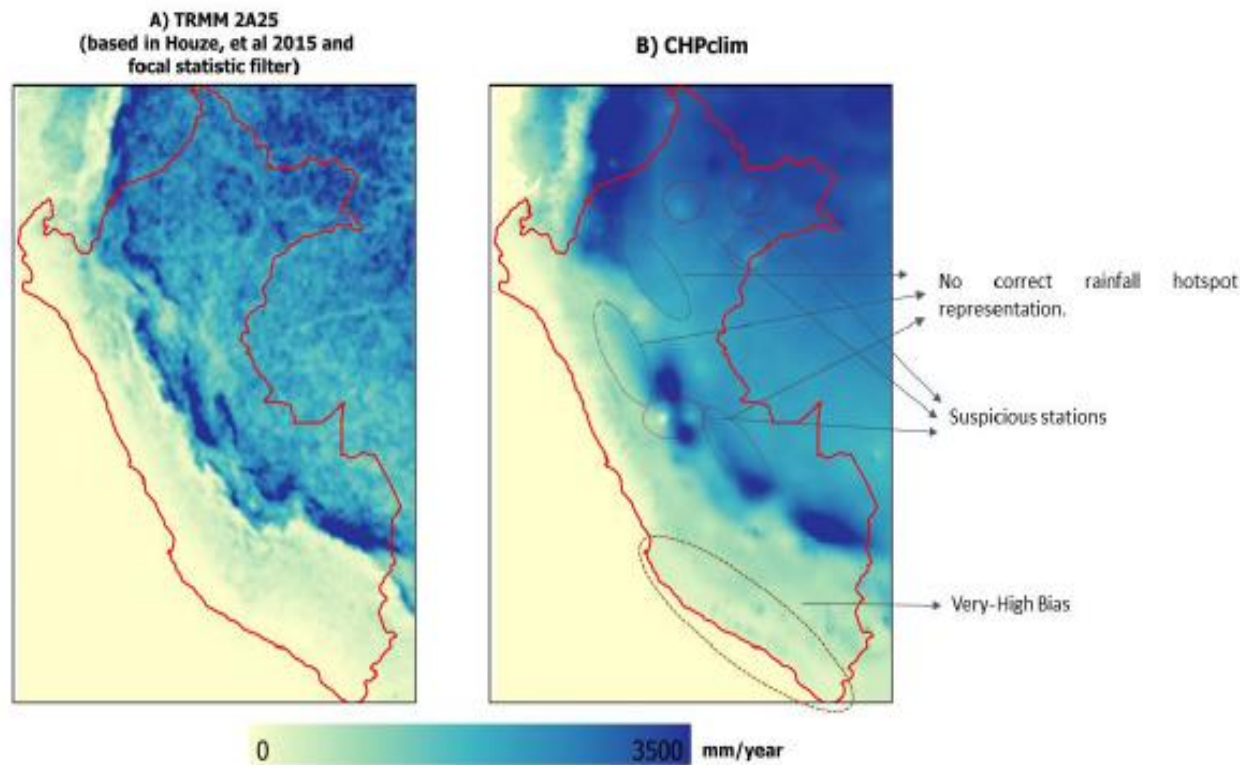


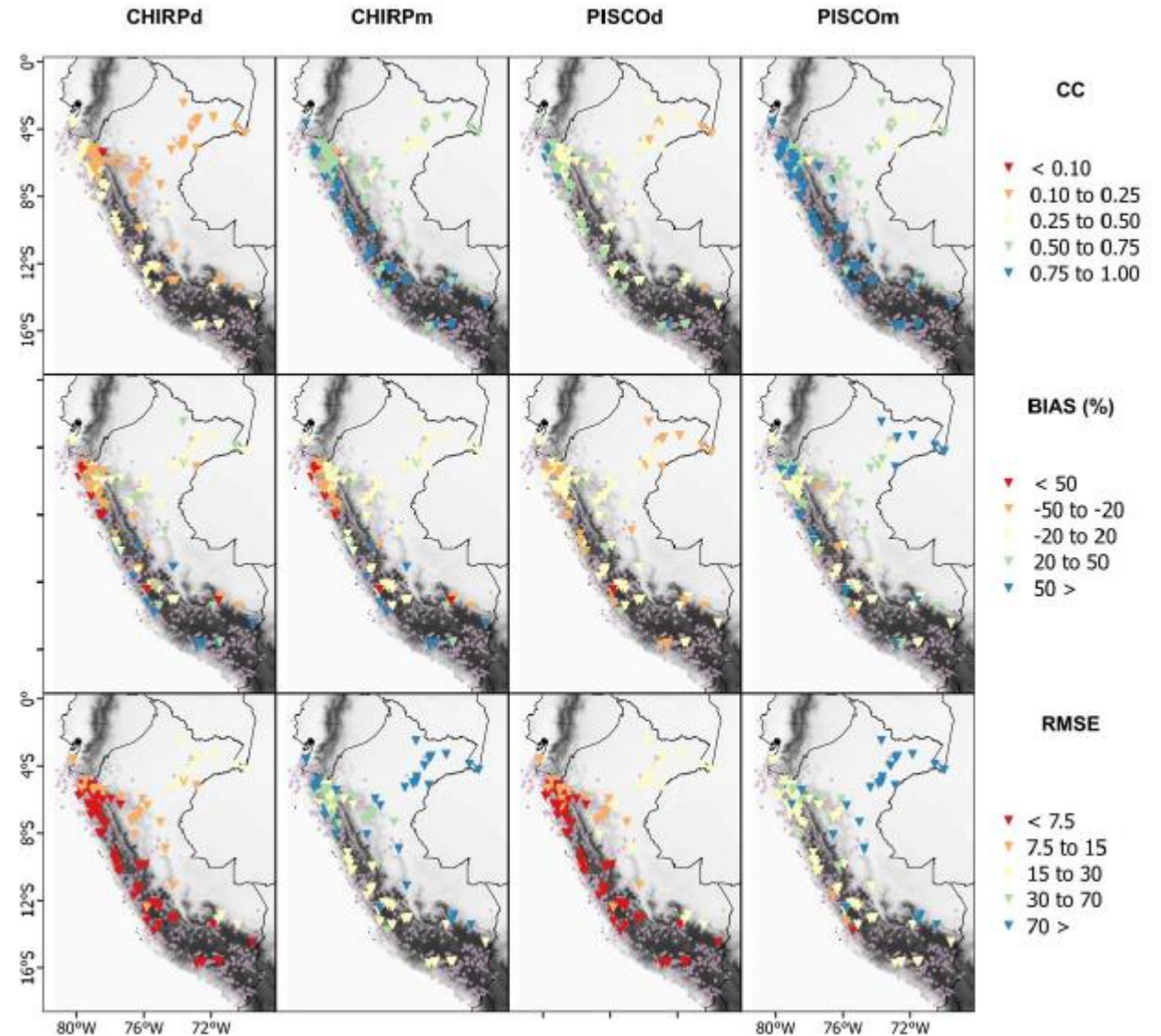
Figura 2. A la izquierda climatología de precipitación (1998-2014) generado a partir de datos TRMM2A25 y a la derecha CHPclim nótese los excesivos valores en la costa sur del país.

Sub-region	Elevation (m.a.s.l)	Climate Driver	Rainfall Regime	N
Peruvian Pacific Coast (PC)	0 – 1500	ITCZ, HCS, ENOS	Wet (Dec - May)	97
			Dry (Jun - Nov)	
Andes western slope (AW)	> 1500	Elevation, ITCZ	Wet (Dec - May)	151
			Dry (Jun - Nov)	
Andes eastern slope (AE)	> 1500	Elevation, Orography, ITCZ	Weak seasonality, drier JJA	128
Andes-Amazon transition (AAT)	500 – 1500	Orography, ITCZ, SALLJ	Weak seasonality, drier JJA	26
Amazon lowland (AL)	0 – 500	ITCZ, trade winds	Weak seasonality, drier JJA	39

	Name	Formula	Perfect Score
Continuous statistics	Correlation Coefficient (CC)	$CC = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2 \sum(Y - \bar{Y})^2}}$	1
	Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{N} \sum (X - Y)^2}$	0
	Percentage Bias (PBIAS)	$PBIAS = 100 \times (\sum (X - Y) / \sum X)$	0
Categorical statistics	Probability of detection (POD)	$POD = A / (A + C)$	1
	False alarm ratio (FAR)	$FAR = B / (A + B)$	0
	Threat score (TS)	$TS = A / (A + B + C)$	1

VALIDACIÓN

PISCOpm y PISCOpd evidencia una mejora significativa del coeficiente de correlación (CC) frente a CHIRPd y CHIRPm sobretudo en la costa y sierra, esto está relacionado a la distancia entre las estaciones utilizadas en la validación y las estaciones utilizadas en PISCOp, el bajo desempeño de PISCOp en la selva es debido a la baja densidad de estaciones que imposibilita la caracterización de la estructura espacial (mediante el variograma) de forma adecuada. A diferencia del CC que nos permite definir la variabilidad de las precipitaciones, tanto el BIAS % como el RMSE nos permite obtener información sobre las intensidades.



VERSIONES DE PISCO HASTA LA ACTUALIDAD

Versión	Covariable	Escala temporal	Método de Interpolación	Imputación	Observaciones	Cobertura temporal
1.0	CHIRPS	Mensual	KED	NO	Fuertes quiebres artificiales	1981-2014
1.1	CHIRPS	Mensual y diaria	KED, IDW y RIDW	NO	Fuertes quiebres artificiales	1981-2015
1.2	CHIRP	Mensual y diario	RIDW	SI	Quiebres Leves artificiales	1981-2016
2.0	CHIRP TRMM2A25	Mensual y diario	KED y RIDWO	SI	Quiebres leves artificiales	1981-2016

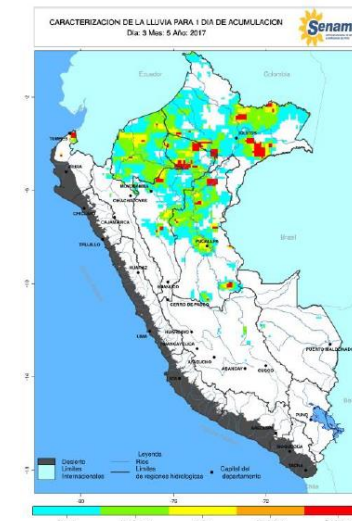
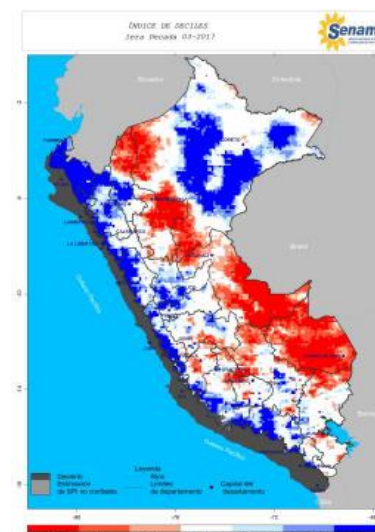
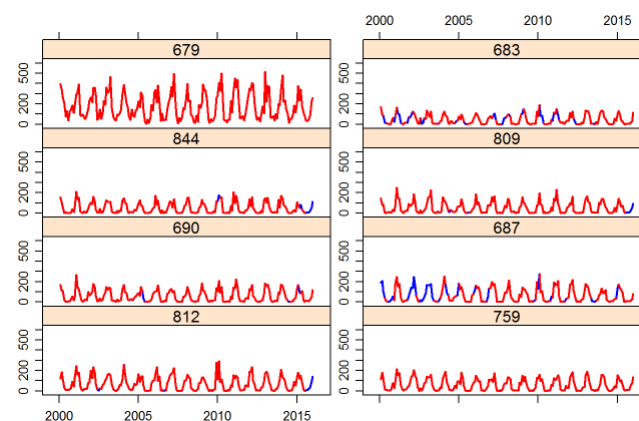
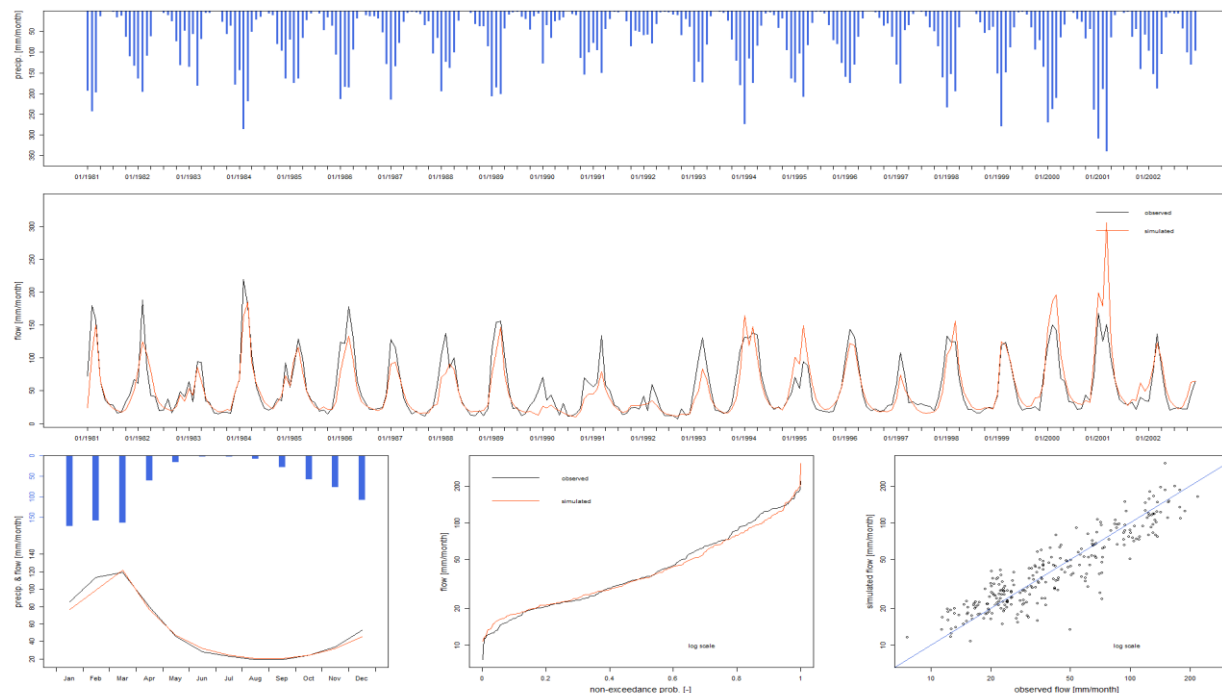
Las primeras versiones de PISCO precipitación (1.0 y 1.1) se caracterizan por la falta de completación de la serie de datos, lo cual repercute directamente en la generación de quiebres artificiales sobre todo a partir del 2010.

La versión 1.0 solo genera datos mensuales, los datos de pluviómetros no presentan completación por lo que se evidencia la formación de fuertes quiebres sobre todo a partir del 2010.

En la segunda versión de PISCO (descrita en este reporte) se generan datos mensuales y diarios, el principal cambio radica en que el control de calidad ha sido mucho más riguroso, además de completar los datos faltantes tanto para las series diarias como las mensuales la covariable cambia de CHIRPS a CHIRPM a fin de mejorar la predicción de las intensidades.

UTILIDADES DE PISCO

- MONITOREO DE SEQUIAS
- MONITOREO DE INUNDACIONES
- MONITOREO DECAIARIO DE PRECIPITACIÓN
- MODELAMIENTO HIDROLOGICO DE CUENCAS
- COVARIABLES PREDICTORAS PARA DESARROLLO DE MICROCLIMAS
- ESTUDIOS HIDROLOGICOS (PREVIO TRATAMIENTO DE DATOS)
- COMPLETACIÓN DE INFORMACIÓN (PREVIO TRATAMIENTO)



DATOS CLIMATICOS CMIP 5 - CMIP 6



Proyecto de inter-comparación de modelos de clima acoplados (CMIP en sus siglas en inglés, Coupled Model Intercomparison Project)

CMIP 5 (rcps) (2010 – 2014)

- En 2008, el Grupo de trabajo del WCRP sobre modelado acoplado, en su duodécima sesión , aprobó el protocolo CMIP5
- 35 Experimentos de modelos climáticos
- Evaluar los mecanismos responsables de las diferencias del modelo en comprender las retroalimentaciones asociadas con el ciclo del carbono y las nubes.
- Examinar la "previsibilidad" del clima y explorar la capacidad de los modelos para predecir el clima en escalas de tiempo decenales y, de manera más general.

CMIP 6 (2013 – Actual)

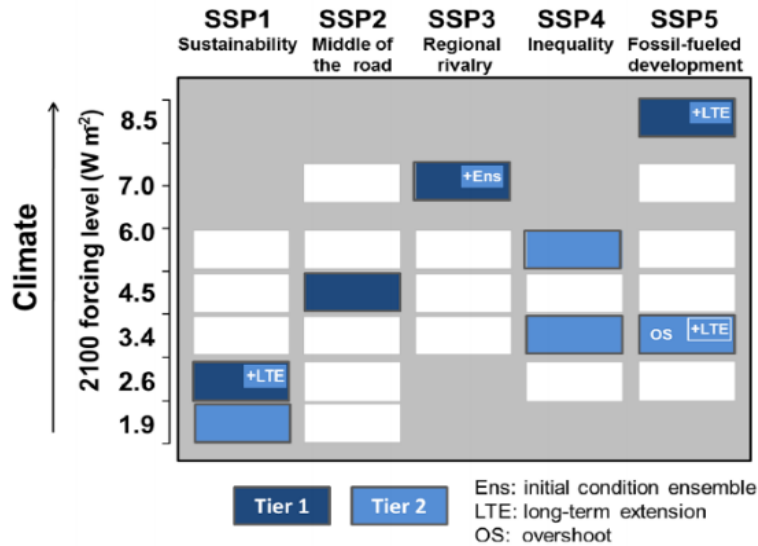
- Actualmente se cuentan con 23 CMIPs Aprobados
- 33 grupos de investigación
- Ahora son llamados ssps
- Contribuye a las proyecciones de un mayor calentamiento en este siglo, alrededor de 0,4 ° C a 0,9 ° C más cálido que en escenarios similares ejecutados en CMIP5

RCP : Vías de concentración representativa
SSP: Caminos socioeconómicos compartidos

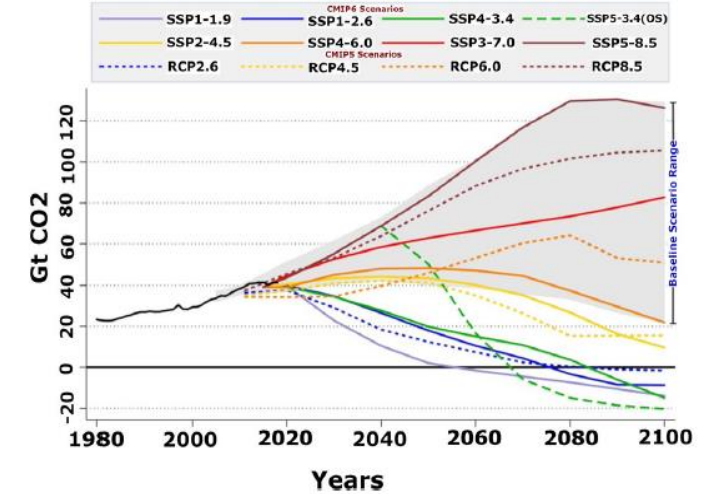
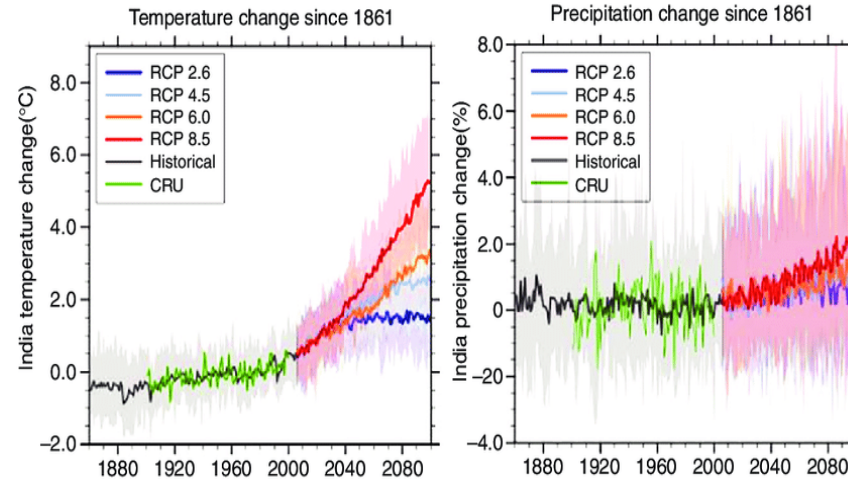


MODELOS CLIMATICOS GLOBALES

Shared socioeconomic pathways



PAPERS



El IPCC AR5 presentó cuatro vías de concentración representativas (RCP) que examinaron diferentes posibles emisiones futuras de gases de efecto invernadero. Estos escenarios (RCP2.6 , RCP4.5 , RCP6.0 y RCP8.5) tienen nuevas versiones en CMIP6. Estos escenarios actualizados se denominan SSP1-2.6, SSP2-4.5, SSP4-6.0 y SSP5-8.5, cada uno de los cuales da como resultado niveles de forzamiento radiativo de 2100 similares a los de su predecesor en AR5.

Glacial mass balance changes in the Karakoram and Himalaya based on CMIP5 multi-model climate projections

Rajiv K. Chaturvedi • Anil Kulkarni •
Yogesh Karyakarte • Jaideep Joshi • G. Bala

Skill assessment of global climate model wind speed from CMIP5 and CMIP6 and evaluation of projections for the Bay of Bengal

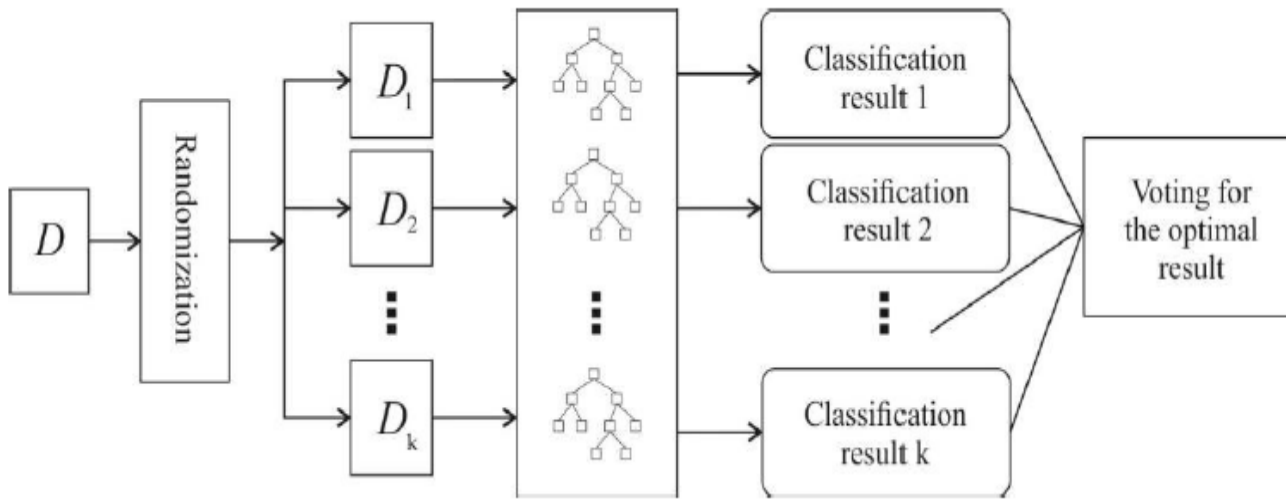
Athira Krishnan¹ • Prasad K. Bhaskaran¹

Received: 14 May 2020 / Accepted: 29 July 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

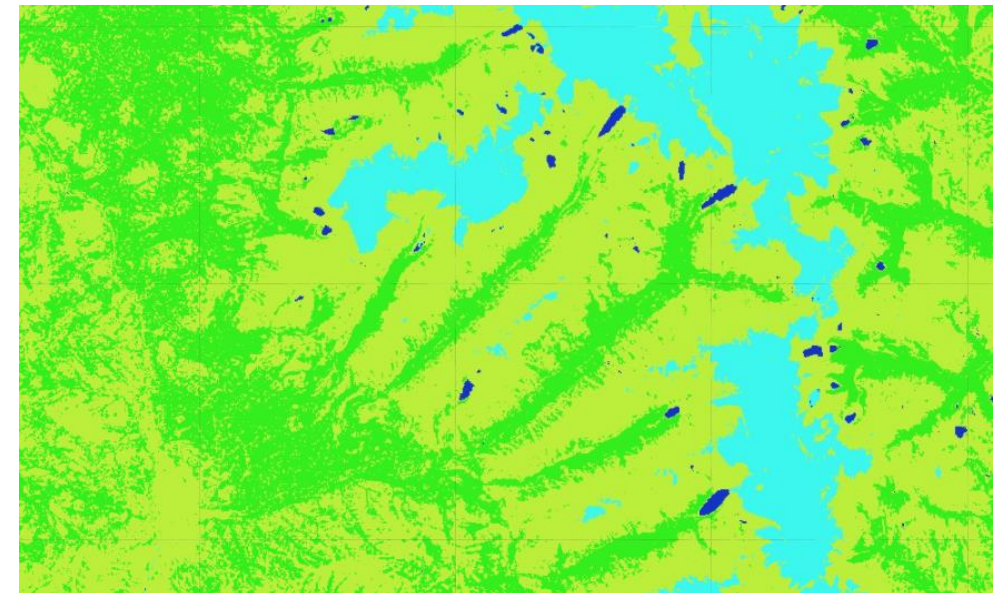
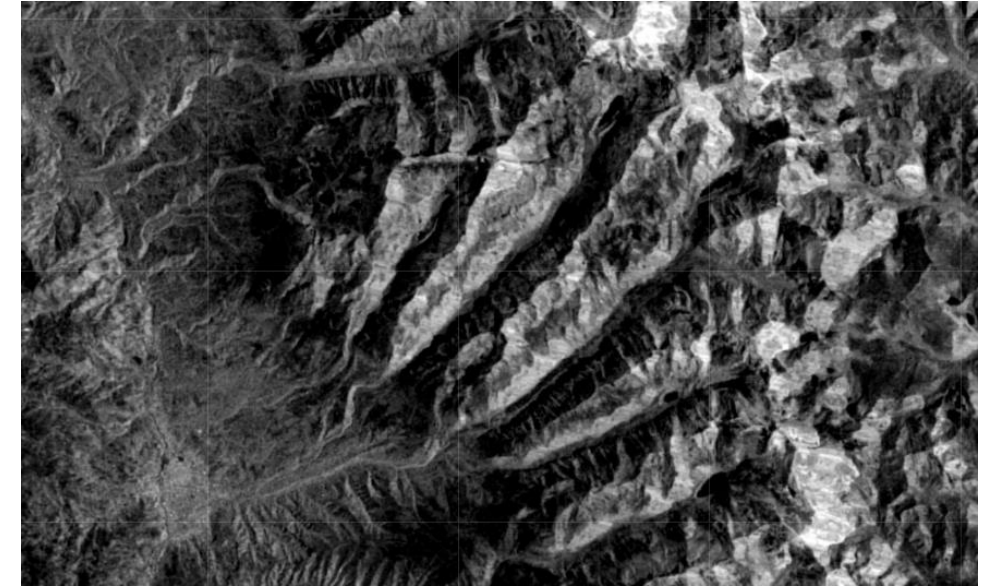
MODELO DE CLASIFICACIÓN RANDOM FOREST

Random forest (RF), propuesto por (L. Breiman, 2001), es un modelo de regresión y clasificación que consiste de una colección de clasificadores estructurados por árbol (decision tree), donde cada uno trabaja en función a los valores de un vector o muestra.

El aprendizaje del árbol crece en el proceso de recopilación de datos usando la metodología Cart(Clasification and regression tree).



Fuente: Wang & Rudin (2015).



CMIP6.ScenarioMIP.NOAA-GFDL.GFDL-ESM4.ssp119.r1i1p1f1.Amon.tasmax.gr1

Data Node: esgdata.gfdl.noaa.gov

Version: 20180701

Total Number of Files (for all variables): 1

Full Dataset Services: [[Show Metadata](#)] [[List Files](#)] [[THREDDS Catalog](#)] [[WGET Script](#)] [[LAS Visualization](#)] [[Show Citation](#)] [[PID](#)]
[[Globus Download](#)] [[Further Info](#)]

CMIP6.ScenarioMIP.NOAA-GFDL.GFDL-ESM4.ssp119.r1i1p1f1.Amon.tasmin.gr1

Data Node: esgdata.gfdl.noaa.gov

Version: 20180701

Total Number of Files (for all variables): 1

Full Dataset Services: [[Show Metadata](#)] [[List Files](#)] [[THREDDS Catalog](#)] [[WGET Script](#)] [[LAS Visualization](#)] [[Show Citation](#)] [[PID](#)]
[[Globus Download](#)] [[Further Info](#)]

CMIP6.ScenarioMIP.NOAA-GFDL.GFDL-ESM4.ssp585.r1i1p1f1.Amon.tasmax.gr1

Data Node: esgdata.gfdl.noaa.gov

Version: 20180701

Total Number of Files (for all variables): 1

Full Dataset Services: [[Show Metadata](#)] [[List Files](#)] [[THREDDS Catalog](#)] [[WGET Script](#)] [[LAS Visualization](#)] [[Show Citation](#)] [[PID](#)]
[[Globus Download](#)] [[Further Info](#)]

CMIP6.ScenarioMIP.NOAA-GFDL.GFDL-ESM4.ssp585.r1i1p1f1.Amon.tasmin.gr1

Data Node: esgdata.gfdl.noaa.gov

Version: 20180701

Total Number of Files (for all variables): 1

Full Dataset Services: [[Show Metadata](#)] [[List Files](#)] [[THREDDS Catalog](#)] [[WGET Script](#)] [[LAS Visualization](#)] [[Show Citation](#)] [[PID](#)]
[[Globus Download](#)] [[Further Info](#)]

CMIP6.ScenarioMIP.NOAA-GFDL.GFDL-ESM4.ssp119.r1i1p1f1.Amon.pr.gr1

Data Node: esgdata.gfdl.noaa.gov

Version: 20180701

Total Number of Files (for all variables): 1

Full Dataset Services: [[Show Metadata](#)] [[List Files](#)] [[THREDDS Catalog](#)] [[WGET Script](#)] [[LAS Visualization](#)] [[Show Citation](#)] [[PID](#)]
[[Globus Download](#)] [[Further Info](#)]

CMIP6.ScenarioMIP.NOAA-GFDL.GFDL-ESM4.ssp585.r1i1p1f1.Amon.pr.gr1

Data Node: esgdata.gfdl.noaa.gov

Version: 20180701

Total Number of Files (for all variables): 1

Full Dataset Services: [[Show Metadata](#)] [[List Files](#)] [[THREDDS Catalog](#)] [[WGET Script](#)] [[LAS Visualization](#)] [[Show Citation](#)] [[PID](#)]
[[Globus Download](#)] [[Further Info](#)]

Resolution: 100km

	Institution	Country		Institution	Country		Institution	Country
1	AWI	Germany	12	DOE	USA	23	MRI	Japan
2	BCC	China	13	EC-Earth-Cons	Europe	24	NASA-GISS	USA
3	BNU	China	14	FGOALS	China	25	NCAR	USA
4	CAMS	China	15	FIO-RONM	China	26	NCC	Norway
5	CasESM	China	16	INM	Russia	27	NERC	UK
6	CCCma	Canada	17	INPE	Brazil	28	NIMS-KMA	Republic of Korea
7	CCCR-IITM	India	18	IPSL	France	29	NOAA-GFDL	USA
8	CMCC	Italy	19	MESSY-Cons	Germany	30	NUIST	China
9	CNRM	France	20	MIROC	Japan	31	TaiESM	Taiwan, China
10	CSIR-CSIRO	South Africa	21	MOHC	UK	32	THU	China
11	CSIRO-BOM	Australia	22	MPI-M	Germany	33	Seoul Nat.Uni	Republic of Korea

DOWNSCALING

Quantile Mapping

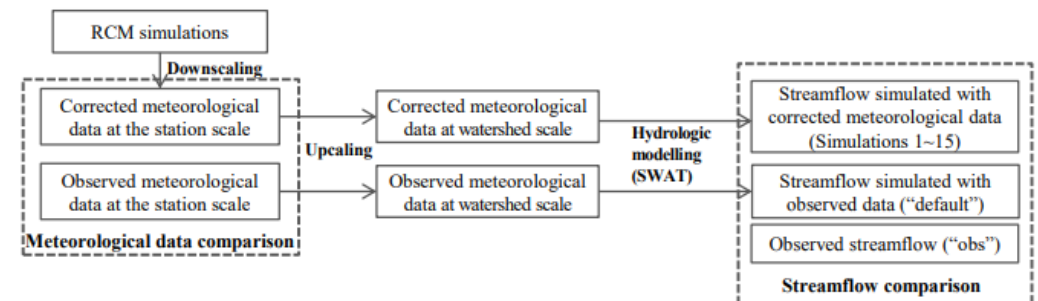
El método QM es un método de corrección de sesgo no paramétrico y es de aplicación general para todas las posibles distribuciones de precipitación sin ningún supuesto sobre la distribución de la precipitación, puede corregir efectivamente el sesgo en la media, la desviación estándar y la frecuencia de los días húmedos también como cuantiles.

Linear Scaling

El método LS tiene como objetivo coincidir perfectamente con la media mensual de valores corregidos con los observados (Lenderink et al., 2007). Opera con valores de corrección mensual basados en las diferencias entre los datos observados y sin procesar. La temperatura generalmente se corrige con un aditivo plazo sobre una base mensual.

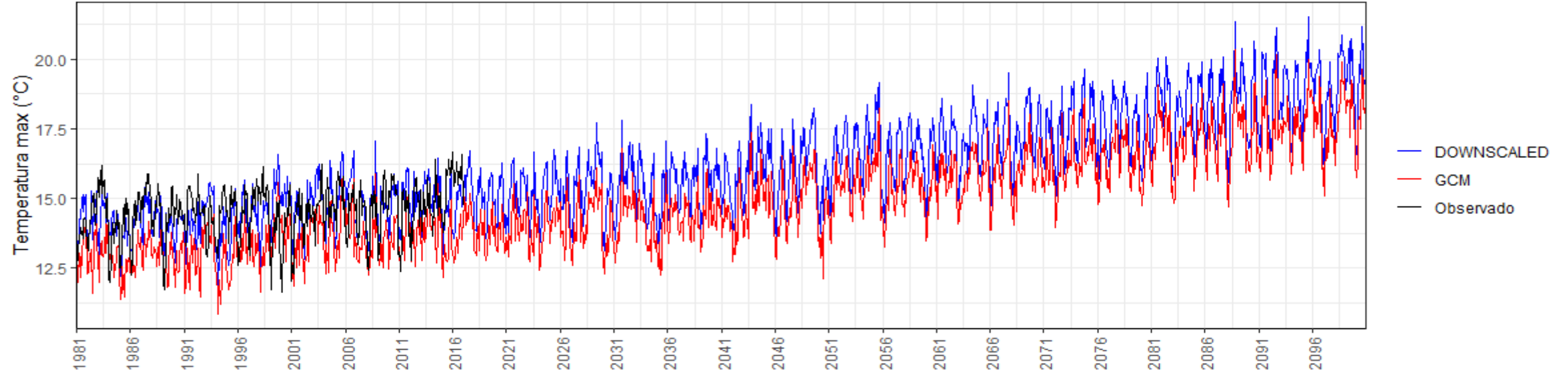
Bias correction for precipitation	Bias correction for temperature
Linear scaling (LS)	Linear scaling (LS)
Local intensity scaling (LOCI)	Variance scaling (VARI)
Power transformation (PT)	Distribution mapping for temperature using Gaussian distribution (DM)
Distribution mapping for precipitation using gamma distribution (DM)	
Quantile mapping (QM)	

Fuente: Fang et al,2015

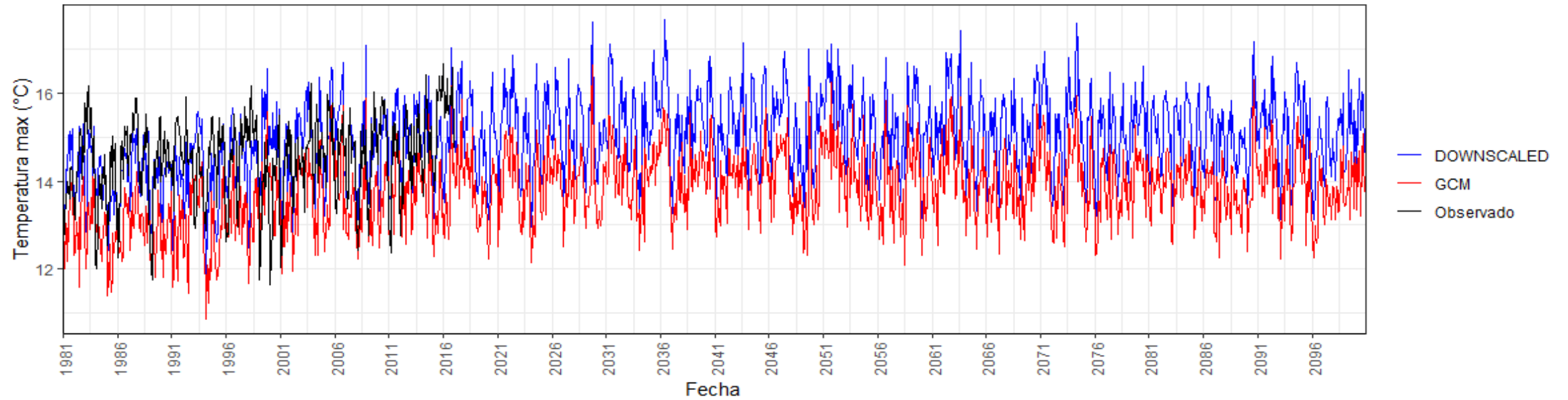


Fuente: Fang et al,2015

Temperatura maxima ssp585



Temperatura maxima ssp119



LIBRERIAS DE R APLICADOS A DIVERSOS CAMPOS

Meteo &
Clima

Hidrología

Teledetección

Extras

Cutoff

AirGR

Rstoolbox

Lubridate

Climdex

SWATplusR

raster

lattice

SPEI

Baseflow

sf

zoo

hydroTSM

HydroPSO

sen2cor

xts

hydropclim

Topmodel

RGEE

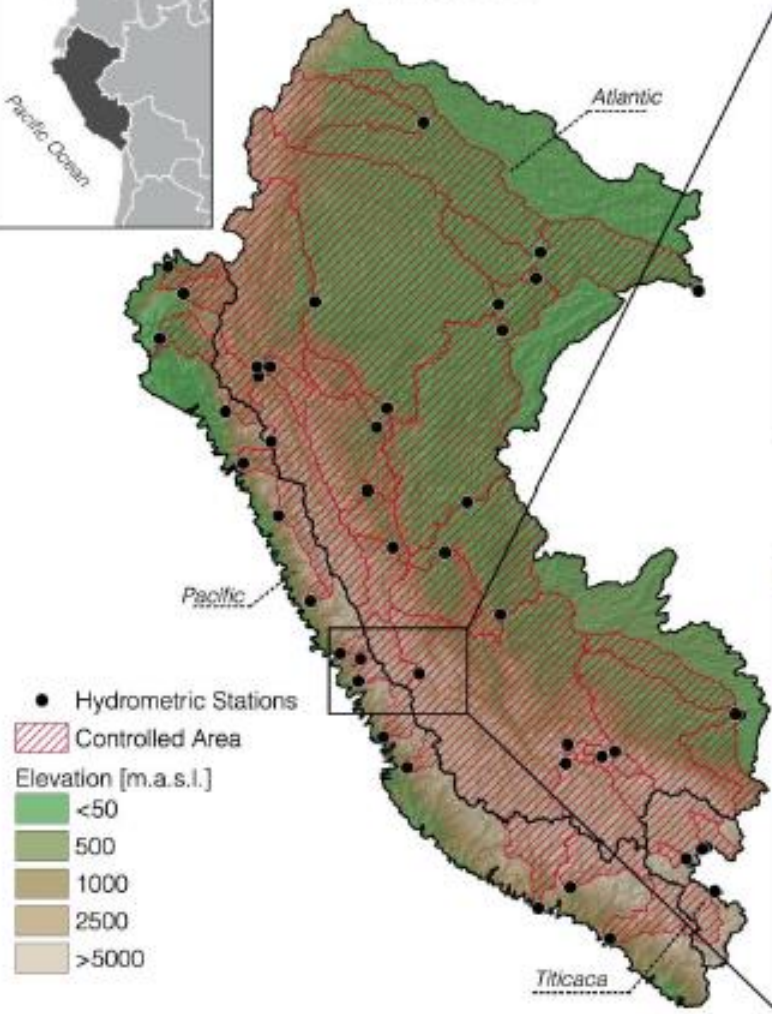
RPISCO

Dominio de Estudio

Área total:
1'480,620 km²



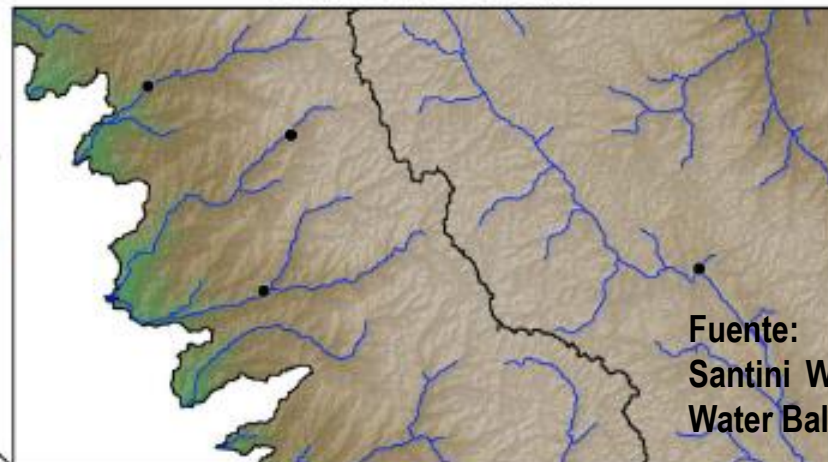
Study Area



Detail of subbasins (mean area: 410 km²)



Detail of river streams



03 Vertientes Hidrográficas

Atlántico (78.7%)

Pacífico (18.2%)

Titicaca (3.1%)

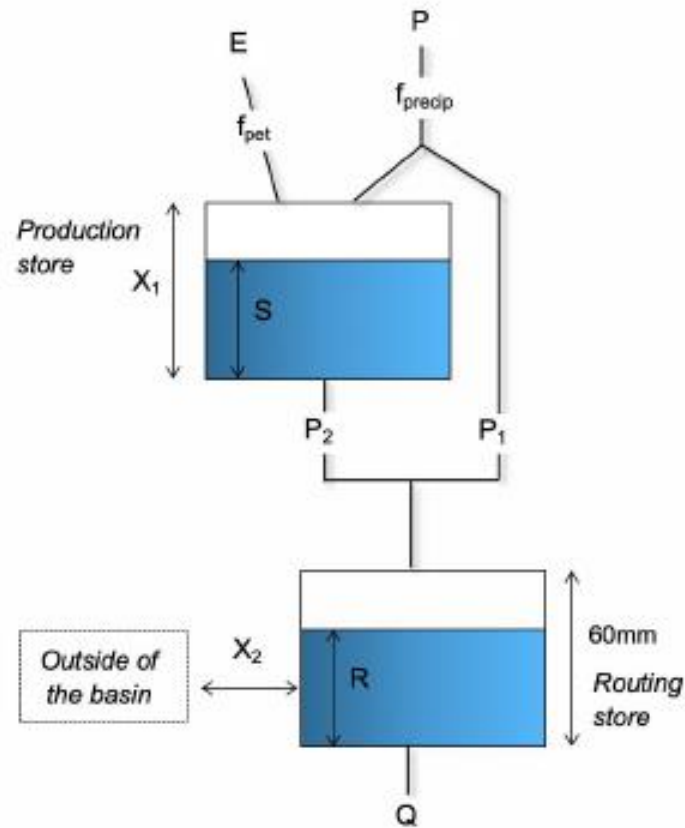
3594

subcuencas y tramos

Fuente: Llauca H, Lavado-Casimiro W, Montesinos C, Santini W, Rau P. PISCO_HyM_GR2M: A Model of Monthly Water Balance in Peru (1981–2020). Water. 2021; 13(8):1048.

Modelo Hidrológico Conceptual – GR2M

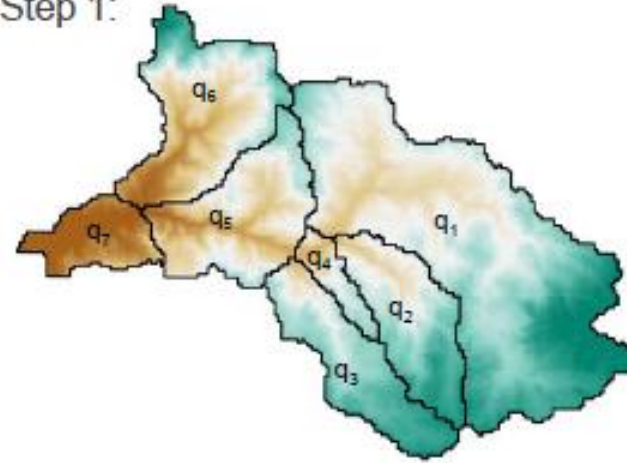
Estructura de GR2M



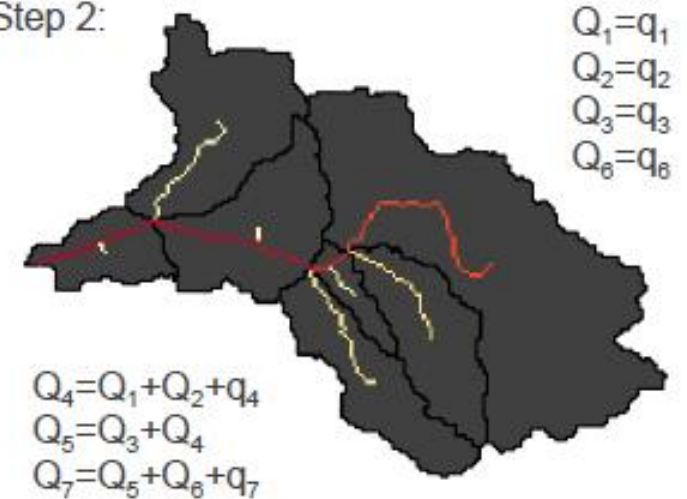
Mouelhi et al. 2006 (JH)

Adaptación semidistribuida:

Step 1:



Step 2:



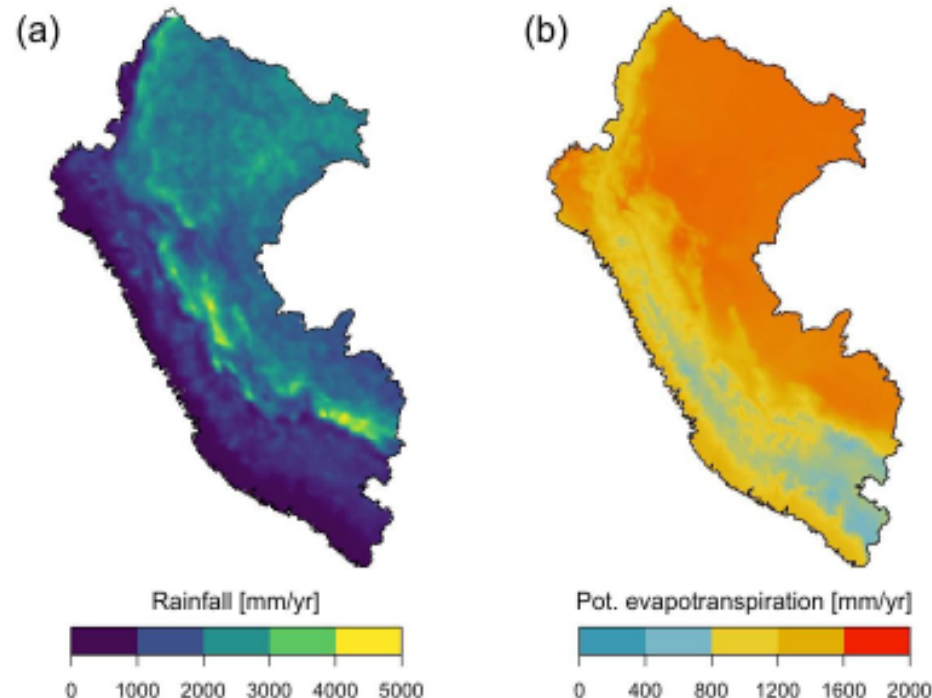
DEM \Rightarrow Flow Direction \Rightarrow Weighted Flow Accumulation

<https://github.com/hllauca/GR2MSemiDistr>

Datos de Entrada

Dataset PISCO

Producto grillado ($0.10^\circ \times 0.10^\circ$) de P, T y PET

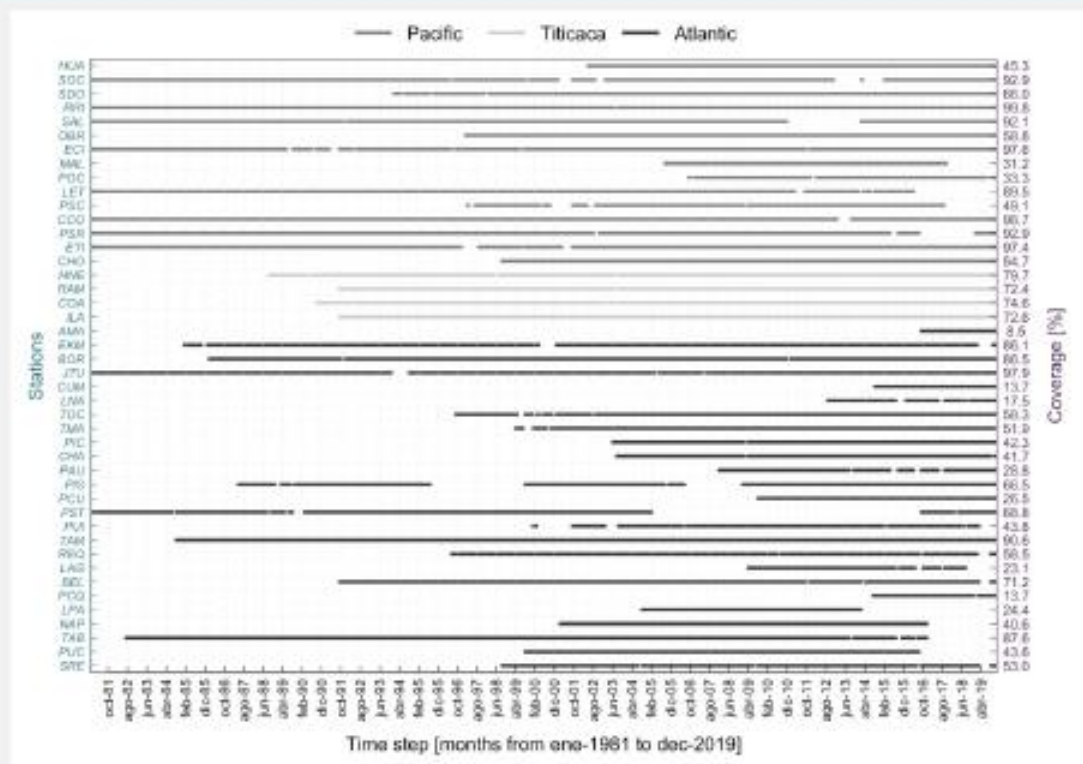


- ❑ Periodo: Enero 1981 – Marzo 2020
- ❑ PET por Hargreaves-Samani.
- ❑ Climatología mensual de PET (operatividad).
- ❑ Series sintéticas de P y PET para el periodo 1978-1980 (Warm-Up)
- ❑ Cálculo de promedios areales para 3954 subcuencas.

Aybar et al. 2019 (HSJ)

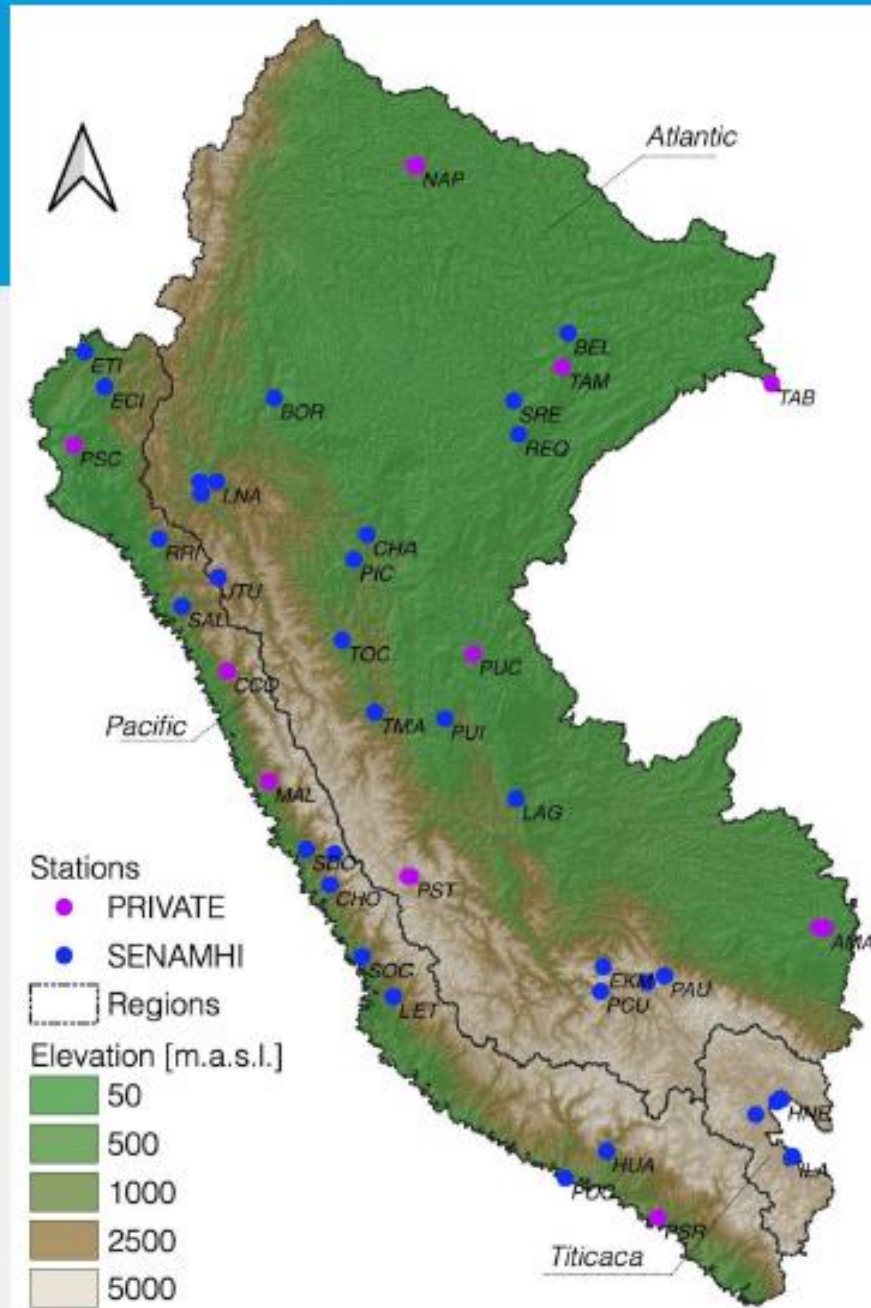
Fuente: Llauca H, Lavado-Casimiro W, Montesinos C, Santini W, Rau P. PISCO_HyM_GR2M: A Model of Monthly Water Balance in Peru (1981–2020). Water. 2021; 13(8):1048.

Datos de Entrada (2)



43 estaciones
hidrométricas

Enero 1981 – Marzo 2020



KGE

Función Objetivo

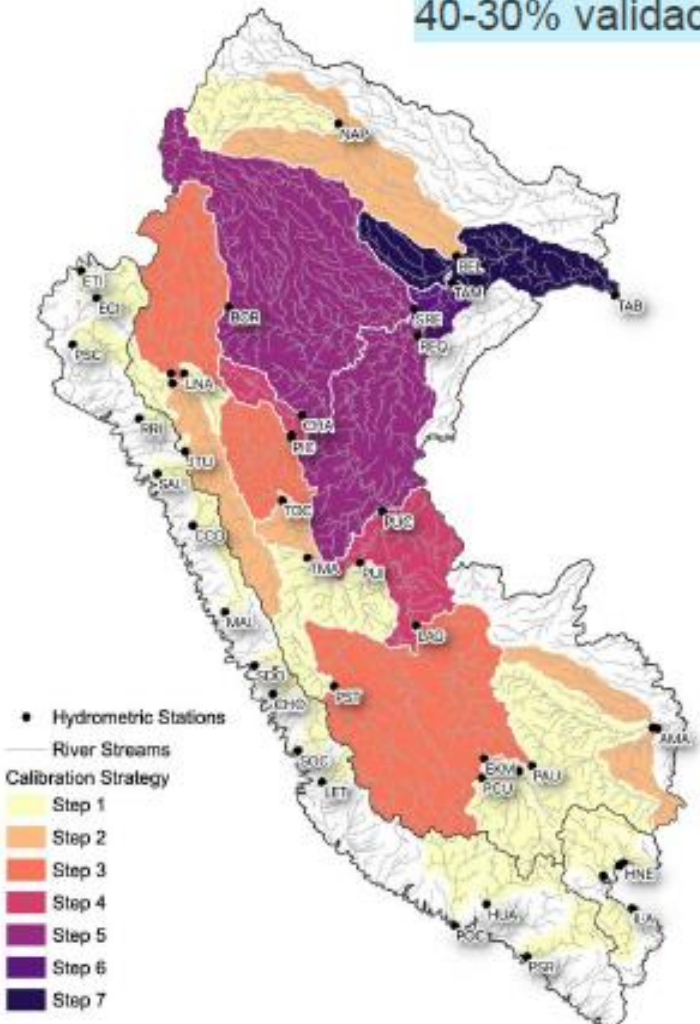
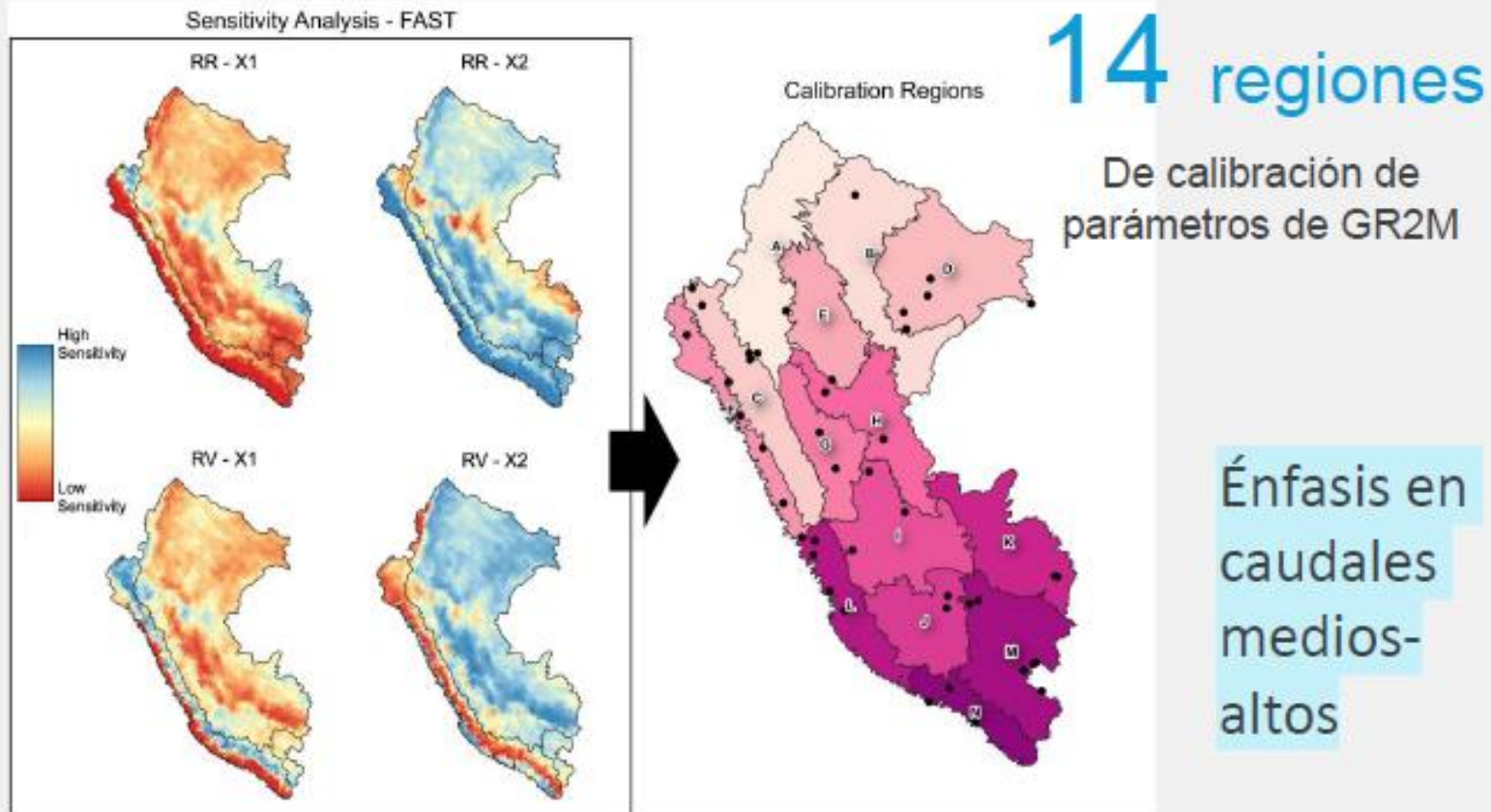
60-70% calibración

40-30% validación

Regionalización & Calibración

Análisis de Sensibilidad de Índices Hidroclimáticos +
Análisis Cluster

Similar a Bock et al. 2016 (HESS)

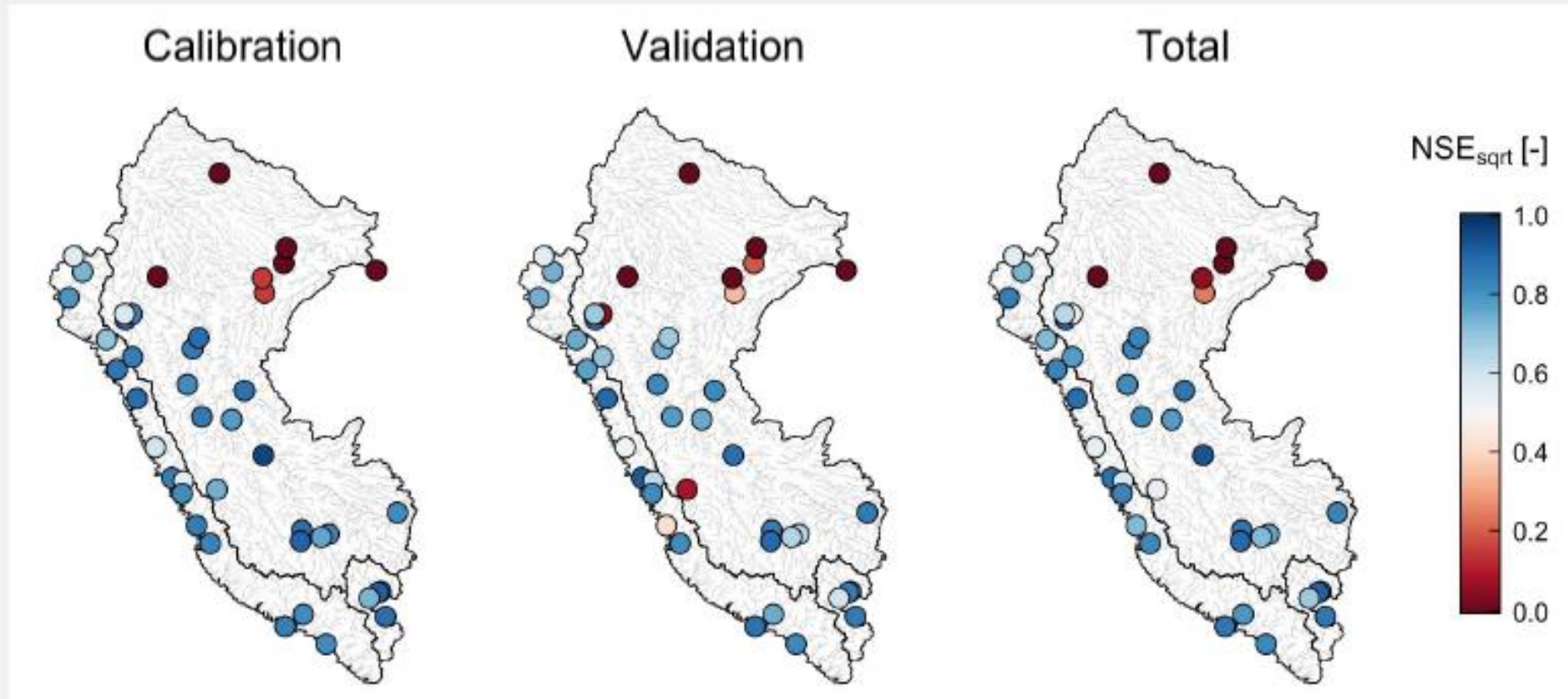


$$NSE_{sqrt} = 1 - \frac{\sum_{i=1}^n (\sqrt{Q_{sim,i}} - \sqrt{Q_{obs,i}})^2}{\sum_{i=1}^n (\sqrt{Q_{obs,i}} - \sqrt{Q_{obs,i}})^2}$$

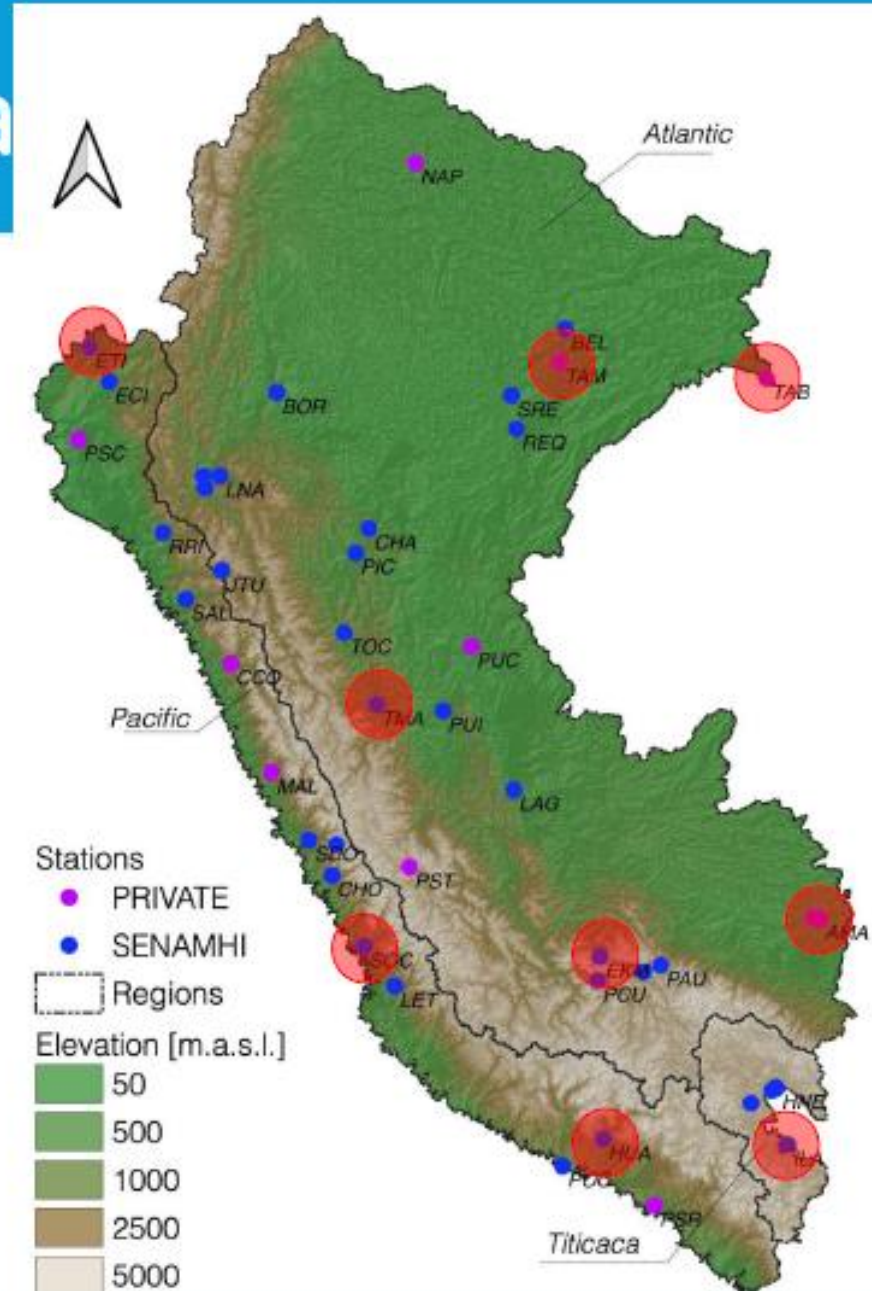
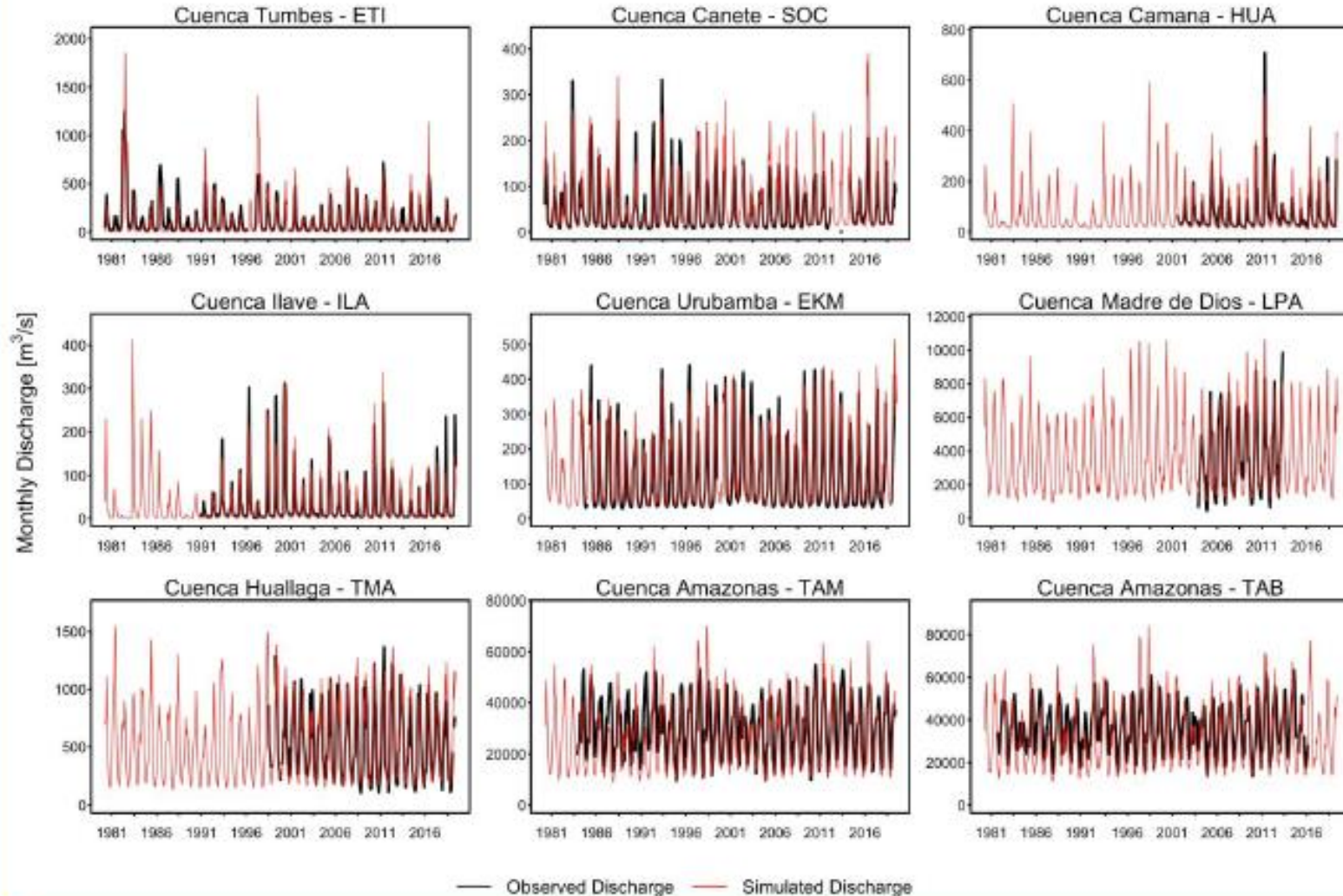
Similar a Bai et al. 2015(JH)

Desempeño del Modelo (2)

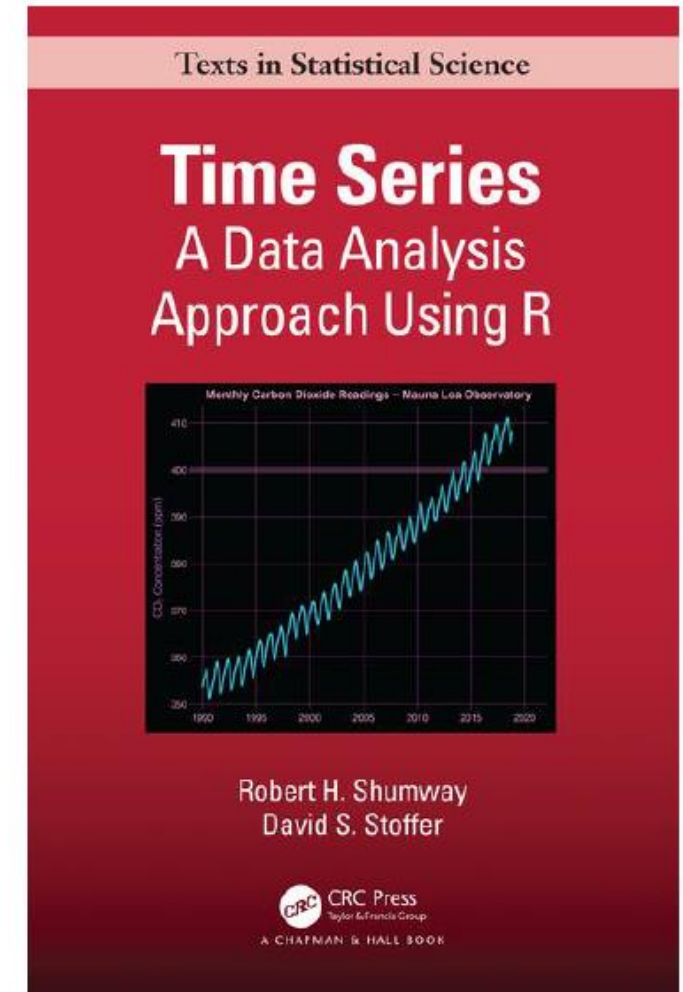
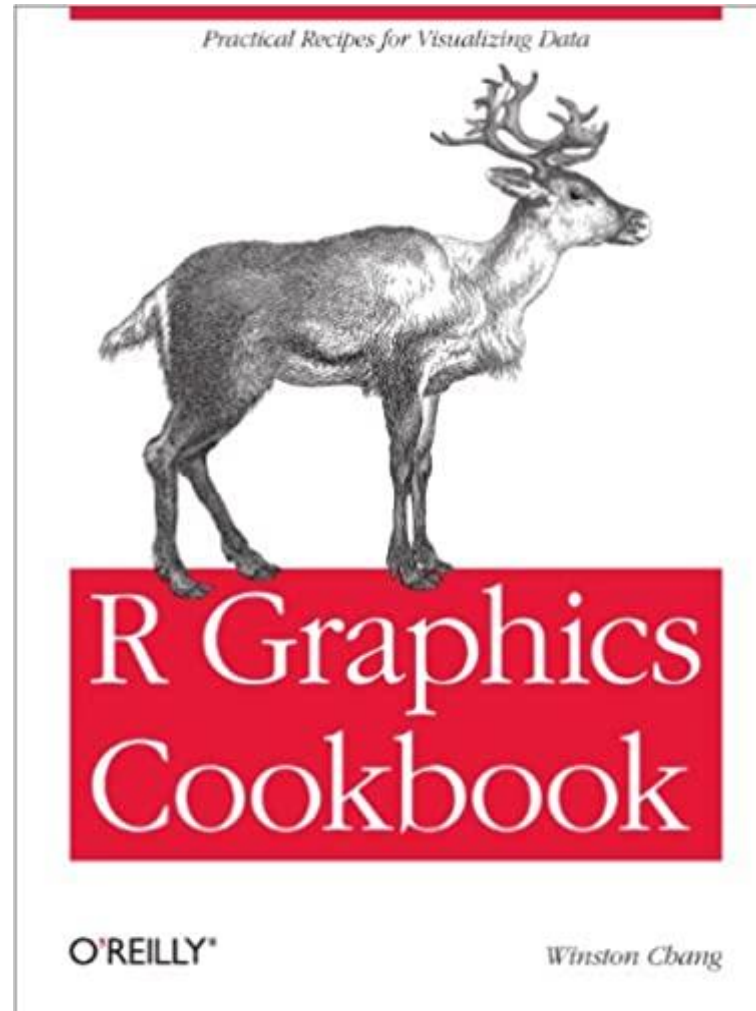
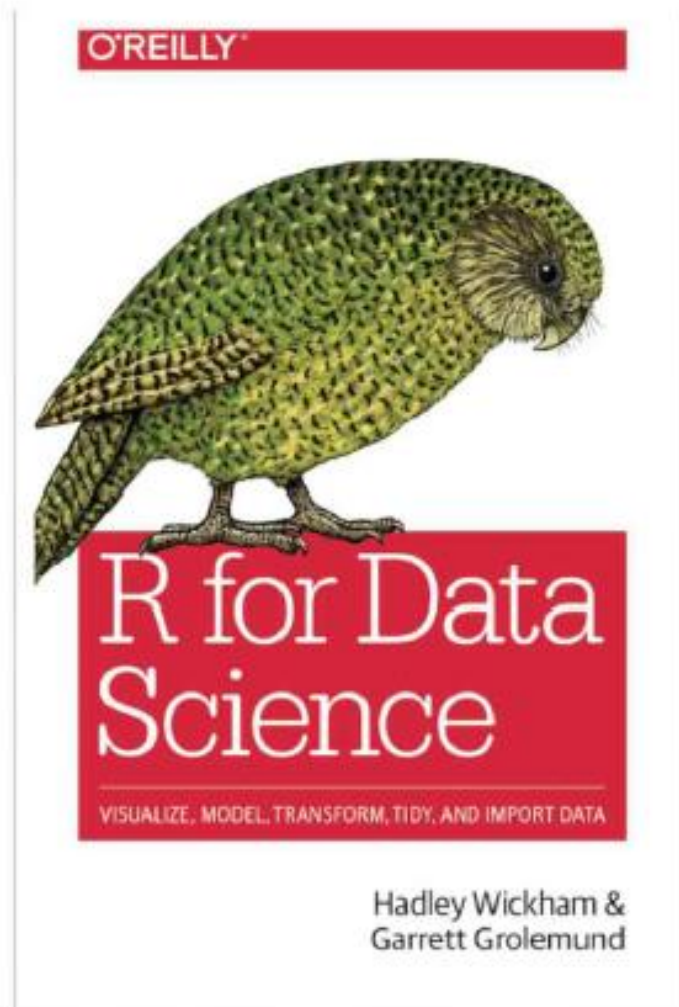
Caudales medios-bajos:



Simulación de caudales mensuales

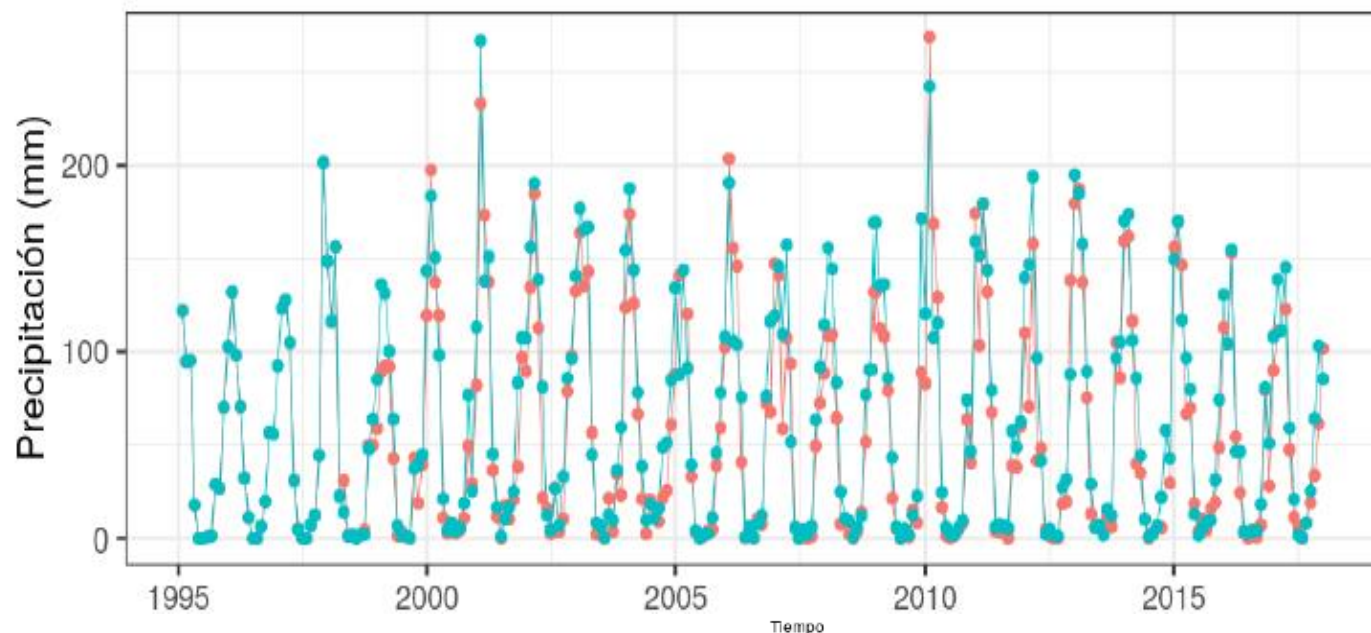


BIBLIOGRAFIA RECOMENDADA

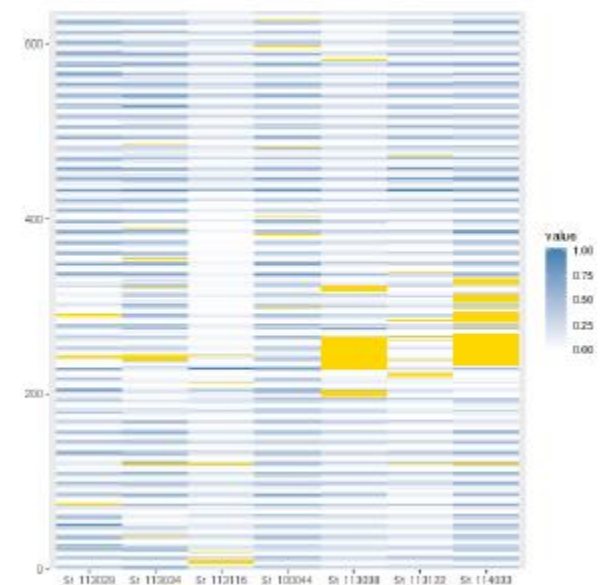


¿PREGUNTAS COMUNES?


- ¿Puedo completar información con PISCO?
- ¿Puedo determinar Pp max en 24h con PISCO?
- ¿Se puede utilizar PISCO para Instrumentos de gestión ambiental?
- ¿Se puede utilizar PISCO para estudiar el cambio climático?
- ¿Se puede hacer modelización hidrológica con PISCO?



series
● observado
● reconstruido



DESCARGA DE PISCO

 Biblioteca de Datos
SENAMHI HSR PISCO

Language
español ▼

Descripción | Modo Experto

← →

sirve desde [IRI/LDEO Climate Data Library](#)

SOURCES SENAMHI HSR PISCO

SENAMHI HSR PISCO

SENAMHI HSR PISCO: Peruvian Interpolated data.

Documentos

[overview](#) an outline showing sub-datasets of this dataset

Bases de Datos y Variables

[PET](#) Potential Evapotranspiration.






[Prec](#) Precipitation.

[Streamflow](#) SENAMHI HSR PISCO Streamflow[v1]


[Temp](#) Temperature.


Última Actualización: *Thu, 23 Jul 2020 17:30:54 GMT*

Compartir

     Like 0

Contáctenos



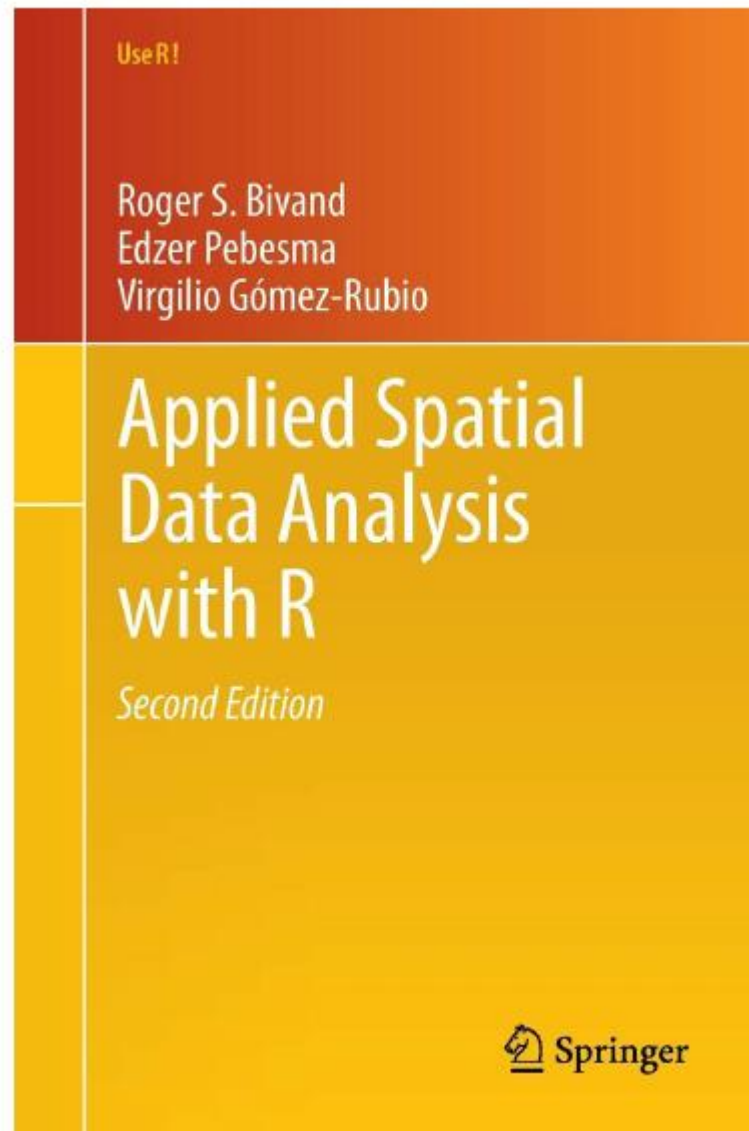
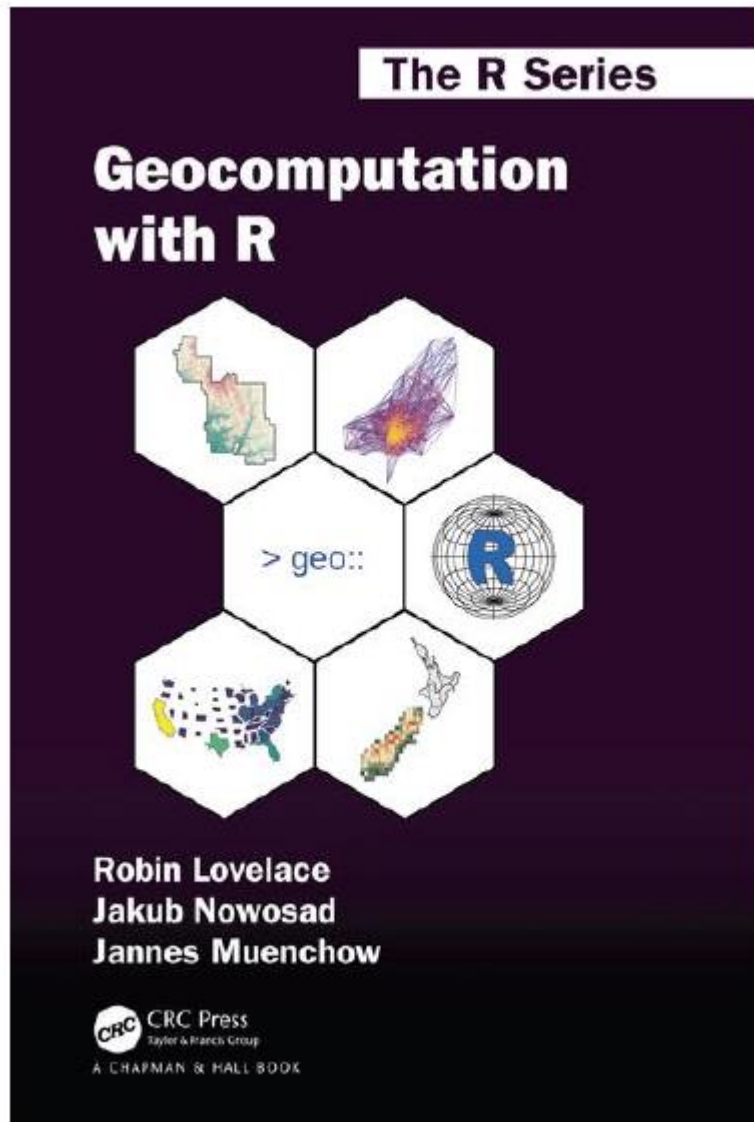


<https://iridl.ldeo.columbia.edu/SOURCES/.SENAMHI/>

¿QUE SIGUE?



RECOMENDACIONES FINALES



PROFE NO ENTENDI

❖ BIBLIOGRAFIA

- Aybar Camacho, C., Fernandez, C., Huerta, A., Lavado, W., Vega, F., & Obando, F. (2019). Construction of a high-resolution gridded rainfall dataset for Perú from 1981 to the present day. Hydrological Sciences Journal.
- Llauca H, Lavado-Casimiro W, Montesinos C, Santini W, Rau P. PISCO_HyM_GR2M: A Model of Monthly Water Balance in Peru (1981–2020). Water. 2021; 13(8):1048

***Muchas gracias por su
atención***



garciatitomiguel@gmail.com



+51 952281280