

Homogeneización de series mensuales de temperatura y precipitación en Bolivia a partir de registros SENAMHI entre 1980 y 2015

Marko Andrade^{1,*}, Javier Terrazas², Rocío Guzmán²

¹Proyecto de Hidroclimatología, Departamento de Física, Facultad de Ciencias y Tecnología Universidad Mayor de San Simón, Cochabamba, Bolivia

²Proyecto de Variables Climáticas, Departamento de Física, Facultad de Ciencias y Tecnología Universidad Mayor de San Simón, Cochabamba, Bolivia

*m.andrade@umss.edu.bo

Resumen

Existe consenso acerca de la importancia de realizar un control de calidad a las series de datos meteorológicos antes de utilizarlos en estudios climáticos de una región, orientando principalmente los esfuerzos a completar registros faltantes y a analizar la homogeneidad de las series históricas. Entre las distintas técnicas estadísticas que permiten el relleno de datos faltantes en series climáticas, los algoritmos basados en inteligencia artificial, como ser los métodos de redes neuronales, han demostrado ser bastante eficientes. Acerca de las inhomogeneidades en registros climáticos, se cuenta también con diferentes métodos numéricos para identificarlas y corregirlas, destacando los métodos basados en homogeneidad relativa. En Bolivia las series climáticas, en su mayor parte provistas por SENAMHI, no son la excepción y requieren desarrollar continuamente nuevos y mejores procedimientos para completar datos faltantes y homogeneizar series. En este trabajo se ha realizado el control de calidad a 1200 series mensuales de precipitación pluvial y temperatura de las estaciones climáticas SENAMHI para el periodo de 1980 a 2015. La metodología de relleno consistió en filtrar las series con menos de 30 % de datos faltantes, rellenarlas utilizando algoritmos de inteligencia artificial basados en redes neuronales, verificar para cada serie rellena el error del ajuste y escoger en cada caso la serie rellena con el menor error. Luego se procedió a evaluar la homogeneidad de cada serie. Se usó un método de homogeneidad relativa implementado en lenguaje R y se descartaron aquellas series con homogeneidad baja. Como resultado se obtuvieron 91 estaciones climáticas para precipitación pluvial y 53 estaciones para temperatura que pasaron el control de calidad. El resultado es consistente con estudios previos y muestra que del total de estaciones climáticas que se tiene en el país, menos del 10 % cuentan con información utilizable para estudios climatológicos que comprendan más de 30 años.

Palabras clave: Control de calidad, Series climáticas de precipitación y temperatura, Redes neuronales, Homogeneización, Climatología de Bolivia.

1. Introducción

Toda serie de datos climáticos proveniente de estaciones de monitoreo debe pasar por una revisión o control de calidad antes de ser utilizada en estudios concretos. Por lo general los esfuerzos se orientan a analizar la homogeneidad de las series históricas de modo que cualquier conclusión basada en tales registros sea suficientemente representativa del comportamiento real del clima (Peterson et al., 1998). Algunas de las causas más comunes de inhomogeneidades en los registros climáticos son los cambios en la instrumentación, en el observador, en la ubicación y exposición de la estación de monitoreo, y en los procedimientos observacionales (Wang, 2008a). Estos cambios artificiales en las series de datos se conocen como puntos de quiebre o puntos de cambio; se han desarrollado distintos algoritmos para identificarlos,

distinguirlos de los cambios reales y, eventualmente, corregirlos. Un segundo problema, también bastante frecuente en el control de calidad de series climáticas, es el referido a la cantidad de datos faltantes. En algunos casos el porcentaje de faltantes es tan alto que la serie debe ser desechada ya que no podría justificarse estadísticamente ningún proceso de relleno artificial. En el caso de Bolivia, las series de precipitación pluvial y de temperatura (máxima y mínima) son provistas por el Servicio Nacional de Meteorología e Hidrología (SENAMHI) que cuenta con estaciones de monitoreo en distintas zonas del país. Se observa que la cantidad de estaciones es baja para la extensión territorial de Bolivia, y además, los registros presentan notorias discontinuidades (datos faltantes) al margen de las inhomogeneidades propias de cualquier serie climática.

Si bien existen reportes previos sobre control de calidad en series climáticas de SENAMHI-Bolivia, se propone en este trabajo una metodología alternativa para el relleno y homogeneización de series mensuales de precipitación pluvial y temperatura. Se ha escogido como periodo de estudio 1980 a 2015 y seleccionado las estaciones de monitoreo SENAMHI que cumplen con la condición de ausentes con un porcentaje menor al 30%. El relleno se realizó utilizando algoritmos de redes neuronales en los cuales cada serie es completada con ayuda de estaciones próximas (con las que tiene buena correlación) y acudiendo también, en algunos casos, a series de datos ajenas a la base de SENAMHI (fuentes externas) con las que la serie a ser completada tiene buena correlación. Se ha podido evidenciar que los métodos de redes neuronales son más eficientes que los métodos estadísticos convencionales para el relleno de este tipo de series temporales. En cuanto a la homogeneización de las series, existen reportes de comparación entre distintos algoritmos utilizados en la actualidad (Venema, 2012) en los que se contemplan los softwares más utilizados por la comunidad en la homogeneización de series mensuales de precipitación y temperatura. La mayor parte de los algoritmos se basan en aproximaciones de homogeneización relativa, la cual asume que estaciones cercanas están expuestas a condiciones climáticas parecidas y que cualquier diferencia entre estaciones cercanas puede utilizarse para detectar inhomogeneidades. Algunos de los paquetes de homogeneización más recientes son (Guijarro, 2017): Climtol 3.0 (Guijarro, 2018), ACMANT 3.0 (Domonkos, 2014), (Domonkos and Coll, 2017), MASH 3.03 (Szentimrey, 2007), RH-testsV4 (Wang & Feng, 2013b), USHCN v52d (Menne & Williams, 2005), HOMER 2.6 (Mestre, 2013). En este trabajo se ha utilizado RHtestsV4, el cual consiste una paquetería en R (R Core Team, 2019) con la cual se puede detectar y corregir múltiples puntos de cambio que podrían estar presentes en series climáticas con errores autorregresivos de primer orden. Las series pueden ser de precipitación, temperatura o cualquier otra variable climática, y de tipo diario, mensual o anual (se excluye las series de precipitación diaria por no tener un comportamiento gaussiano, requiriendo una variante de RHtests que en este trabajo no fue necesario). RHtestsV4 se basa en el test t de máxima penalidad (Wang, 2007) y en el test F de máxima penalidad (Wang, 2008b), incluidos ambos en un algoritmo de testeo recursivo. Dependiendo de si la

serie temporal a homogeneizar (serie base) cuenta o no con una serie de referencia (serie ya homogénea y con alta correlación con la serie base), se utiliza el test t o el test F (test t si se tiene serie de referencia y test F si no se tiene serie de referencia). En las secciones subsiguientes se explica con más detalle tanto el procedimiento de relleno como el de homogeneización aplicado a las series climáticas seleccionadas para Bolivia. El trabajo concluye con la presentación de las series mensuales de precipitación y temperatura completas y homogéneas, almacenadas en un repositorio virtual de acceso libre para toda la comunidad.

2. Métodos

En este estudio se ha utilizado la base de datos SISMET del Servicio Nacional de Meteorología e Hidrología de Bolivia (SENAMHI, 2020) para el periodo 1980 a 2015. Si bien la base cuenta con información climática de más de 1200 estaciones, no todas cubren el rango requerido en el trabajo. Se seleccionaron, de inicio, 190 estaciones para precipitación y 92 para temperatura máxima y mínima según el criterio de cubrir el período de estudio con un porcentaje de datos faltantes menor a 30 %. Las Figs. 1 y 2 muestran la ubicación geográfica de las estaciones, tanto para precipitación como para temperatura. Se observa una mayor densidad de estaciones en la zona central del país en comparación con el noreste y el sudoeste que están poco monitoreados. Para fines de comparación, se han considerado también datos provenientes de fuentes externas. Una de esas fuentes proviene de la Unidad de Investigación Climática (*Climatic Research Unity*, CRU) de la Universidad East Anglia (Harris, I. et al., 2014). Los datos CRU han sido generados por interpolación de datos disponibles de estaciones meteorológicas de todo el mundo, logrando una resolución de 0.5° latitud/longitud. Si bien esta resolución no es suficiente para, por ejemplo, estudios hidrológicos, ya permite tener una idea de la climatología del país y sirve de base en este trabajo para fines de comparación. Las Figs. 3 y 4 muestran la climatología de Bolivia para el período 1980 a 2015 en base a medias mensuales de temperatura (°C) y precipitación (mm/día) provistas por CRU. En el caso de la temperatura, es notorio el gradiente altitudinal al pasar de la zona oriental a las tierras altas del occidente. En cuanto a la precipitación, destaca la zona del Chapare en el trópico de Cochabamba con los valores más altos (vea la mancha verde intensa en la región

central del país para la precipitación DJF en la Fig.4), así como la baja precipitación en el altiplano. Una segunda fuente de información consultada en este trabajo fue la plataforma PCA - Princeton Climate Analytics (PCA, 2019). Se trata de un servicio de información climática de la Universidad de Princeton que combina datos satelitales con observaciones en tierra y modelos de reanálisis para proveer series hidroclimatológicas con resolución espacial de 0.25° y resoluciones temporales diarias, mensuales y anuales. Para el caso de Bolivia, PCA ofrece el Monitor de Inundaciones y Sequías para Latinoamérica, de donde fueron extraídas series temporales mensuales de precipitación, temperatura máxima diaria, temperatura mínima diaria, radiación solar, escorrentía y evaporación, para distintas locaciones coincidentes con las de las estaciones climáticas del país. La Fig. 5 muestra la serie temporal de precipitación extraída de dicha plataforma para la ciudad de Cochabamba.

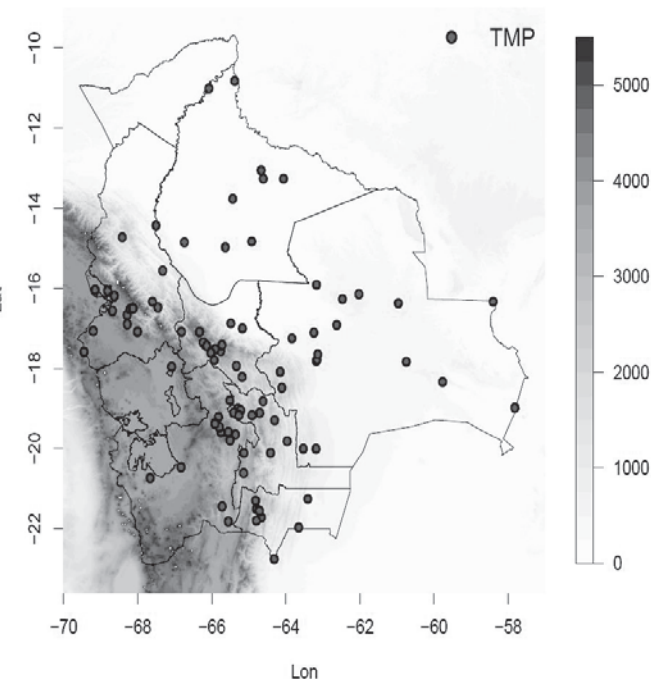


Figura 2. Estaciones seleccionadas con datos de temperatura: 92 en toda Bolivia antes de la homogeneización

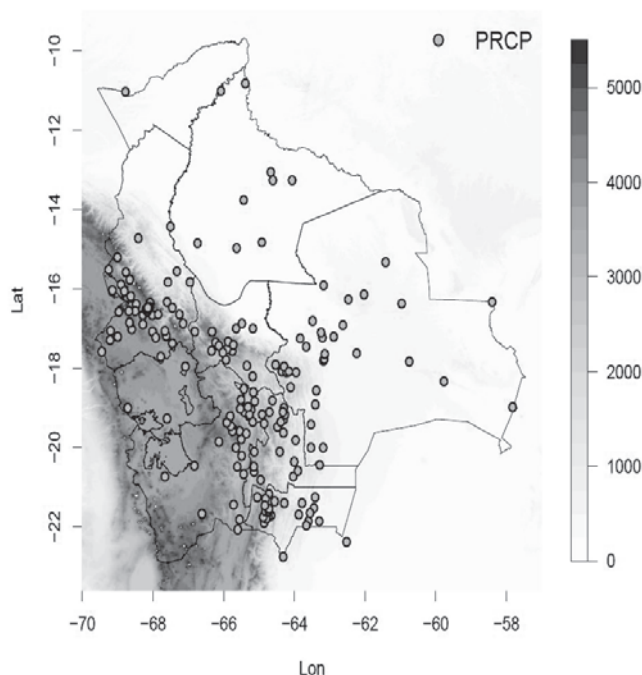


Figura 1. Estaciones seleccionadas con datos de precipitación: 190 en toda Bolivia antes de la homogeneización

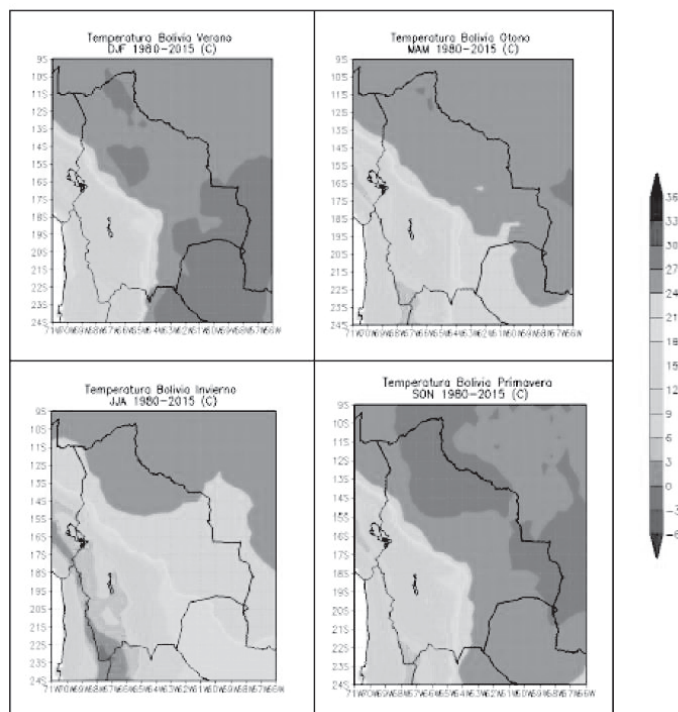


Figura 3. Temperatura promedio estacional Bolivia. Período 1980-2015 en base a CRU TS v.4.02

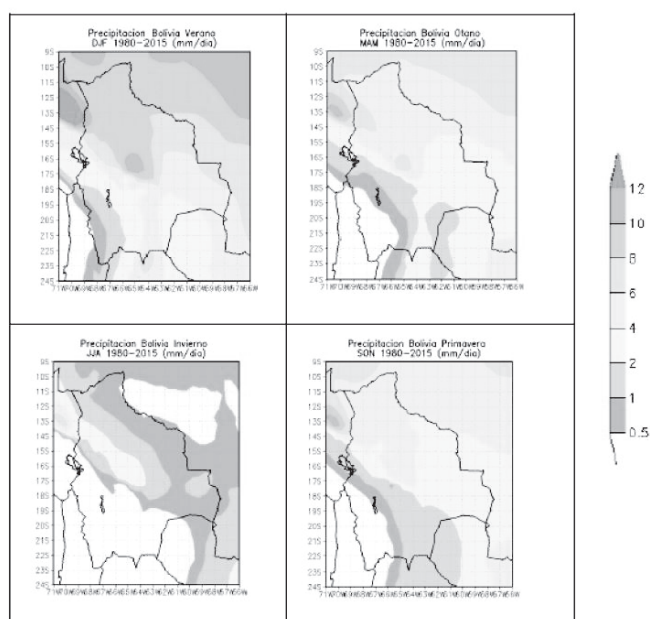


Figura 4. Precipitación promedio estacional Bolivia. Período 1980-2015 en base a CRU TS v.4.02

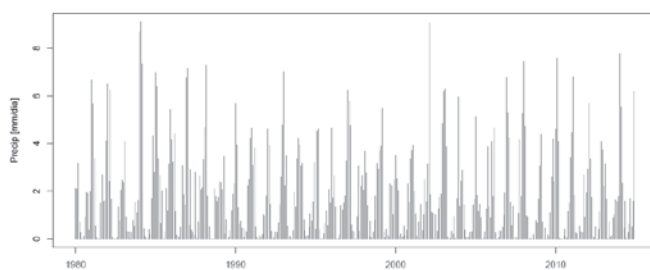


Figura 5. Serie temporal de precipitación mensual en mm/día para la ciudad de Cochabamba según la PCA. Período 1980 a 2015

2 Herramientas para el control de calidad de series climáticas

El software utilizado para la homogeneización fue RHTestsV4 (Wang, 2007). Consiste en una paquetería escrita en lenguaje R, muy potente, de uso libre y multiplataforma. Detecta y ajusta múltiples puntos de cambio (desplazamientos) que podrían existir en una serie de datos que pueden tener errores autorregresivos de primer orden. RHtestsV4 se basa en la prueba t de máxima penalidad (Wang, 2007) y en la prueba F de máxima penalidad (Wang, 2008b), incluidas ambas en un algoritmo de testeo recursivo. Dependiendo de si la serie temporal a homogeneizar (serie base) cuenta o no con una serie de referencia (serie ya homogénea y con

alta correlación con la serie base), se utiliza la prueba t o la prueba F (test t si se tiene serie de referencia y test F si no se tiene serie de referencia). En el *caso sin serie de referencia*, la prueba F de máxima penalidad permite que las series de tiempo tengan una tendencia lineal durante todo el período del registro de datos, es decir, sin cambio en el componente de la tendencia. Los resultados sin el uso de una serie de referencia son, sin embargo, menos confiables que en el caso con referencia. Por este motivo, en el presente trabajo se trató de utilizar series de referencia tanto como fue posible.

En cuanto a las técnicas de relleno, se optó por implementar un algoritmo de **aprendizaje profundo** (DL o *deep learning*), que es un subcampo específico del aprendizaje automático (*machine learning*). Representa la idea de capas sucesivas de modo que la cantidad de capas que contribuyen a un modelo de datos se denomina profundidad del modelo (Chollet, F., 2018). Las redes neuronales con múltiples capas se denominan Redes Neuronales Profundas (DNN por sus siglas en inglés). Una de las librerías de aprendizaje profundo disponible para el desarrollo de redes neuronales profundas es Keras (Keras, 2019). Esta librería es una API de redes neuronales de alto nivel, escrita en Python y desarrollada con un enfoque que permite su implementación **rápida**. En la Fig. 6 se observa la composición de una red neuronal profunda (DNN), donde x_1 ; x_2 y x_3 son las características de entrada, seguidas de dos capas de neuronas ocultas, y finalmente la capa de salida z_1 y z_2 .

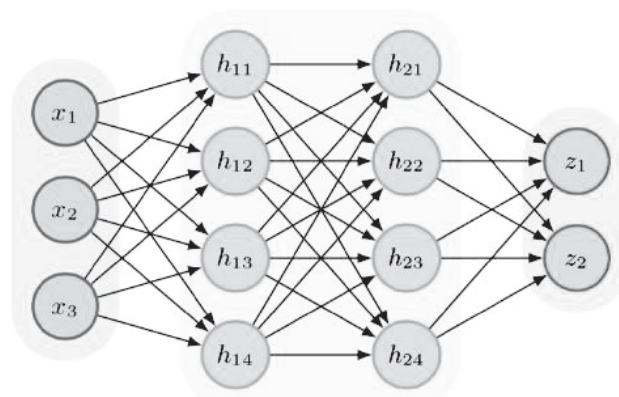


Figura 6. Estructura de una red neuronal profunda (DNN) de 2 capas ocultas.

Cada neurona contiene una función de activación y un valor límite. El valor límite es el mínimo valor que una

entrada debe tener para activar la neurona. Una función de activación está diseñada para el límite de salida de la neurona, usualmente para valores entre 0 a 1, o -1 a +1. Existen varios tipos de funciones de activación disponibles en DL (Moolayil, 2019); en este trabajo se han utilizado las funciones Sigmoid y ReLu con valores límite de 0 a 1, que se muestran en la Fig. 7.

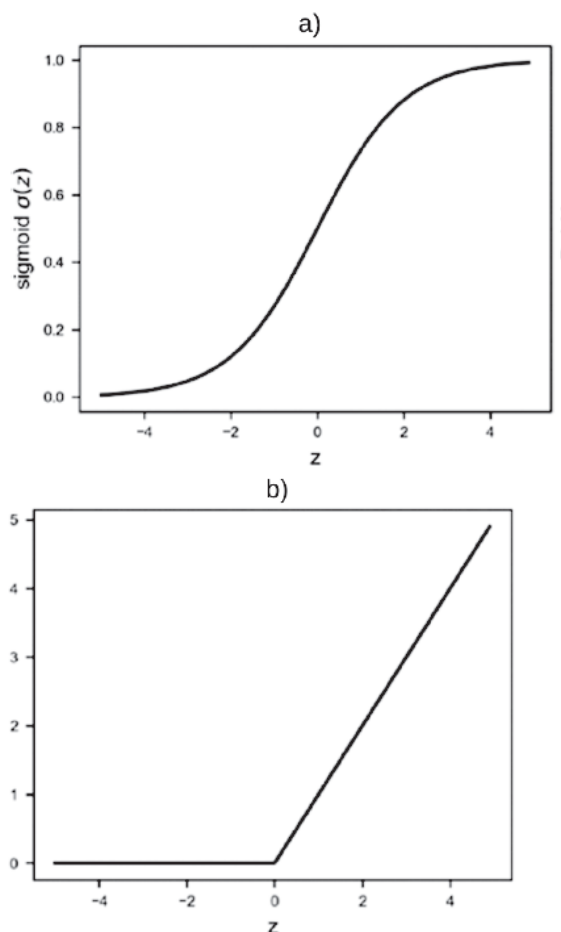


Figura 7. Funciones de activación. a) Sigmoid y b) ReLU (Moolayil, 2019) utilizadas en este trabajo.

El algoritmo DL utiliza una “función de costo” que es la función que se debe minimizar para obtener el mejor valor de cada parámetro en el modelo. Las funciones de costo más populares son el Error Cuadrático Medio (MSE) y el Error Absoluto Medio (MAE). Para reducir el costo se utilizan optimizadores, funciones que buscan el mejor valor de los parámetros. Su elección y uso adecuado constituye la parte más importante del llamado “entrenamiento” de la red neuronal. Básicamente una función optimizador es un algoritmo matemático que emplea reglas del cálculo diferencial

(derivadas, gradientes) para encontrar óptimos en la red a partir de variaciones en el peso relativo de cada neurona. En este trabajo se utilizaron los optimizadores de Descenso de Gradiente Estocástico (SGD, *Stochastic Gradient Descent*) y el de Estimación de Momento Adaptativo (AdAM, *Adaptive Moment Estimation*).

1.2 Relleno y homogeneización de las series

Se siguieron los siguientes pasos:

1. Descarga de datos crudos y preprocesamiento a un formato estándar.
2. Filtro de series en base a datos ausentes y puntos de cambio (sin referencia).
3. Homogeneización 1: Se realizó la agrupación para la homogeneización con y sin referencia por cada grupo.
4. Reconstrucción de series PRCP y TMP con redes neuronales profundas, DNN.
5. Homogeneización 2: Se repite el paso 3 pero ahora sin agrupación y se realizan ajustes finales.

Estos pasos se encuentran en el diagrama de flujo de la Figura 8.

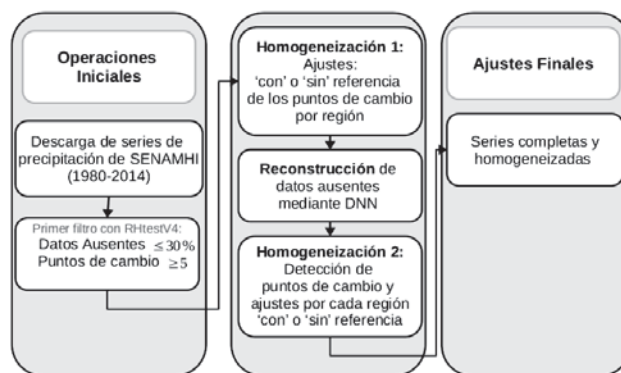


Figura 8. Secuencia metodológica implementada en el proceso de homogeneización y reconstrucción de las series climáticas de precipitación y temperatura

El filtro del paso 2 consistió en eliminar series con más de 30% de datos ausentes (sin importar su ubicación en la serie temporal) y eliminar series con más de 5 puntos de cambio. En este paso ya se usó RHTests bajo la modalidad “sin referencia”; por tanto, este filtro evaluó únicamente la consistencia interna de las series. A continuación, se procedió a agrupar estaciones tanto para PRCP como para TMP en función de su altura s.n.m. y su ubicación espacial. La agrupación permite determinar la significancia de los puntos de cambio. Se

obtuvieron 35 regiones para precipitación y 26 para temperatura. La Fig. 9 muestra la regionalización para la precipitación.

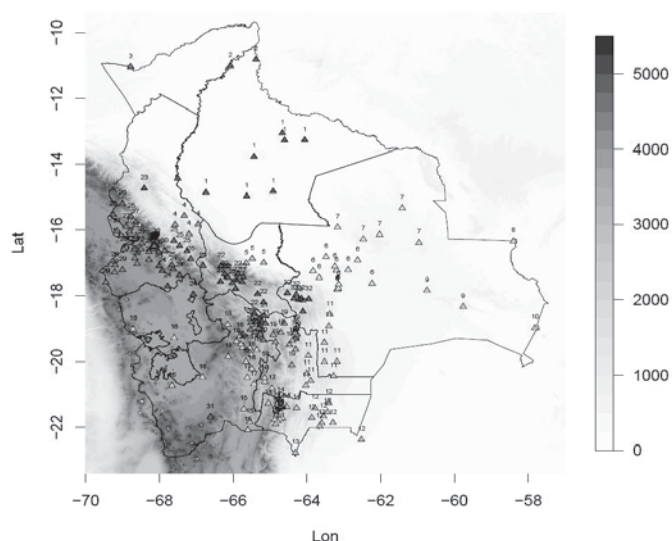


Figura 9. Regionalización del dominio de estudio: 35 regiones para precipitación.

El paso 4 (reconstrucción de series mediante DNN) es el más novedoso del trabajo. Se utilizaron funciones de activación Linear y ReLU según el diagrama de flujo de la Figura 10. Por un lado, se acumularon las estaciones completas y homogéneas (las mejores estaciones de partida) junto con las extraídas de la PCA (3 variables: precipitación, evaporación y escorrentía) para conformar los predictores del método DNN; por el otro lado, se ordenaron las series incompletas identificando como objetivo las variables a ser reconstruidas en cada una.

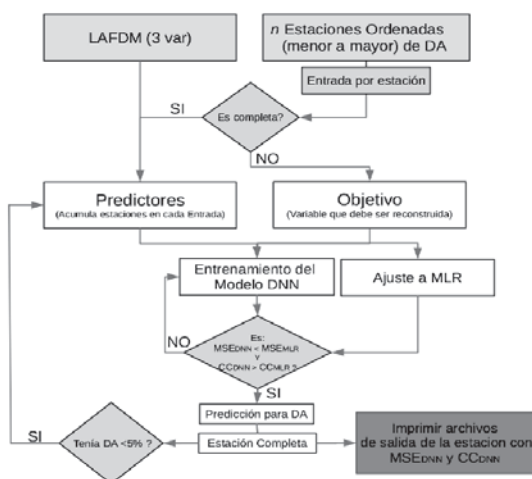


Figura 10. Diagrama de flujo para el relleno o reconstrucción de series de precipitación y temperatura mediante redes neuronales profundas (DNN).

Abreviaciones: LAFDM: Datos de la plataforma Princeton Analytics para Latinoamérica, DA: Datos ausentes, DNN: Redes neuronales profundas, CC: Coeficiente de correlación, MLR: Regresión lineal múltiple.

Se procedió a ajustar mediante regresión lineal múltiple (MLR) y mediante DNN cada una de las series objetivo, luego se calculó el Error Cuadrático Medio (MSE) y el Error Absoluto Medio (MAE) en ambos casos, junto con los coeficientes de correlación para poder distinguir si el ajuste por DNN es mejor que el MLR. La ventaja del método DNN es que el proceso de “entrenamiento” de las redes permite mejorar el ajuste hasta que, eventualmente, supera al MLR. Como se ve en el diagrama de flujo, el proceso de entrenamiento se repite hasta que el error del DNN sea menor que el del MLR.

La Fig. 11 muestra la estructura de la red neuronal profunda, DNN, que se utilizó en el trabajo. Se muestran las 4 capas escondidas, cada una con 27 neuronas y sus respectivas activaciones, y los datos de salida de la red ya entrenada.

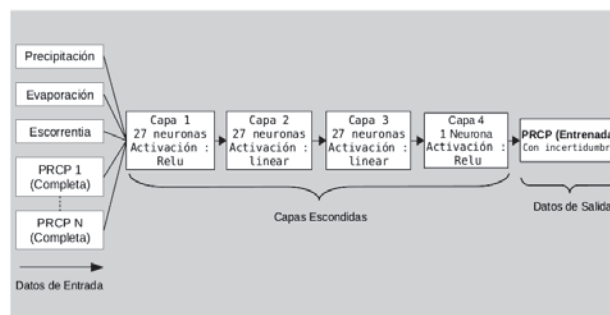


Figura 11. Estructura de la red neuronal para la reconstrucción de PRCP: Datos de entrada, 4 capas escondidas de 27 neuronas c/u, y datos de salida.

Una vez rellena la serie con el dato ajustado por DNN se pasa al siguiente dato ausente. El proceso se detiene, finalmente, cuando se ha relleno la totalidad de datos ausentes en la serie.

Como último paso (paso 5), se realiza una segunda corrida en RHTestsV4 para verificar que no se haya introducido puntos de cambio nuevos durante el relleno. Las series que pasan esta última prueba de homogeneización son, finalmente, consideradas completas y homogéneas.

3. Resultados del Proceso de Relleno y Homogeneización

Los resultados para cada estación fueron similares a los

que se muestra en las figuras siguientes. La Fig. 12 muestra el ajuste MLR para la estación Tiraque, en unidades normalizadas. Se destaca una parte de la serie para mostrar que el modelo contempla el dato de precipitación cero (valores negativos excluidos) como límite inferior. La Fig. 13 muestra la misma estación, pero ahora con el modelo ajustado por DNN. Si bien ambos modelos lucen similares, el cálculo del error permite distinguirlos. Para el caso mostrado, se obtuvo un MSE de 0.1243 en el MLR y de 0.1107 en el DNN. Esta mejora de alrededor de 1% a favor del DNN se observó en todas las series. Del mismo modo, la correlación del ajuste DNN (0.8717) fue superior que la del ajuste MLR (0.8531).

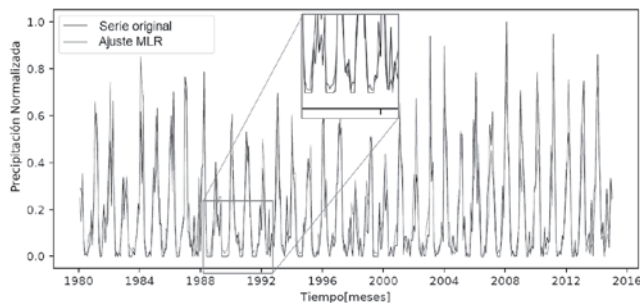


Figura 12. Modelo MLR para la estación TP3304 (Tiraque)-Precipitación. En negro los valores originales de la serie y en rojo los valores modelados.

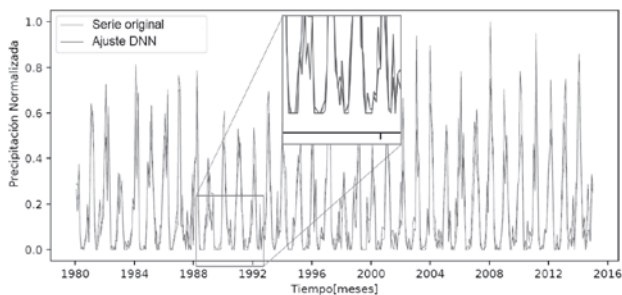


Figura 13. Modelo DNN para la estación TP3304 (Tiraque) - Precipitación. En azul los valores originales de la serie y en rojo los valores modelados.

Una vez obtenido un ajuste DNN superior al MLR, se utilizó tal modelo para rellenar los datos faltantes en cada una de las series. Por ejemplo, la Fig. 14 muestra el resultado para la estación Mizque, en mm de precipitación, para el rango de estudio. Para temperatura, máxima y mínima, el procedimiento arrojó resultados análogos. La Fig. 15 muestra el relleno de la serie de Temperatura Mínima en la estación La Tamborada.

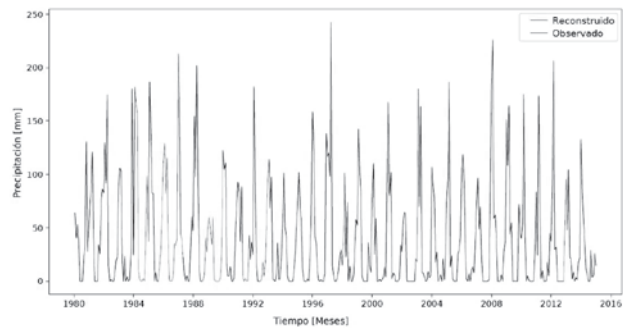


Figura 14. Estación TP2045 (Mizque), precipitación, reconstruida a partir del modelo DNN. En rojo los datos rellenados.

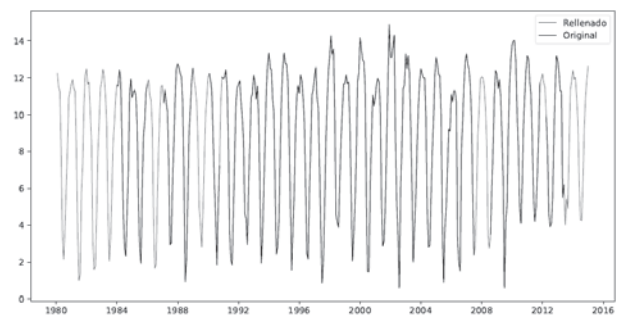


Figura 15. Estación TP2597 (La Tamborada), temperatura mínima, reconstruida a partir del modelo DNN. En rojo los datos rellenados.

Tras el relleno de series con DNN se obtuvieron 190 estaciones de precipitación y 93 de temperatura, previo a la última homogeneización. Se muestra en la Fig. 16 la salida del proceso de homogeneización en RHTestsV4 para precipitación y en la Fig. 17 para temperatura. El cálculo detecta los puntos de cambio (puntos donde se pierde la homogeneidad) en la serie. Los ejemplos de las Figs. 16 y 17 corresponden a series que fueron descartadas por tener uno o más puntos de cambio tras el relleno. Una vez revisadas todas las estaciones y filtradas las que traían puntos de cambio, se redujo la serie de estaciones con datos completos y homogéneos a 91 para precipitación y 53 para temperatura. Este es el resultado final del trabajo. Todas las series se encuentran disponibles para descarga libre en el repositorio <https://github.com/MarkoAndrade/PMVC-UMSS>.

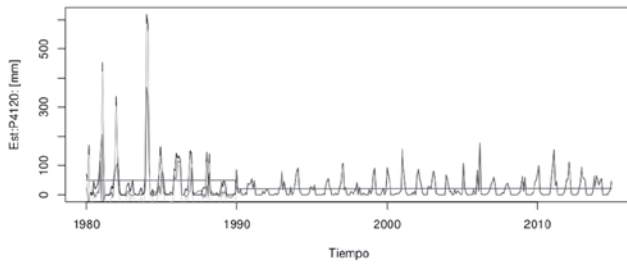


Figura 16. Homogeneización de la estación P4120 (Base) con la referencia CO3874 en RHTestsV4 para Precipitación. La línea azul muestra una discontinuidad (punto de cambio).

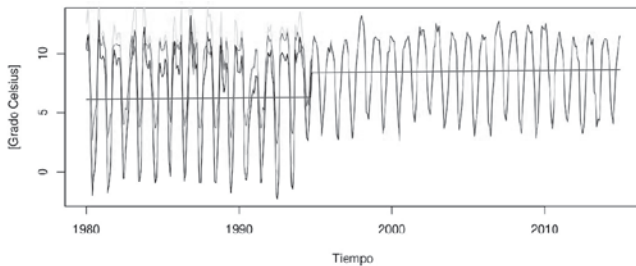


Figura 17. Homogeneización de la estación CP2775 en RHTestsV4 para Temperatura Mínima. La línea azul muestra una discontinuidad (punto de cambio).

4. Conclusiones

El trabajo permitió desarrollar una nueva alternativa para el control de calidad de series climáticas, la cual tiene como principal novedad la inclusión de métodos basados en Redes Neuronales Profundas (DNN). Asimismo, incorporó a RHTests como una herramienta de homogeneización diferente a las que se habían utilizado en estudios previos. Ambas mejoras permitieron obtener series mensuales de precipitación y temperatura para toda Bolivia, completas y homogéneas, con mejores ajustes que los obtenidos con procedimientos tradicionales. El desarrollo acelerado de los métodos basados en redes neuronales y su inclusión en la solución de problemas climatológicos, hace pensar que a futuro estos métodos serán preferibles en comparación con los métodos estadísticos tradicionales.

En cuanto al resultado final, no discrepa demasiado de reportes previos sobre la temática. Muestra que los datos provistos por las estaciones climatológicas del país son, en su mayor parte, incompletos y faltos de homogeneidad. Esto hace que, tras un control de calidad como el del presente trabajo, se tengan que desechar alrededor del 90% de las estaciones disponibles de inicio, reduciendo considerablemente la información

climática disponible en el país. Esta situación obliga a buscar continuamente mejoras para el relleno y homogeneización de series climáticas en Bolivia.

Agradecimientos

El presente trabajo se realizó en el marco del Proyecto de Modelación de Variables Climáticas, PMVC, del Dpto. de Física de la UMSS. Los autores agradecen a SENAMHI-Bolivia por la provisión de los datos crudos de su base SISMET. Asimismo, a Gerald Corzo del IHE-DELFT por la colaboración en el desarrollo de las redes neuronales.

5. Referencias bibliográficas

- Chollet, F. (2018). Deep Learning mit Python und Keras: Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek. MITP-Verlags GmbH & Co. KG.
- Domonkos, P. (2014). Homogenization of precipitation time series with ACMANT. Theor. Appl. Climatol., 122:303-314.
- Domonkos, P. and Coll, J. (2017). Homogenization of temperature and precipitation time series with ACMANT3: method description and efficiency tests. Int. J. Climatol. 37: 1910–1921.
- Guijarro, J.A. (2018). Package 'climatol'. <https://cran.r-project.org/web/packages/climatol/climatol.pdf>. Accedido en enero de 2019.
- Guijarro, J. A. e. a. (2017). Comparison of homogenization packages applied to monthly series of temperature and precipitation: The multitest project. Ninth seminar for homogenization and quality control in climatological databases and Fourth conference on spatial interpolation techniques in climatology and meteorology, Budapest, Hungary, 03 – 07 April 2017. World Meteorological Organization.
- Harris, I., Jones, P., Osborn, T., and Lister, D. (2014). Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset.
- Keras (2019). Consultado el 25 de julio de 2019. <https://keras.io>.
- Menne, M. J. and Williams, C. N. J. (2005). Detection of undocumented changepoints using multiple test statistics and composite reference series. J. Climate 18:4271-4286.

- Mestre, O. e. a. (2013). HOMER: a homogenization software - methods and applications. *Időjárás*, 117:47-67.
- Moolayil, J., Moolayil, and John, S. (2019). *Learn Keras for Deep Neural Networks*. Springer.
- PCA (2019). Consultado el 24 de julio de 2019. https://platform.princetonclimate.com/PCA_Platform/lafdmLanding.html.20
- Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., et al. (1998). Homogeneity adjustments of in situ atmospheric climatedata: a review. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 18(13):1493–1517.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- SENAMHI (2019). Servicio Nacional de Meteorología e Hidrología de Bolivia, t.
- Szentimrey, T. (2007). *Manual of homogenization software MASHv3.02*. Hungarian Meteorological Service, 65 pp.
- Venema, V. K. C. e. a. (2012). Benchmarking homogenization algorithms for monthly data. *Clim.Past*, 8, 89–115.
- Wang, X. and Feng, Y. (2013a). Rhtestv4 user manual. *UserManual.doc* Available online at <http://cccma.seos.uvic.ca/ETCCDMI/software.shtml>.
- Wang, X. L. (2008a). Accounting for autocorrelation in detecting mean-shifts in climate data series using the penalized maximal t or F test. *J. Appl. Meteor. Climatol.*, 47, 2423-2444.
- Wang, X. L. (2008b). Penalized maximal F-test for detecting undocumented mean-shifts without trend-change. *J. Atmos. Oceanic Tech.*, 25 (No. 3), 368-384.
- Wang, X. L. and Feng, Y. (2013b). *RHtestsV4 User Manual*. Climate Research Division, Atmospheric Science and Technology Directorate, Science and Technology Branch, Environment Canada. 28 pp. [Available online at <http://etccdi.pacificclimate.org/software.shtml>].
- Wang, X. L. e. a. (2007). Penalized maximal t test for detecting undocumented mean change in climatedata series. *J. Appl. Meteor. Climatol.*, 46 (No. 6), 916-931