
Relación espacio-temporal

entre estaciones utilizadas
para el relleno de datos
de precipitación en Chone, Ecuador

Spatial-temporal relation between
stations used for filling missing rainfall data
in Chone, Ecuador

María Fernanda Ayala¹

David Carrera-Villacrés^{2,3}

Alfonso Tierra²

¹ Universidad de las Fuerzas Armadas (ESPE), Departamento de Ciencias de la Tierra y la Construcción, Carrera de Ingeniería Geográfica y Medio Ambiente. Sangolquí, Ecuador

² Universidad de las Fuerzas Armadas (ESPE), Centro Geoespacial.

³ Universidad Central del Ecuador, Facultad de Ingeniería, Geología, Minas, Petróleos y Ambiental (FIGEMPA). Quito, Ecuador.

mfayala@espe.edu.ec

dvcarrera@espe.edu.ec

artierra@espe.edu.ec

Resumen

En el presente estudio se aplica la metodología de regresión lineal para completar los datos faltantes de los registros de precipitación de la estación climatológica principal de Chone, a partir de la información de seis estaciones cercanas con una temporalidad similar a la estación de interés. Los datos obtenidos fueron validados mediante un test de rachas y el método de curva de doble masa. El objetivo del trabajo fue encontrar un grado de relación entre las distancias de las estaciones auxiliares a la estación de Chone (d) y los promedios de los coeficientes de correlación (r) de las estaciones auxiliares, obtenidos en el proceso de relleno de datos faltantes. Se obtuvieron así coeficientes de correlación (r) con valores superiores a 0,5 representando correlaciones positivas moderadas e intensas. Finalmente, se determinó que las variables d y r están linealmente relacionadas en un sentido inverso.

Palabras clave: regresión lineal; test de rachas; representatividad; datos ausentes.

Abstract

In this study, the methodology of linear regression was used to complete the missing rainfall records of the Chone main climatological station from the information of six nearby stations with a similar temporality with the station of interest. The obtained data were validated by a run test and the double mass curve method. The objective of this study was to find a degree of relationship between the distances of the auxiliary stations to the Chone station (d) and the averages of the correlation coefficients (r) of the auxiliary stations obtained in the process of filling the missing data. Correlation coefficients (r) were obtained with values higher than 0.5, representing moderate and intense positive correlations. Finally, it was determined that the variables d and r were linearly related in inverse sense.

Key words: linear regression; run test; representativity; missing data.

1. Introducción

La climatología es una rama dinámica de la ciencia con una amplia gama de funciones y aplicaciones. En la actualidad, se están desarrollando nuevas investigaciones con objeto de estudiar la aplicación de estudios climáticos en áreas como la agricultura, la silvicultura, los ecosistemas, la energía, la industria, la producción, el transporte, el turismo, la gestión de recursos hídricos y de desastres, la pesca y el desarrollo de las zonas costeras. Todos los productos y servicios climáticos, como lo es la información derivada de datos meteorológicos y climáticos obtenidos en el pasado hasta las estimaciones sobre el clima futuro, deben sustentarse en datos recopilados de observaciones y registros amplios de una serie de variables clave que permiten caracterizar el clima en distintas escalas temporales (OMM, 2011).

Las condiciones climáticas pueden afectar la producción agraria, por lo tanto, es fundamental tener conocimiento de los recursos disponibles en el territorio. Sin embargo, en las estaciones climatológicas del Ecuador es común que los registros estén incompletos. La ausencia de datos suele deberse a situaciones que afectan la calidad de los datos registrados como lo es el movimiento de los instrumentos, la transformación del espacio físico del entorno donde se encuentra la estación o no registrar el dato en el momento adecuado (Medina *et al.*, 2008; Trujillo *et al.*, 2015). La disponibilidad de series históricas climáticas completas y de períodos largos, así como la homogeneidad de los datos son de vital importancia para responder a las necesidades de las investigaciones sobre el clima, además de ser un requisito esencial para realizar estudios confiables y representativos (Carrera *et al.*, 2015; Medina *et al.*, 2008; OMM, 2011). El instituto encargado de suministrar informa-

ción sobre el tiempo, el clima y los recursos hídricos del pasado, presente y futuro en el Ecuador es el Instituto Nacional de Meteorología e Hidrología (INAMHI). En muchos casos, se ha verificado la falta de registros de datos en las series históricas de distintas variables meteorológicas proporcionadas por el INAMHI; por consiguiente, es necesario generar los datos a partir de información existente (Campozano *et al.*, 2015; Carrera *et al.*, 2016; Gutiérrez *et al.*, 2008; Román y De la Paz, 2010).

Tal fue el caso del presente estudio donde se evidenció la ausencia de datos en el registro de precipitación media mensual medida en milímetros (mm), de la estación climatológica principal de Chone (M0162), desde 1962 hasta 2015. Ante el problema de falta de datos, varias metodologías han sido desarrolladas para completar los registros (ej. Método de la razón normal, regresión lineal, completación por regresiones múltiples, métodos multivariantes, combinación lineal ponderada, transformada de *wavelett*, método de las isoyetas, redes neuronales, interpolación por distancia inversa ponderada, entre otros), los cuales han presentado resultados satisfactorios (Campozano *et al.*, 2015; Carrera *et al.*, 2016; Londhe *et al.*, 2015; Pegram, 1997; Pizarro *et al.*, 2009; Trujillo *et al.*, 2015; Vicente-Serrano *et al.*, 2010).

En el presente trabajo se utilizó la metodología de regresión lineal debido a que es uno de los métodos que propone la OMM para el relleno de datos faltantes, y es recomendado en la estimación de datos mensuales y anuales, además de ser la mejor metodología del relleno de datos en series temporales en la región andina y costera del Ecuador (Carrera *et al.*, 2016; OMM, 2011; Pizarro *et al.*, 2009).

El objetivo de este trabajo fue encontrar un grado de relación entre las distancias de las estaciones auxiliares a la de Chone (d) y los

promedios de los coeficientes de correlación (r) de las estaciones auxiliares obtenidos en el proceso de relleno de datos faltantes de la serie de tiempo de precipitación para poder predecir cuál es la probabilidad de que la función explique las variables d y r .

2. Materiales y métodos

2.1 Área de estudio

La ciudad de Chone dispone de una estación climatológica principal localizada al noreste de su zona urbana; de acuerdo con los datos de dicha estación, el clima es tropical, la temperatura del aire registra una media anual de 25,6°C, máxima media anual de 33,6°C y mínima media anual de 20,1°C. La precipita-

ción máxima en 24 horas registró un valor de 51,8 mm en el año 2015 (INAMHI, 2015). En la **Figura 1** se muestra la representación gráfica del área de estudio dibujada mediante el software *ArcGis 10.1*; utilizando la cartografía base de libre acceso obtenida del geportal del Instituto Geográfico Militar del Ecuador (IGM, 2017). Se puede apreciar la ubicación de la estación climatológica principal de Chone (M0162), simbolizada mediante un triángulo de color azul. Para el estudio se recopiló la información de seis estaciones cercanas representadas por triángulos de color rojo. Se eligieron dichas estaciones auxiliares debido a su cercanía con la estación M0162; además, se cumplió con la condición de que su base de datos tuviese una temporalidad similar a la estación en estudio.

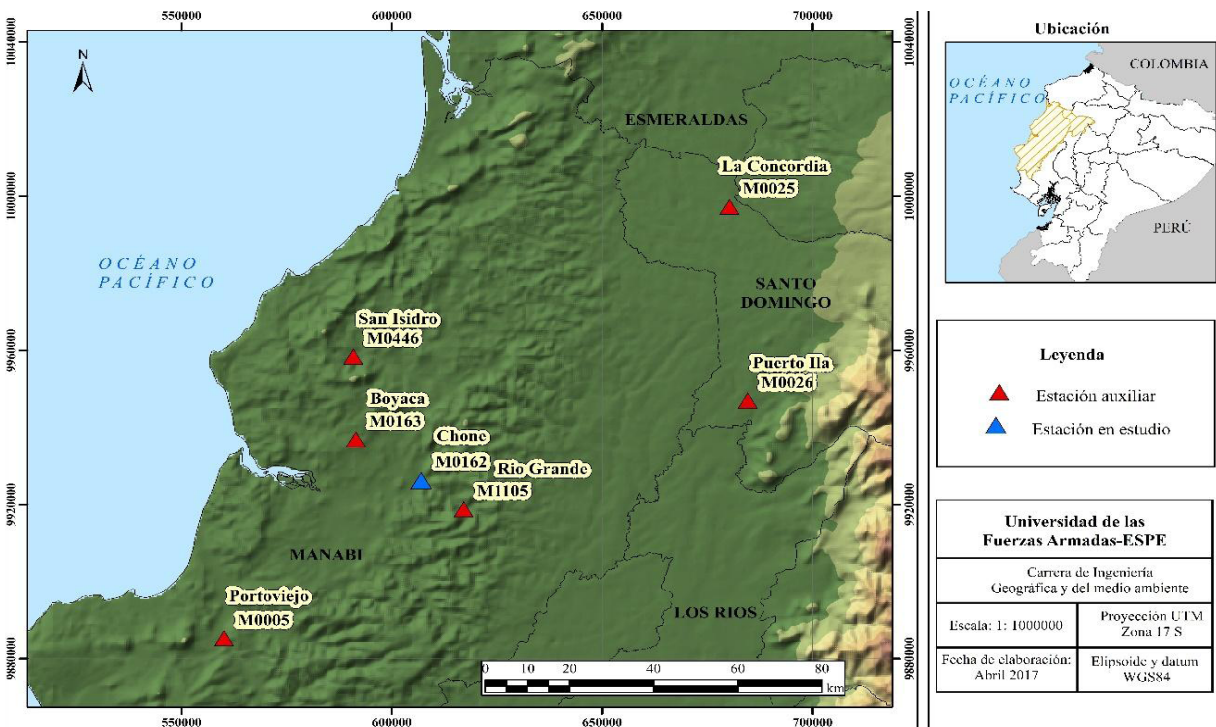


Figura 1 Ubicación de estaciones meteorológicas en la zona de estudio y estaciones auxiliares aledañas

2.2 Relleno de datos

La guía de prácticas climatológicas de la Organización Meteorológica Mundial (OMM, 2011) propone diversas metodologías utilizadas para el relleno de datos faltantes de bases de datos meteorológicas. Una de ellas es la metodología empleada por Carrera *et al.* (2016), la cual supone una relación lineal entre variables, para esto se establece una regresión y correlación lineal entre una estación auxiliar y la estación en estudio con carencia de información, mediante una ecuación de dos variables (1)

$$y = a + bx \quad (1)$$

Donde,

y = valor estimado de la variable para la estación carente,

x = valor de la variable registrada por la estación auxiliar,

a, b = parámetros de regresión.

Por medio de una gráfica de dispersión se determinó el grado de relación y se utilizó el coeficiente de correlación como una forma de establecer la calidad de los datos y la intensidad de la relación del conjunto de variables. Este método se recomienda para la estimación de datos mensuales y anuales, además es la mejor metodología para el relleno de datos en series temporales para la región andina y costera del Ecuador (Carrera *et al.*, 2016; Lind *et al.*, 2004; Pizarro *et al.*, 2009).

Un coeficiente de correlación, también llamado r de Pearson, de -1 o de $+1$ indica una correlación perfecta. Si no existe en absoluto alguna relación estadística entre el conjunto de variables, r será cero. Un coeficiente de correlación r cercano a 0 indica que la relación es muy débil. En el Cuadro 1 se resume la intensidad y la dirección del coeficiente de correlación (Lind *et al.*, 2004).

En términos hidrológicos, se considera aceptable una regresión cuyo valor de r sea mayor a $0,8$ o menor que $-0,8$. En el caso de no existir una buena correlación, el modelo de regresión lineal permitirá tener una buena aproximación de la estimación de los datos (Cruz, 2014; Pizarro *et al.*, 1993).

La metodología utilizada para validar los datos estimados fue el test de rachas y la curva de doble masa. El test de rachas se usó para contrastar si el orden de aparición de dos valores de la variable era aleatorio (López, 2007). Una racha es la sucesión de valores de la misma categoría, se clasifica a los valores por encima de la mediana de los datos representados como NA , y los valores por debajo de la mediana se los denomina NB , por definición NA y NB tienen aproximadamente el mismo valor ($NA = NB \pm 1$), al número total de rachas (NS) le corresponde una probabilidad de excedencia entre 10% y 90% (Barón *et al.*, 2004; Caballero, 2013; Cando, 2010). A continuación, en el Cuadro 2 se muestra el dominio de aceptación del test de rachas en función del número de datos (Caballero, 2013).

El segundo método aplicado para la validación de los datos rellenados de las series de tiempo de cada una de las variables, fue el método curva doble masa o de acumulación, el cual permite establecer la homogeneidad de los datos (Cruz, 2014). Según Martín-Vide y García (2013), una serie climática es homogénea cuando sus valores sólo reflejan el comportamiento natural (climático) de la correspondiente variable, sin interferencias o perturbación por hechos artificiales, es decir sus variaciones son fruto exclusivo de variaciones en el clima. Sólo cuando se garantice la homogeneidad de la serie, se podrá proceder al análisis temporal de la misma. El método consiste en realizar una gráfica de dispersión

Cuadro 1 Intensidad y dirección del coeficiente de correlación

Coefficiente de correlación r	Intensidad de relación lineal
$r = -1$	Correlación negativa perfecta
$-1 < r < -0,5$	Correlación negativa intensa
$r = -0,5$	Correlación negativa moderada
$-0,5 < r < 0$	Correlación negativa débil
$r = 0$	Sin correlación
$0 < r < 0,5$	Correlación positiva débil
$r = 0,5$	Correlación positiva moderada
$0,5 < r < 1$	Correlación positiva intensa
$r = 1$	Correlación positiva perfecta

Cuadro 2 Valores críticos del test de rachas

Test de las rachas de Thom							
Valores críticos. Nivel de confianza del 95%							
Nº de datos	10	20	30	40	50	60	80
Nº mínimo de rachas	2	6	10	14	18	22	31
Nº máxima de rachas	6	15	21	27	33	39	50

donde las parejas de puntos son obtenidas por las acumulaciones sucesivas de dos series de valores del mismo periodo, si los valores son proporcionales se distribuirán en una línea recta (Oñate, 2009).

2.3 Distancia entre estaciones meteorológicas

Las observaciones meteorológicas deben ser representativas de una zona conforme a su aplicación. La representatividad es definida por la escala temporal y espacial de cada fenómeno. La OMM en 1981 presentó un Manual del Sistema Mundial de Observación, donde establece una clasificación para las escalas horizontales de los fenómenos meteorológicos, la cual se muestra en el Cuadro 3. Una estación situada en un terreno accidentado o

Cuadro 3 Escalas meteorológicas horizontales

Escala	Rango
Pequeña escala	< 100 km
Meso escala	100 a 1000 km
Gran escala	1000 a 5000 km
Escala planetaria	> 5000 km

en la costa probablemente no sea representativa a gran escala o meso escala. Sin, embargo, incluso en estaciones no representativas, la homogeneidad de las observaciones en el tiempo puede permitir a los usuarios emplear eficazmente los datos (OMM, 1996).

Se empleó la clasificación anterior para representar la distancia horizontal entre las estaciones mediante un mapa, utilizando la cartografía base del IGM y los archivos *shape*

de tipo punto proporcionados por el INAMHI, donde se encuentra la ubicación de las estaciones meteorológicas auxiliares: M005 (Portoviejo), M0025 (La Concordia), M0026 (Puerto Ila), M0163 (Boyacá), M0446 (San Isidro) y M1105 (Río Grande), así como la estación estudiada: M0162 (Chone). Finalmente, se aplicó una regresión lineal que relacionó a las variables d y r , mediante la fórmula (1), siendo x el valor de las distancias entre cada una de las estaciones auxiliares a la estación de Chone, e y el valor de los coeficientes de correlación promedio de las estaciones auxiliares obtenidos en el proceso de relleno de datos.

3. Resultados

3.1 Relleno de datos

Para realizar el relleno de datos de la serie de tiempo de precipitación, primero se determinó la cantidad de datos faltantes. La metodología de la regresión lineal simple se aplicó tomando los promedios mensuales de todos los años de manera vertical, en el caso de la ausencia de los datos mensuales de todo un año en específico como indica la metodología utilizada por Carrera *et al.* (2016), cuando los datos faltantes fueron pocos meses de un año, se realizó una regresión lineal tomando los datos de manera horizontal de cada año faltante. Por lo cual fue necesario analizar si el dato faltante se podía estimar mediante una regresión que tomara los datos verticalmente u horizontalmente, teniendo en cuenta el formato en el que se recopiló la información. Se realizó una regresión lineal para cada dato faltante con cada una de las estaciones auxiliares. A las estaciones M1105, M0163 y M0446 se las usó solamente para el relleno de años en específico, debido a que fueron instaladas

en una época más reciente que la estación M0162 de Chone. Se determinó el coeficiente de correlación de cada regresión lineal. Para la estimación finalmente se usó la ecuación de regresión lineal de aquella estación que tuviera el coeficiente de correlación con un valor más alto, debido a que esta herramienta es recomendada por la OMM (2011).

Se utilizaron 583 datos del total de la base de datos de precipitación media medida en milímetros (mm) proporcionada por el INAMHI, desde 1962 hasta 2015. Se tomó la decisión de eliminar 12 datos correspondientes a los meses del año 1978 y se los incluyó en los datos faltantes a ser rellenos, debido a que en este año se presentaron valores anómalos de precipitaciones muy altas con una media de 740,5 mm; este valor se lo consideró erróneo debido a que en los años de episodios de fenómeno El Niño (1965, 1966, 1969, 1972, 1973, 1976, 1983, 1992, 1998), donde ocurren precipitaciones excepcionalmente altas que pueden alcanzar valores varias veces superiores a las ocurridas en una época de lluvia normal o fuerte (ACOLIT, 2008), se presentaron precipitaciones con valores medios entre 84,0 mm y 202,4 mm siendo muy inferiores a los del año 1978. Además, se compararon estos valores con la base de datos de las estaciones auxiliares M0005, M0025 y M0026. Se comprobó que en 1978, no se registraron precipitaciones de tal magnitud. Adicionalmente, el valor de precipitación media mensual del año 1978 supera los records de precipitación máxima mensual en el país correspondiente a 586,9 mm en una de las zonas más húmedas del mismo (INAMHI, 2015).

La cantidad de datos faltantes que se estimó en el relleno fueron 65. A continuación, en el **Cuadro 4**, se señala (negrita) la estación auxiliar utilizada para el relleno de los datos. Se

Cuadro 4 Proceso de cálculo por regresión lineal tomando los datos **a)** verticalmente y **b)** horizontalmente

Relleno vertical de la serie de precipitación				
Mes	Estación auxiliar	Ecuación	r	Datos rellenos
Enero	M0025	$y=0,2713x+86,30$	0,44	4
	M0026	$y=0,2173x+117,29$	0,29	
	M0005	$y=1,0567x+116,08$	0,63	
Febrero	M0025	$y=0,4666x+53,56$	0,50	4
	M0026	$y=0,093x+261,51$	0,11	
	M0005	$y=1,0251x+172,55$	0,59	
Marzo	M0025	$y=0,509x-38,38$	0,69	5
	M0026	$y=0,477x+35,68$	0,63	
	M0005	$y=1,101x+126,47$	0,69	
Abril	M0025	$y=0,2897x+7,87$	0,58	5
	M0026	$y=0,3066x+36,82$	0,49	
	M0005	$y=1,1094x+105,92$	0,63	
Mayo	M0025	$y=0,3016x-26,75$	0,57	5
	M0026	$y=0,5681x-60,00$	0,66	
	M0005	$y=1,7653x+17,01$	0,88	
Junio	M0025	$y=0,2707x-2,40$	0,68	5
	M0026	$y=0,4895x-13,76$	0,74	
	M0005	$y=2,2383x+17,05$	0,74	
Julio	M0025	$y=0,3741x-7,37$	0,84	4
	M0026	$y=0,5957x-9,99$	0,86	
	M0005	$y=1,3386x+14,30$	0,89	
Agosto	M0025	$y=0,3322x-4,15$	0,95	4
	M0026	$y=0,4614x-6,81$	0,76	
	M0005	$y=2,7836x+6,53$	0,85	
Septiembre	M0025	$y=0,1307x+1,87$	0,81	4
	M0026	$y=0,1629x+1,50$	0,81	
	M0005	$y=1,5559x+6,20$	0,85	
Octubre	M0025	$y=0,2289x-2,37$	0,82	4
	M0026	$y=0,3571x-6,85$	0,81	
	M0005	$y=2,8822x+12,17$	0,56	
Noviembre	M0025	$y=0,4405x-9,86$	0,90	4
	M0026	$y=0,4068x-5,87$	0,80	
	M0005	$y=2,1014x+9,84$	0,79	
Diciembre	M0025	$y=0,5012x-30,56$	0,74	4
	M0026	$y=0,4569x-30,23$	0,88	
	M0005	$y=1,4994x+25,39$	0,87	

Cuadro 4 Continuación

Relleno horizontal de la serie de precipitación				
Mes	Estación auxiliar	Ecuación	r	Datos rellenados
1967	M0025	$y=0,6074x-30,73$	0,77	2
	M0026	$y=0,6266x-42,44$	0,84	
	M0005	$y=2,2556x+17,65$	0,99	
	M0163	$y=1,5535x-4,672$	0,96	
1983	M0025	$y=0,6555x-28,66$	0,85	1
	M0026	$y=0,9381x-108,51$	0,79	
	M0005	$y=1,7229x+65,89$	0,80	
2008	M0025	$y=0,5042x-42,05$	0,93	2
	M0026	$y=0,613x-41,90$	0,93	
	M0005	$y=1,5288x+10,76$	0,99	
	M0446	$y=1,6174x-25,49$	0,91	
2009	M0025	$y=0,2368x+10,94$	0,70	1
	M0026	$y=0,2744x+8,11$	0,88	
	M0005	$y=1,7574x+12,57$	0,92	
2011	M0025	$y=0,2487x+10,05$	0,73	3
	M0026	$y=0,3557x-3,60$	0,99	
	M0005	$y=2,5995x+10,36$	0,83	
2012	M0025	$y=0,3988x-12,92$	0,90	2
	M0026	$y=0,5x-10,59$	0,87	
	M0005	$y=1,4755x+24,73$	0,98	
	M1105	$y=0,5542x-0,28$	0,99	
	M0163	$y=0,9616x+20,776$	0,93	
2013	M0025	$y=0,5911x-6,31$	0,95	1
	M0026	$y=0,6453x-22,63$	0,97	
	M0005	$y=2,0252x+13,19$	0,99	
	M1105	$y=0,7155x+6,6836$	0,99	
	M0163	$y=2,9098x+48,388$	0,73	
2015	M0025	$y=0,2725x+39,07$	0,64	1
	M0026	$y=0,4941x+5,021$	0,79	
	M0005	$y=1,55127x+60,32$	0,83	

muestra también el coeficiente de correlación r de cada ecuación de las estaciones auxiliares, y la cantidad de datos rellenados con la regresión lineal; en la **Figura 2** se muestra el gráfico de dispersión realizado en Excel, del mes de diciembre, de manera similar se obtuvo los gráficos de todos los meses y años de la serie.

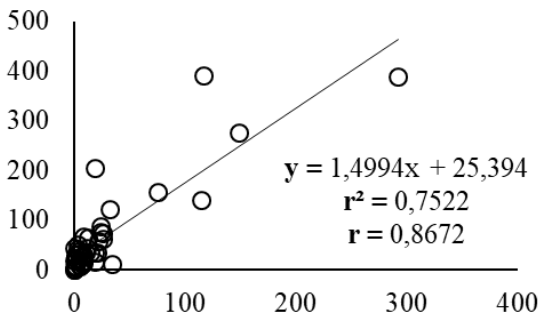


Figura 2 Gráfica de dispersión del mes de diciembre de la serie de precipitación

3.2 Validación de los resultados

Los datos estimados se validaron mediante un test de rachas. En el **Cuadro 5** se muestra el test de rachas aplicado a la serie de precipitación rellenada en los años con datos faltantes; de manera similar se empleó el test para los datos rellenados mensualmente. Se puede observar que el número total de rachas (NS), concuerda con las recomendaciones del dominio de aceptación del test de Caballero (2013).

El siguiente método para validar fue el de curva de doble masa; en la **Figura 3**, se observa como ejemplo la gráfica de curva doble de masa realizada en Excel, de la serie de precipitación en el año 2013. De igual manera, se aplicó este método de validación en el resto de años rellenados de la serie.

Las ecuaciones de regresión lineal permitieron estimar los datos de cada fecha en la que hubo datos perdidos, finalmente, se pudo

rellenar la serie; en la **Figura 4** se muestra la gráfica realizada mediante el software Matlab de **a)** la serie original de precipitación y **b)** la serie completada mediante la metodología de regresión lineal. Así, finalmente se tiene un total de 648 datos de precipitación, desde el año 1962 hasta el 2015.

Se determinó la distancia entre la estación climatológica principal de Chone M0162 y las estaciones auxiliares. A continuación, en la **Figura 5** se puede apreciar visualmente la distancia horizontal entre las estaciones. Se usó la clasificación de escalas meteorológicas horizontales de la OMM para identificar las distancias que debían representarse en el mapa dibujado mediante el software ArcGis 10.1.

Se usó la información obtenida de los coeficientes de correlación de las estaciones auxiliares y se determinó un promedio r para tener un dato de cada estación auxiliar. Se buscó el grado de relación de r con las distancias d obtenidas entre la estación de Chone y las estaciones auxiliares, mediante una regresión lineal simple realizada en Excel (**Figura 6**). En el **Cuadro 6**, se muestran los valores de r y d , de cada una de las estaciones auxiliares.

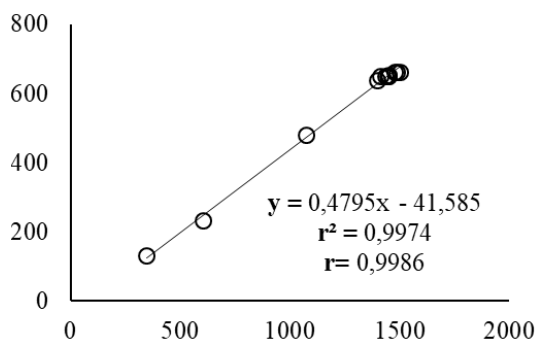
4. Discusión de los resultados

La metodología de la regresión lineal simple utilizada para rellenar los datos faltantes de la serie de tiempo de precipitación, permitió estimar 65 datos, los cuales representan el 10% del total de la base de datos. Este valor cumplió con el umbral propuesto en el estudio de Campozano *et al.* (2015), en el cual se recomienda un umbral correspondiente al 25% del total de la base de datos, dicho criterio fue tomado debido a la escasez de información de datos meteorológicos registrados por el INAMHI en la zona andina ecuatoriana.

Cuadro 5 a) Serie con los datos rellenos. **b)** Test de rachas

a) Datos rellenos													
Años	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sept	Oct	Nov	Dic	Mediana
1967	408,1	411,5	87,6	50,1	54,3	8,5	24,0	17,7	39,0	3,8	0,0	6,4	31,5
1983	496,6	593,0	541,5	284,1	681,6	614,8	310,1	133,8	114,5	9,2	8,5	62,8	297,1
2008	354,6	393,0	202,2	95,1	25,4	10,5	29,0	12,2	5,4	12,0	12,0	11,1	18,8
2009	201,4	206,7	99,9	94,6	23,3	6,6	0,6	27,3	29,1	12,6	0,0	41,9	28,2
2011	147,4	108,3	91,6	265,4	4,2	20,8	66,9	6,9	0,3	13,6	2,6	72,3	43,9
2012	257,9	429,4	388,3	96,0	195,7	59,0	5,8	3,7	0,4	3,9	15,5	23,9	41,5
2013	347,2	257,9	467,5	323,0	18,1	22,0	7,2	5,3	1,9	26,6	9,7	15,8	20,1
2015	231,2	301,0	322,3	139,4	168,1	99,8	104,7	5,8	1,6	50,2	8,7	90,3	102,3

b) Prueba de rachas														
Años	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sept	Oct	Nov	Dic	NA	NS
1967	+	+	+	+	+	-	-	-	+	-	-	-	6	4
	1	0	0	0	0	1	0	0	1	1	0	0		
1983	+	+	+	-	+	+	+	-	-	-	-	-	6	4
	1	0	0	1	1	0	0	1	0	0	0	0		
2008	+	+	+	+	+	-	+	-	-	-	-	-	6	4
	1	0	0	0	0	1	1	1	0	0	0	0		
2009	+	+	+	+	-	-	-	-	+	-	-	+	6	5
	1	0	0	0	1	0	0	0	1	1	0	1		
2011	+	+	+	+	-	-	+	-	-	-	-	+	6	5
	1	0	0	0	10	0	1	1	0	0	0	1		
2012	+	+	+	+	+	+	-	-	-	-	-	-	6	2
	1	0	0	0	0	0	1	0	0	0	0	0		
2013	+	+	+	+	-	+	-	-	-	+	-	-	6	6
	1	0	0	0	1	1	1	0	0	1	1	0		
2015	+	+	+	+	+	-	+	-	-	-	-	-	6	4
	1	0	0	0	0	1	1	1	0	0	0	0		

**Figura 3** Curva de doble masa de la serie de precipitación en el año 2013

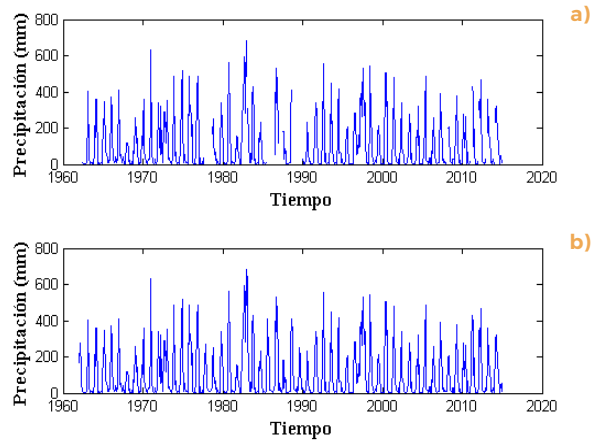


Figura 4 a) serie original y. b) serie completa de la variable precipitación

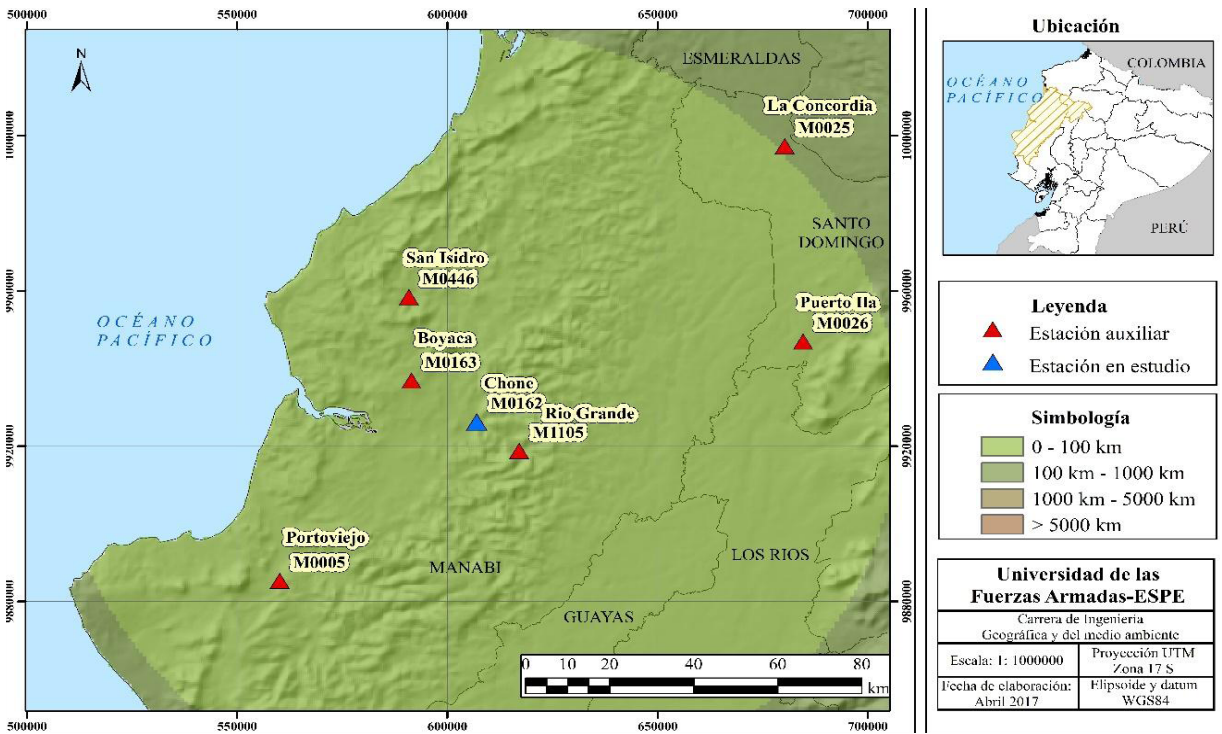


Figura 5 Distancias entre la estación de Chone M0162 y las estaciones auxiliares

Cuadro 6 Coeficientes de correlación y distancias de estaciones auxiliares

Estación auxiliar	Nombre de la estación	r	d (km)
M0005	Portoviejo-UTM	0,82	53,7
M0025	La Concordia	0,75	102
M0026	Puerto Ila	0,75	80
M1105	Río Grande INAMHI	0,99	12
M0163	Boyacá	0,96	18,8
M0446	San Isidro-Manabí	0,91	36,02

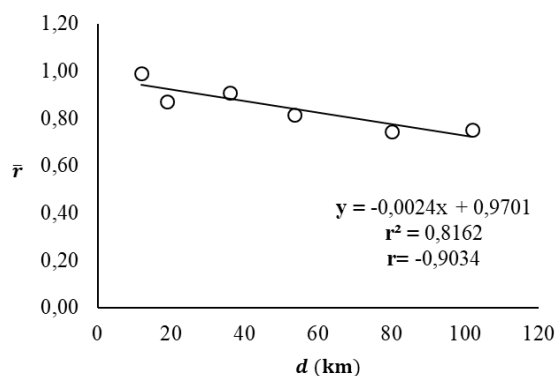


Figura 6 Gráfico de dispersión que relaciona r y d

Los coeficientes de correlación de las ecuaciones obtenidas para el relleno de datos tienen un valor de $r > 0,5$ representando correlaciones positivas moderadas e intensas como se puede corroborar en el Cuadro 1. En el estudio de Carrera *et al.* (2016), se obtienen valores similares para la región andina y costera del Ecuador concluyendo que el método de regresión lineal simple fue el mejor para rellenar datos, debido a la gran cantidad de datos faltantes que provee el INAMHI. El estudio realizado por Pizarro *et al.* (2009) en la región del Maule en Chile, concluyó que el método de regresión lineal simple, a pesar de ser el más

recomendado en la literatura, en la práctica no se obtienen los mejores resultados, al ser comparado con otros métodos de completación como la regresión múltiple usando dos y tres estaciones auxiliares.

En cuanto a las distancias horizontales entre las estaciones auxiliares utilizadas se observó que 5 de ellas son representativas para la zona de estudio, al ser clasificadas como pequeña escala por estar a menos de 100 km de la estación climatológica principal de Chone, según la clasificación de la OMM mostrada en el Cuadro 3. La estación M0025 se encuentra a una distancia de 102 km, y obtuvo coeficientes de correlación en su mayoría positiva (moderada a intensa) en el proceso de relleno de datos.

La regresión lineal empleada para relacionar d y r estableció una correlación negativa intensa, lo cual significó que estas variables están linealmente relacionadas en un sentido inverso, es decir a mayor d se tendrá un menor r . Resultados similares se obtuvieron en la investigación de El Kenawy *et al.* (2013), en un estudio realizado al noreste de España, donde la correlación entre estaciones fue mayor entre las más cercanas y fue disminuyendo gradualmente entre aquellas que se encontraban a grandes distancias; además en la investigación, se pudo mejorar la correlación entre las estaciones al eliminar el ruido de las series.

Finalmente, se determinó que la estación mejor correlacionada a la estación de Chone es la estación de Río Grande, siendo ésta la estación más cercana. Corroborando así los criterios de representatividad de las estaciones para las observaciones meteorológicas de una zona expuestas por la OMM.

5. Conclusiones

Sobre la base de datos de la serie de tiempo de precipitación utilizada en el estudio se concluyó que la metodología de regresión lineal obtuvo buenos resultados con valores de $r > 0,5$ representando correlaciones positivas moderadas e intensas. Además de ser una metodología de fácil aplicación, los datos obtenidos del relleno fueron validados mediante el test de rachas, cuyos valores de NS entraron en su dominio de aceptación y la técnica de curva doble de masa permitió garantizar la homogeneidad de la serie. Se estimaron así 65 datos faltantes, los cuales representaron el 10% del

total de la base de datos de la serie de tiempo de precipitación. Las distancias horizontales entre las estaciones auxiliares a la estación de Chone fueron menores a 102 km lo que permitió usar las observaciones, ya que fueron representativas de la zona de estudio con coeficientes de correlación promedio superiores a 0,7 mostrando una correlación positiva intensa. Finalmente, el grado de relación entre las variables d y r , estableció una correlación negativa intensa lo cual significó que estas variables están linealmente relacionadas en un sentido inverso.

6. Agradecimientos

Los autores agradecen al INAMHI por proveer los datos meteorológicos necesarios para la realización de esta investigación.

7. Referencias citadas

- ASESORES y CONSULTORES DEL LITORAL (ACOLIT). 2008. *Actualización de los estudios definitivos del proyecto de propósito múltiple Chone. Estudio Hidrológico*. Documento técnico de apoyo N° 9. Manabí, Ecuador.
- BARÓN, J.; RÍUS, F. y E. SÁNCHEZ. 2004. *Bioestadística: métodos y aplicaciones*. Manuales de métodos y aplicaciones. Universidad de Málaga, España.
- CABALLERO, I. 2013. *Análisis de la Homogeneidad de las Series de Precipitación de Guipúzcoa*, 119p. Nota técnica N° 13. Ministerio de Agricultura, Alimentación y Medio Ambiente. Agencia Estatal de Meteorología (AEMET), España.
- CAMPOZANO, L.; SÁNCHEZ, E.; AVILES, A. & E. SAMANIEGO. 2015. "Evaluation of infilling methods for time series of daily precipitation and temperature: The case of the Ecuadorian Andes". *Revista Científica Maskana*, 5(1): 99-115.
- CANDO, L. E. 2010. *Incidencia de la variabilidad climática en los caudales mínimos del Ecuador*. Facultad de Ingeniería Civil y Ambiental. Escuela Politécnica Nacional. Quito, Ecuador. Tesis de Grado.
- CARRERA-VILLACRÉS, D.; GUEVARA-GARCÍA, P.; MAYA-CARRILLO, M. & T. CRISANTO-PERRAZO. 2015. "Variations of species concentration in inorganic water dam purpose multiple Chone, Ecuador based on precipitation,

- evaporation and evapotranspiration". *Procedia Earth and Planetary Science*, 15: 641-646.
- CARRERA-VILLACRÉS, D. V.; GUEVARA-GARCÍA, P. V.; TAMAYO-BACACELA, L. C.; BALAREZO-AGUILAR, A. L.; NARVÁEZ-RIVERA C. A. y D. R. MOROCHO-LÓPEZ. 2016. "Relleno de series anuales de datos meteorológicos mediante métodos estadísticos en la zona costera e interandina del Ecuador, y cálculo de la precipitación media". *Idesia (Arica)*, 34(3): 81-90.
- CRUZ, G. P. 2014. "Variables hidrometeorológicas asociadas al cambio climático en Girardot y la Región del Alto Magdalena". *Ambiente y Desarrollo*, 18(35): 134-147.
- EL KENAWY, A.; LÓPEZ-MORENO, J. I.; STEPANEK, P. & S. M. VICENTE-SERRANO. 2013. "An assessment of the role of homogenization protocol in the performance of daily temperature series and trends: application to northeastern Spain". *International Journal of Climatology*, 33(1): 87-108.
- GUTIÉRREZ, C.; GÓNGORA, E. y P. MELO. 2008. *Estudio Hidrológico de Inundaciones en la Cuenca Alta del Río Chone (Garrapata, Mosquito y Grande)*. Estudios e Investigaciones Hidrológicas Superficiales INAMHI, CLIRSEN y FAO. Quito, Ecuador.
- INSTITUTO GEOGRÁFICO MILITAR (IGM). 2017. *Cartografía de libre acceso*. Geoportal IGM. Disponible en: <http://www.geoportaligm.gob.ec/portal/index.php/descargas/cartografia-de-libre-acceso/> [Consulta: junio, 2017].
- INSTITUTO NACIONAL DE METEOROLOGÍA E HIDROLOGÍA (INAMHI). 2015. *Boletín Climatológico Anual 2015*. Editor: José Olmedo. Quito, Ecuador.
- LIND, D. A.; MASON, W. G.; LIND, R. D. D. A.; MARCHAL, W. G. y R. D. MASON. 2004. *Estadística para administración y economía*. Alfaomega Grupo Editor. Bogotá, Colombia.
- LONDHE, S.; DIXIT, P.; SHAH, S. & S. NARKHEDE. 2015. "Infilling of missing daily rainfall records using artificial neural network". *ISH Journal of Hydraulic Engineering*, 21(3): 255-264.
- LÓPEZ, C. P. 2007. *Minería de datos: técnicas y herramientas*. Editorial Paraninfo. Madrid, España.
- MARTÍN-VIDE, J. y M. D. C. M. GARCÍA. M. 2013. "La difícil determinación de la evolución el número de días de tormenta en España. El caso de Barcelona". *Polígonos. Revista de Geografía*, (24): 77-94.
- MEDINA, R. D.; MONTOYA, E. y A. JARAMILLO. 2008. *Estimación estadística de valores faltantes en series históricas de lluvia*. Organización Meteorológica Mundial (1981). Manual del Sistema Mundial de Observación. Volumen I, Aspectos Mundiales, OMM, N° 544. Ginebra, Suiza.
- OÑATE-VALDIVIESO, F. 2009. *Hidrología. Apuntes de clase*. Escuela de Ingeniería Civil. Universidad Técnica Particular de Loja. Ecuador.

- ORGANIZACIÓN METEOROLÓGICA MUNDIAL (OMM). 1981. *Guía de Prácticas Climatológicas*. N° 100. Ginebra, Suiza.
- ORGANIZACIÓN METEOROLÓGICA MUNDIAL (OMM). 1996. *Guía de Instrumentos y Métodos de Observación Meteorológicos*. N° 8. Sexta edición. Ginebra, Suiza.
- ORGANIZACIÓN METEOROLÓGICA MUNDIAL (OMM). 2011. *Guía de Prácticas Climatológicas*. OMM, N° 100. Ginebra, Suiza.
- PEGRAM, G. 1997. "Patching rainfall data using regression methods". 3. Grouping, patching and outlier detection. *Journal of hydrology*, 198(1): 319-334.
- PIZARRO, R.; GONZÁLEZ, P.; WITTERSSHEIN, M.; SAAVEDRA, J. y C. SOTO. 1993. *Elementos técnicos de hidrología III*. Proyecto Regional Mayor Unesco-Rostlac.
- PIZARRO, R.; AUSENSI, P.; ARAVENA, D.; SANGÜESA, C.; LEÓN, L. y F. BALOCCHI. 2009. "Evaluación de métodos hidrológicos para la completación de datos faltantes de precipitación en estaciones de la región del Maule, Chile". *Aqua-LAC: revista del Programa Hidrológico Internacional para América Latina y el Caribe*, 1(2): 172-184.
- ROMÁN, A. y M. DE LA PAZ. 2010. *Instructivos de procesamiento de información hidrometeorológica*. Facultad de Ingeniería Civil y Ambiental. Escuela Politécnica Nacional. Quito, Ecuador. Tesis de Grado.
- TRUJILLO, T.; MARÍA, A.; ARTEAGA RAMÍREZ, R.; VÁZQUEZ PEÑA, M. A. y L. A. IBÁÑEZ CASTILLO. 2015. "Relleno de series diarias de precipitación, temperatura mínima, máxima de la región norte del Urabá Antioqueño". *Revista mexicana de ciencias agrícolas*, 6(3): 577-588.
- VICENTE-SERRANO, S. M.; BEGUERÍA, S.; LÓPEZ-MORENO, J. I.; GARCÍA-VERA, M. A. & P. STEPANEK. 2010. "A complete daily precipitation database for northeast Spain: reconstruction, quality control, and homogeneity". *International Journal of Climatology*, 30(8): 1.146-1.163.