

Interpolation of monthly precipitation amounts in mountainous catchments with sparse precipitation networks

Alexandra P. Jacquin^{1*}, and Jaime C. Soto-Sandoval¹

Most studies dealing with the interpolation of precipitation gauge data have focused in areas where the meteorological network is relatively dense, implying that it is still unknown what interpolation methods are more appropriate in the case of mountain catchments with scarce gauge data. This study evaluates the applicability of Kriging with External Drift (KED) and the Optimal Interpolation Method (OIM) for interpolation of monthly precipitation in these situations. Thiessen Polygons (TP) are used as benchmark. The study area corresponds to the upper subcatchment of Aconcagua River, Central Chile. Cross-validation experiments revealed that all these methods show similar performance in the lower zone of the study area, but OIM outperforms TP and KED at high elevations. Optimal Interpolation Method generally produces the smallest bias in the Andean zone of the study area, with mean errors whose absolute values are smaller than 9% of mean monthly precipitation. From April to September, the root mean squared errors of OIM are between 14% and 33% smaller than those of TP and KED in this zone. Although KED achieves a good agreement to mean monthly values at high elevations (mean errors smaller than 19% in absolute value), its performance is comparable to that of TP in terms of root mean squared errors. Long-term water balances did not provide evidence against the applicability of KED and OIM. Nevertheless, the results of the cross-validation experiments indicate that OIM is a better alternative than KED for the interpolation of monthly precipitation in the study area.

Key words: Kriging with External Drift, Optimal Interpolation Method, Thiessen Polygons.

INTRODUCTION

Precipitation information is essential for the evaluation of water resources availability, water resources planning and design of hydraulic works. However, it is often found that such information is either not available at the location or interest or that data is not representative of catchment wide conditions. Even in zones where long-term precipitation records exist, it is possible that the meteorological network is too sparse or that its spatial distribution is such that precipitation data required is not directly available, implying that precipitation amounts must be estimated through interpolation. The issue of rain gauge interpolation is even more critical in catchments having a complex topography, where the spatial distribution of precipitation is likely to be subjected to orographic effects. In the case of catchments located in the Chilean Andes, in particular, the scarcity of precipitation information poses serious difficulties in the development of water resources studies. In spite of the improvements made in recent years, the meteorological network in the Chilean Andes is still insufficiently dense and large

uncertainties persist with respect to the spatial distribution of precipitation. As pointed out by Falvey and Garreaud (2007), the fact that the snow cover dominates the higher Cordillera throughout winter is a major impediment for collecting long-term precipitation data, a problem that will probably remain unsolved in the near future. As a result, even though previous research has intended to quantify the variation of precipitation amounts at a large-scale, confirming that precipitation fields are affected by elevation and that a spatial trend in long-term average precipitation exists (Garreaud and Rutlland, 1997; Falvey and Garreaud, 2007; Viale and Núñez, 2011), it is not possible to assertively describe this variability at a local scale. The choice of an interpolation method that allows a reliable estimation of these orography influenced precipitation fields in such poor information scenario is not trivial. In these cases, traditional precipitation interpolation methods utilized in engineering hydrology (e.g. Thiessen Polygons) are most probably unable to provide reliable estimations of precipitation amounts at different elevations.

Methods available for interpolation of precipitation gauge data include, for example, Thiessen Polygons, Classical Polynomial Interpolation, Inverse Distance Weighting, Multiquadric Functions fitting, Optimal Interpolation Method, and Kriging techniques (Sen, 2009). Studies have been conducted on long-term mean monthly and annual precipitation data (Goovaerts, 2000;

¹Pontificia Universidad Católica de Valparaíso, Facultad de Ingeniería, Av. Brasil 2147, Valparaíso, Chile.

*Corresponding author (alexandra.jacquin@ucv.cl).

Received: 1 April 2013.

Accepted: 7 October 2013.

doi:10.4067/S0718-58392013000400012

Subyani, 2004; Diodato and Ceccarelli, 2005; Feki et al., 2012), annual and monthly data (Chen et al., 2002; Ruelland et al., 2008), daily data (Haberlandt and Kite, 1998; Chang et al., 2005; Buytaert et al., 2006; Ahrens, 2006; Ruelland et al., 2008; Skaugen and Andersen, 2010), and also data from individual storms (Amani and Lebel, 1997; Syed et al., 2003; Buytaert et al., 2006). This research has revealed that the choice of an appropriate interpolation method depends on the time step used in the analysis, the density of the precipitation network, the precipitation regime and the catchment's geomorphologic characteristics. Most of the examples found in the literature use data from catchments with no significant elevation changes, while there are very few existing studies dealing with catchments subjected to strong orographic effects. Nevertheless, the results of these studies have shown that precipitation interpolation in this case benefits from choosing a method that makes explicit consideration of the existence of a relationship between precipitation and elevation (Goovaerts, 2000; Diodato and Ceccarelli, 2005; Buytaert et al., 2006; Feki et al., 2012). Unfortunately, previous research has mainly focused in areas where the precipitation network is relatively dense and the issue of what precipitation interpolation methods is more appropriate remains unresolved in the case of mountain catchments with scarce gauge data.

This study is concerned with the evaluation of the applicability of Kriging with External Drift (KED) and the technique known as Optimal Interpolation Method (OIM) for the estimation of monthly precipitation amounts in mountainous catchments with sparse precipitation networks. The methods KED and OIM are both able to account for the existence of a relationship between precipitation and elevation, but they differ in the manner in that this information is incorporated in the modeling process. In the case of KED, the modeler must make a prior assumption on the nature of the analytical expression between the long-term expectation of precipitation and elevation, but it is not necessary that this function is explicitly specified. By contrast, the application of OIM requires the explicit estimation of the long-term mean expectation and the standard deviation of precipitation at the location where interpolations are to be performed. Regression functions relating long-term mean precipitation and elevation, and sample standard deviation and elevation, can be used for this purpose. The choice of regression function types depends on the characteristics of these relationships, as revealed by field data. Precipitation estimates obtained with the Thiessen Polygons (TP) method, unable to explicitly account for the existence of a relationship between precipitation and elevation, are used as a benchmark. The study area corresponds to the upper subcatchment of Aconcagua River, located in the Andes of Central Chile, a mountainous zone with elevation ranging from 959 to 5930 m a.s.l. Given the strong elevation changes that exist in the study area, it is

an appropriate location for testing the ability of KED and OIM to perform the interpolation of monthly precipitation in mountainous catchments.

MATERIALS AND METHODS

Study area and data

The study area corresponds to the upper subcatchment of Aconcagua River (32°36' and 33°11' S, 69°59' and 70°32' W), in the Andes of Central Chile. The catchments of Aconcagua River at Chacabuquito and Juncal River at Juncal fluviometric station are used as case studies (Figure 1). Aconcagua River at Chacabuquito is a snowmelt dominated mountainous catchment with a surface of 2110 km², with elevation ranging from 950 m a.s.l. at the catchment outlet to 5930 m a.s.l. in the higher Cordillera. Juncal River is a catchment located between 2200 and 5930 m a.s.l, having an area of 235 km². Most of the upper Aconcagua's catchment is not inhabited and economic activity is limited to mining, but melting water that the catchment provides during the summer months is essential for the development of agricultural activity and the supply of drinking water to the population downstream Aconcagua River. Monthly data from nine precipitation stations in the period April 1965 to March 2001 are chosen (Table 1). Five of these precipitation stations (numbered 1 to 5 in Table 1) are located in the valley immediately downstream Chacabuquito fluviometric station. Four precipitation stations (numbered 6 to 9 in Table 1) are located in the Andes Mountains, actually within Aconcagua River at

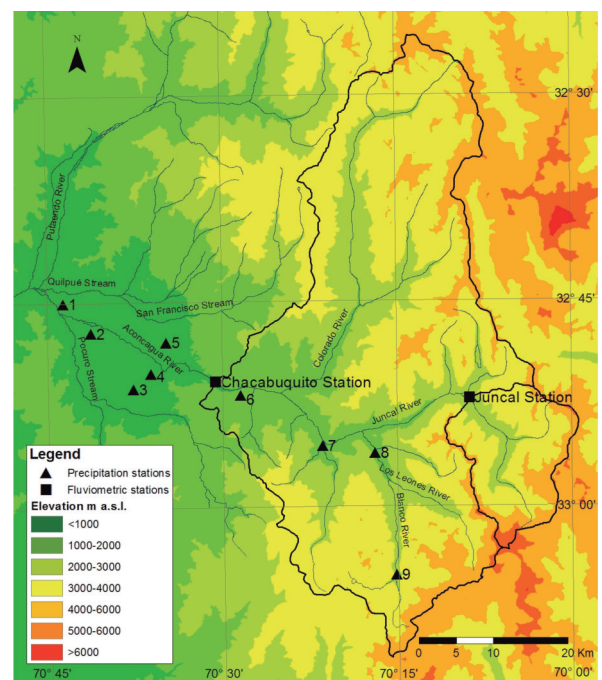


Figure 1. Catchment case study. Elevation information derived from ASTER GDEM Version 1 (METI/NASA, 2009). ASTER GDEM is a product of METI and NASA.

Table 1. Precipitation stations, location and long-term statistics of annual precipitation.

ID	Station	Owner	Latitude	Longitude	Elevation m a.s.l.	MAP ¹	SDAP ¹
			°S	°W		mm	
1	San Felipe	DGA	32.75	70.73	640	215.1	141.8
2	Curimón	DMC	32.79	70.69	712	219.4	148.8
3	CalleLarga	DMC	32.86	70.63	780	228.7	145.5
4	Los Andes	DGA	32.84	70.60	820	264.6	163.1
5	San Esteban	DMC	32.80	70.58	832	229.9	143.2
6	Vilcuya	DGA	32.86	70.47	1100	354.5	199.2
7	Rieccillos	DGA	32.93	70.36	1290	549.0	334.9
8	Saladillo	DMC	32.94	70.28	1580	560.2	364.9
9	Lagunitas	Codecco	33.08	70.25	2765	843.9	422.6

MAP: Mean annual precipitation; SDAP: Standard deviation of annual precipitation; DGA: Dirección General de Aguas; DMC: Dirección Meteorológica de Chile.

¹Statistics for the period April 1965-March 2001 (sample size of 36 yr).

Chacabuquito catchment area, with elevations ranging between 1100 and 2765 m a.s.l. None of them is located within Juncal River catchment. Although this information is probably insufficient to confidently perform catchment wide interpolations of precipitation fields in the higher Aconcagua, it is still relevant to assess to what extent precipitation estimates obtained with different methods are at least credible, bearing in mind that more gauge data are not available and that this poor information scenario is not unlikely to occur in water resources studies in the Chilean Andes.

Table 1 also shows the long-term mean and the sample standard deviation of annual precipitation at each station during the period of analysis. These values indicate a clear relationship between annual precipitation and elevation that becomes more evident above 1000 m a.s.l. In fact, Spearman's rank correlation coefficient between annual precipitation and elevation for the whole data set during the period of analysis equals 0.57. A similar conclusion can be derived from Table 2, which shows the long-term mean and standard deviation of monthly precipitation for the groups of stations 1 to 5 and for the group of stations 6 to 9. Throughout the hydrological year, long-term mean monthly values corresponding to the group of stations 1 to 5, located in the valley, are lower than those corresponding

Table 2. Long-term mean and sample standard deviation of monthly precipitation.

Month	Mean ¹		Standard deviation ¹	
	Valley	Andes	Valley	Andes
	mm			
Apr	9.8	30.0	14.9	46.4
May	32.3	80.0	33.2	70.5
June	57.2	139.7	70.2	165.5
July	63.3	144.2	78.6	181.6
Aug	35.6	80.3	41.5	98.5
Sep	16.6	46.0	19.9	54.0
Oct	7.7	22.0	11.3	27.5
Nov	4.5	14.0	10.5	24.7
Dec	0.9	5.8	3.1	14.1
Jan	0.7	5.4	2.0	12.5
Feb	0.7	3.9	3.1	8.8
Mar	2.2	5.7	4.5	9.5

¹Statistics for the period April 1965-March 2001 (sample size of 36 yr).

to the group of stations 6 to 9, located in the Andes. In addition to this, Table 2 shows that precipitation in the study area is highly seasonal. It concentrates between April and September, while precipitation amounts during the rest of the year are quite low. Falvey and Garreaud (2007) note that this seasonality in precipitation amounts is due to the dominance of the South Pacific Anticyclone during spring-summer, inhibiting the arrival of frontal systems, which is followed by a northwards shift of this high pressure centre during winter that favors the occurrence of precipitation events. Some storms due to convective activity are also observed between December and February in the higher Cordillera (Garreaud and Rutland, 1997).

Thiessen Polygons

The TP method uses precipitation amounts at the closest measuring station as estimations of precipitation P at the point of interest \vec{x}_0 . This is the most computationally inexpensive of all methods applied in this study, but it has the great disadvantage of being unable to explicitly account for the existence of a relationship between precipitation and elevation.

Kriging with External Drift

Studies dealing with the interpolation of precipitation amounts have shown that although these relatively complex methods do not always result in the best point estimates, the application of a Kriging method that is adequate to the problem at hand can produce robust interpolations of gauge data time series (Haberlandt and Kite, 1998; Ruelland et al., 2008). In Kriging methods, spatial distribution of precipitation is seen as a regionalized variable, that is, a variable that varies spatially in a structured and organized manner at a global scale, but which suffers erratic local variations (Émery, 2001). At a given time step, the precipitation field is a realization of a random function P that simultaneously assigns a precipitation value $P(\vec{x})$ to every point \vec{x} in the horizontal plane. The variogram of $P(\vec{x})$ is defined as:

$$\gamma(\vec{d}_{ij}) = \frac{1}{2} \text{VAR}[P(\vec{x}_i) - P(\vec{x}_j)] \quad [1]$$

which is assumed to depend only on the distance vector $\vec{d}_{ij} = \vec{x}_i - \vec{x}_j$ and not on the location of the points \vec{x}_i, \vec{x}_j . In this study, it is assumed that variograms of monthly precipitation are affected by anisotropies such that the variogram ordinates depend on the effective distance between \vec{x}_i and \vec{x}_j , defined as:

$$d_{ij} = \sqrt{(1-s) \{(x_i - x_j)^2 + (y_i - y_j)^2\} + s \cdot \lambda^2 \cdot (z_i - z_j)^2} \quad [2]$$

where (x,y) represent the horizontal coordinates and z represents the elevation of terrain point \vec{x} , and s ($0 < s < 1$) is a factor of vertical importance. A scale factor $\lambda = 10$ is introduced in order to transform vertical distances into a scale that is analogous to that of horizontal distances between data points.

In most kriging applications (Haberlandt and Kite, 1998; Goovaerts, 2000; Subyani, 2004; Ruelland et al., 2008; Feki et al., 2012), the ordinates of the experimental variogram are estimated according to the relationship:

$$\gamma^*(d) = \frac{1}{2N(d)} \sum_{i=1}^{N(d)} [P(\vec{x}_i + \vec{d}) - P(\vec{x}_i)]^2 \quad [3]$$

where $N(d)$ represents the number of pairs of data points that are located at a distance d . This approach is not appropriate for the inference of variograms in the study area for two reasons. Firstly, the precipitation network is so sparse that the number of pairs of data points with similar distances reduces to one in most cases. Secondly, it can be shown that the fact that the expectation of precipitation $m(\vec{x})$ varies space-wise causes that the estimation provided by Equation [3] is biased with respect to the true variogram of $P(\vec{x})$. An alternative strategy would be the calculation of the empirical variogram in terms of the residuals $Y(\vec{x}) = P(\vec{x}) - m(\vec{x})$, applying Equation [3] and substituting $P(\vec{x})$ by $Y(\vec{x})$. However, even though the experimental variogram of the residuals $Y(\vec{x})$ would provide an unbiased estimation of the variogram of $P(\vec{x})$, the experimental variogram of the estimated residuals $\hat{Y}(\vec{x}) = P(\vec{x}) - \hat{m}(\vec{x})$ underestimates the true variogram of $P(\vec{x})$. This underestimation arises from the fact that the true expectation $m(\vec{x})$ is unknown, implying that an estimation $\hat{m}(\vec{x})$ has to be used instead, and this bias would persist even if an optimal estimation procedure were used for the estimation of $\hat{m}(\vec{x})$ (Chilès and Delfiner, 1999; Émery, 2001). In order to circumvent these problems, and at the same time account for the seasonality of precipitation in the study area, it is postulated that precipitation of each month of the year is a different regionalized variable, of which there is available a number of realizations that is equal to the number of years T in the data set. For each month of the year, the ordinates of the sample variogram are estimated by means of the relationship:

$$\gamma^*(d_{ij}) = \frac{1}{2(T-1)} \sum_t [(P^t(\vec{x}_i) - \bar{P}(\vec{x}_i)) - (P^t(\vec{x}_j) - \bar{P}(\vec{x}_j))]^2 \quad [4]$$

which is essentially the same as that proposed by Tabios and Salas (1985), with the exception that the unbiased estimator of the variance $VAR[(P(\vec{x}_i) - P(\vec{x}_j))]$ is used herein. The terms $P^t(\vec{x}_i)$ and $P^t(\vec{x}_j)$ represent the observations at time step t , while and represent the long-term mean of the corresponding month of the year, at locations \vec{x}_i and \vec{x}_j , respectively. The final step in variographic analysis consists in fitting a theoretical model to the experimental variogram. A power model, with nugget effect where necessary, was considered satisfactory for modeling the variograms of monthly precipitation throughout the year. For example, Figure 2 shows the experimental variogram and the theoretical model fitted to April's data.

The KED method assumes that the external drift of precipitation $m(\vec{x})$ is a function of a smoothly varying secondary variable. A logarithmic trend model for the

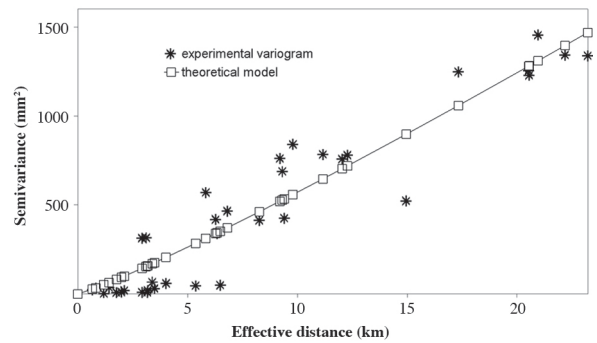


Figure 2. Experimental and theoretical variogram of monthly precipitation in April.

expectation of precipitation of each month of the year, given by:

$$m(\vec{x}) = \alpha_0 + \alpha_1 \ln(z(\vec{x})) \quad [5]$$

where $z(\vec{x})$ represents elevation at location \vec{x} , while α_0 and α_1 are constants, was found to provide an appropriate representation of the external drift of monthly precipitation in the study area, for all months of the year. Linear and parabolic relationships were also tested, but results indicate that a logarithmic relationship between the expectation of monthly precipitation and elevation is the most appropriate external drift model for the application of KED to the interpolation of monthly precipitation in the study area. Finally, precipitation at a given point of interest \vec{x}_0 is estimated according to:

$$P^*(\vec{x}_0) = \sum_{i=1}^n \lambda_i P(\vec{x}_i) \quad [6]$$

where the coefficients λ_i are obtained from the system of linear equations that results from the conditions of unbiasedness and minimal variance of the estimation (Émery, 2001).

Optimal Interpolation Method

Previous studies on interpolation of gauge data have applied the OIM to climatological precipitation records (Sen, 2009), large-scale monthly and daily data sets (Chen et al., 2002; Xie et al., 2007), and data from precipitation events (Creutin and Obled, 1982). It has been observed that this statistical method is able to provide credible large-scale interpolations of precipitation fields. Using the expectation of precipitation $m(\vec{x}_0)$ of the corresponding month of the year as an initial guess throughout the period of analysis, precipitation at a given time step can be estimated by means of the relationship:

$$P^*(\vec{x}_0) = m(\vec{x}_0) + \sum_{i=1}^n W_i \frac{P(\vec{x}_i) - m(\vec{x}_i)}{\sigma(\vec{x}_i)} \sigma(\vec{x}_0) \quad [7]$$

where $\sigma(\vec{x}_i)$ and $\sigma(\vec{x}_0)$ represent the standard deviation of precipitation at station \vec{x}_i and at location \vec{x}_0 , respectively. The interpolation weights W_i are obtained from the system of linear equations:

$$\sum_{i=1}^n W_i \rho_{ij} = \rho_{0j}, \quad j = 1 \dots n \quad [8]$$

where ρ_{ij} represents the correlation coefficient between stations \vec{x}_i and \vec{x}_j , and ρ_{0j} corresponds to the correlation coefficient between station \vec{x}_i and location \vec{x}_0 .

The application of OIM requires prior estimation of the expectation $m(\vec{x}_0)$, the standard deviation $\sigma(\vec{x}_0)$ and the correlation coefficients ρ_{0j} at the point of interest \vec{x}_0 , where precipitation data is not available. For this purpose, a logarithmic relationship between long-term mean precipitation in each month of the year and elevation was fitted, according to Equation [5]. A logarithmic relationship between standard deviation and elevation is also used for the estimation of $\sigma(\vec{x}_0)$. For example, Figure 3 presents the logarithmic curves fitted to data in the case of April's monthly precipitation. The correlation coefficients ρ_{0j} are estimated by fitting a negative-exponential model (Sen, 2009) to the spatial correlation data, as shown in Figure 4 for the case of April's data.

Computational experiments

The goodness of fit of monthly precipitation estimates is first evaluated by means of leave-one-out type cross-validation experiments. Goodness of fit statistics is obtained using purposely-built computing code developed in MATLAB® R2011. Firstly, root mean squared errors (RMSE) and mean errors are calculated over all precipitation stations available, during the period of analysis April 1965 to March 2001. A limitation of

cross-validation methods is that data must come from the same population, in order to obtain goodness of fit statistics that are reliable and unbiased. Time variability is not a serious concern herein, as preliminary analysis showed that the time series used in the experiments are not subjected to evident time trends and seasonality is dealt with by treating each month of the year separately. However, space variability poses difficulties, because the statistical properties of monthly precipitation are known to be affected by elevation. Accordingly, subsequent cross-validation experiments are conducted separately for the group of precipitation stations located in the valley immediately downstream Chacabuquito and for the group of stations located in the Andes. In this manner, the interpolation capability of each method is assessed in terms of its goodness of fit at each of these elevation zones. Orographic differences other than elevation are not considered for classifying the stations into groups. To start with, all of the stations whose data is used in the experiments are located in the windward side of the Andes Mountains. Other orographic differences are not considered either, because precipitation data in the Chilean Andes is so scarce that it is still unclear to what extent terrain characteristics such as slope, aspect and site exposure to prevailing winds may affect the local variability of precipitation in the study area.

Due to the limitations in the spatial extent of data, cross-validation experiments described earlier only provide information about the plausibility of precipitation estimates below 3000 m a.s.l., where only about 64% of Aconcagua River at Chacabuquito catchment area is located. With the aim of assessing the plausibility of the catchment wide spatial distribution of precipitation estimates prescribed by these methods, long-term annual water balances are performed in both catchment case studies. Yearly values of long-term mean real areal precipitation (RAP) can be roughly evaluated from the relationship:

$$RAP = Q/A + ET \quad [9]$$

where Q/A represents the long-term mean observed discharge per unit area at the catchment outlet's and ET represents long-term average yearly actual evapotranspiration in the catchment. Equation [9] relies on the assumptions that long-term water accumulation is small compared to Q/A and ET , and that losses due to groundwater flow are unimportant, which are reasonable considering that the Andes rocky soils are rather impermeable and vegetation is scarce (Falvey and Garreaud, 2007). Data used for the calculation of Q/A in the case of Aconcagua River at Chacabuquito catchment covers the totality of the period of analysis, from April 1965 until March 2001. However, data considered in the case of Juncal River catchment corresponds to the period April 1970 to March 2001, because discharge information at Juncal station is only available from April 1970. Long-term average yearly actual evapotranspiration estimates from the 5 arc minutes resolution map by FAO (2009)

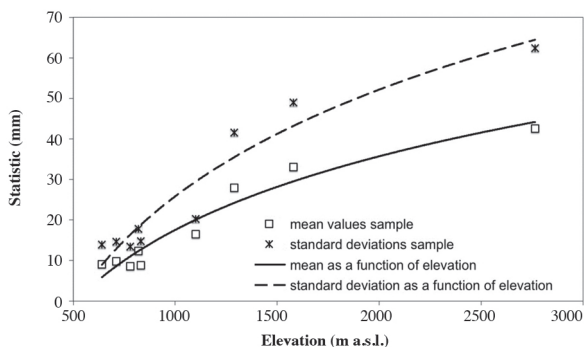


Figure 3. Long-term mean and standard deviation of monthly precipitation in April.

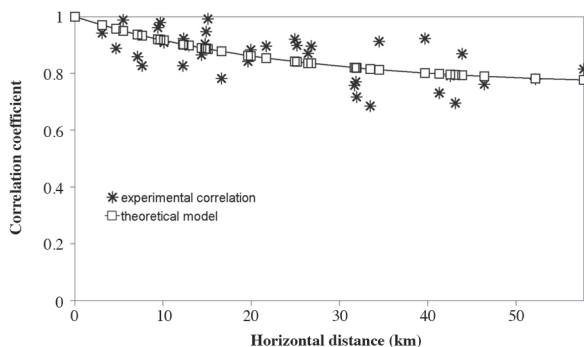


Figure 4. Spatial correlation function of monthly precipitation in April.

were used for the calculation of evapotranspiration losses. RAP values obtained with Equation [9] are compared with the long-term average yearly precipitation estimates provided by each method. For this purpose, precipitation estimates at each time step are calculated for a 500 m resolution grid, where the elevation of the grid nodes was derived from ASTER GDEM Version 1 (METI/NASA, 2009). Yearly amounts of estimated areal precipitation (EAP), averaged during the period April 1965-March 2001 in Aconcagua at Chacabucuito and during the period April 1970-March 2001 in Juncal River catchment, are subsequently calculated.

RESULTS AND DISCUSSION

Cross-validation experiments

Table 3 shows the root mean squared errors (RMSE) and the mean errors calculated considering all nine precipitation stations available. Positive and negative mean errors indicate overestimation and underestimation of precipitation amounts, respectively. In terms of this average goodness of fit statistics, it can be observed that, throughout the year, the OIM generally exhibits the best fit to the observations. In particular, the small mean errors obtained by the OIM are due to the good agreement between empirical mean monthly precipitation values and the logarithmic curves used for the estimation of expectation as a function of elevation (Figure 3). Table 3 reveals that, considering the totality of the rain gauges, the interpolation methods TP and KED obtain similar RMSE. Nevertheless, the mean errors reported in Table 3 show that the KED method has a substantially smaller bias than TP, especially during the wet season April to September. As explained earlier in the Material and Methods section, these results must be taken with caution, as the space variability of the statistical properties of precipitation limits the validity of goodness of fit statistics calculated over all precipitation stations.

Cross-validation experiments conducted separately

Table 3. Error statistics calculated considering all precipitation stations available.

Month	Root mean squared error			Mean error		
	TP	KED	OIM	TP	KED	OIM
	mm					
Apr	19.0	20.8	16.2	-1.5	0.7	0.0
May	33.5	35.2	27.4	-3.5	1.1	0.1
June	68.6	73.9	50.5	-13.3	-6.0	0.0
July	50.6	59.8	44.7	-8.9	-1.2	0.3
Aug	27.8	29.0	24.9	-4.9	-0.3	0.1
Sep	29.4	27.0	22.3	-4.6	-1.3	0.0
Oct	13.2	12.1	13.1	-1.4	0.5	0.0
Nov	8.1	9.5	7.3	-0.4	0.8	0.0
Dec	7.9	7.9	6.1	-0.8	-0.3	0.2
Jan	9.6	9.7	6.4	-0.9	-0.5	0.4
Feb	5.3	4.2	4.0	-0.9	-0.2	0.2
Mar	5.8	5.2	4.5	-0.5	-0.2	0.1

TP: Thiessen Polygons; KED: Kriging with External Drift; OIM: Optimal Interpolation Method.

for the group of stations in the valley immediately downstream Chacabucuito fluviometric station and those located in the Andes provide information on the ability of the interpolation methods under scrutiny for producing monthly precipitation estimates at different elevation zones. Figures 5 and 6 show RMSE values for the group of stations located in the valley and for the group of stations located in the Andes, respectively. Figure 5 reveals that the RMSE obtained by all methods are very similar in the case of the stations located in the valley, but Figure 6 shows that this situation changes in the case of the stations located in the Andes. Even though the KED method is able to account for the existence of an external drift in precipitation, its goodness of fit is comparable to that of the simpler TP method in terms of RMSE at high elevations. By contrast, Figure 6 shows that the monthly precipitation estimates provided by the OIM have a much better agreement than TP and KED with the gauge data available in the Andean region of the study area, especially during the wet season April to September. In this period, the RMSE obtained by the OIM are between 14 and 33% smaller than those of TP and KED in the Andes stations.

Figure 7 and Figure 8 show mean errors for the group of stations located in the valley and for the group of stations located in the Andes, respectively. Unlike the case of the RMSE, mean errors obtained by interpolation methods in the valley stations are clearly distinguishable. Figure 7 reveals that TP generally overestimates mean precipitation in the valley stations; absolute values of these mean errors

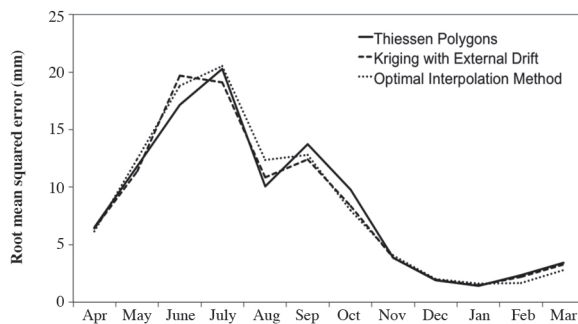


Figure 5. Root mean squared errors for the group of stations located in the valley.

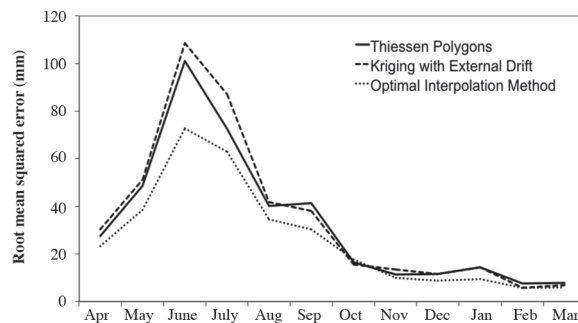


Figure 6. Root mean squared errors for the group of stations located in the Andes.

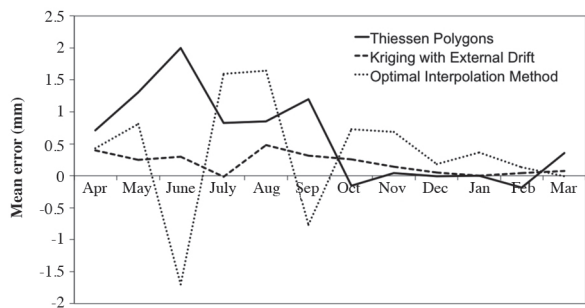


Figure 7. Mean errors for the group of stations located in the valley.

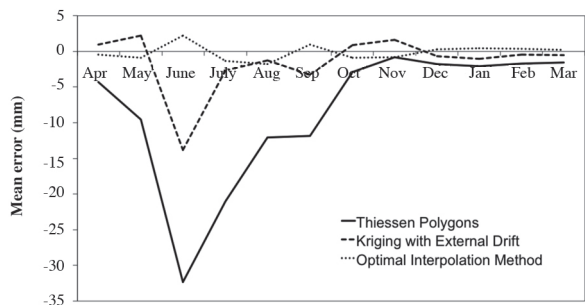


Figure 8. Mean errors for the group of stations located in the Andes.

generally remain below 10% of mean monthly values reported in Table 2, but they reach 26% and 16% of mean monthly precipitation in February and March, respectively. Mean errors of the KED method in valley stations are also positive, but smaller and quite stable throughout the year. Most notably, these errors are smaller than 6% of the corresponding mean monthly precipitation amounts presented in Table 2. In this zone, OIM shows an erratic behavior, obtaining mean errors whose absolute values are comparable to those of TP in some months. The impact of the choice of interpolation method in the bias of the precipitation estimates is even greater at high elevations. As expected, the mean errors shown in Figure 8 reveal that the TP method notably underestimates mean monthly precipitation in the case of stations located in the Andes. Between April and September, absolute values of these mean errors vary between 6% and 44% of the respective mean monthly values reported in Table 2. These results show that TP is likely to produce large negative biases when used for the estimation of monthly precipitation in the Andean Mountains. The KED method provides a better estimate of mean monthly precipitation in this zone, with mean errors smaller than 19% in absolute value in all months. Finally, the application of the OIM generally results in the smallest bias in the case of the Andes stations, with mean errors whose absolute values are smaller than 9% of mean monthly precipitation.

Catchment wide interpolation of precipitation fields

Figure 9 shows long-term annual precipitation estimates averaged at different elevation zones in Aconcagua at

Chacabuquito catchment. It can be observed that the annual precipitation obtained by the TP method, whose precipitation estimates at each time step are bounded above by the largest observation among the gauging stations available, reach 843 mm yr⁻¹ in the higher elevation zones, which corresponds to annual precipitation at Lagunitas (Table 1), as expected. In addition to this, Figure 9 shows that the KED method and OIM provide similar estimations of long-term annual precipitation corresponding to different elevations.

According to the information provided by FAO (2009), long-term average actual evapotranspiration is close to 311 mm yr⁻¹ in Aconcagua River at Chacabuquito catchment and close to 325 mm yr⁻¹ in Juncal River catchment. Considering that the mean discharge per unit area between April 1965 and March 2001 equals 500 mm yr⁻¹ at Chacabuquito station, RAP over the catchment must have been close to 811 mm yr⁻¹ in that period. Similarly, the mean discharge per unit area observed at Juncal station in the period April 1970 to March 2001 equals 808 mm yr⁻¹, implying that RAP in this catchment must have been close to 1133 mm yr⁻¹.

Table 4 shows estimated annual precipitation (EAP) obtained using the interpolation methods applied in this study, in addition to the quotients between EAP and RAP, both for Aconcagua River at Chacabuquito and Juncal River catchment. These results show that TP largely underestimate yearly areal precipitation over both catchments, which is not surprising considering the cross-validation results shown in Figure 8 and discussed earlier. The low ratios EAP/RAP shown in Table 4, and the fact that these quotients in the case of Juncal River catchment

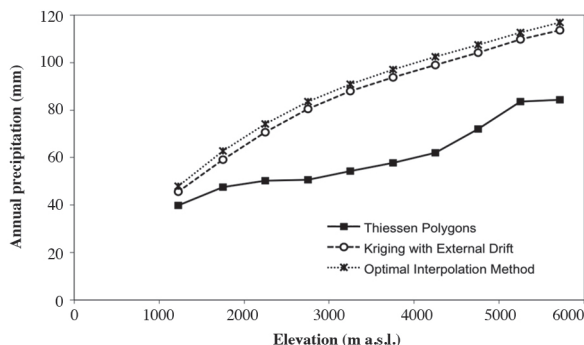


Figure 9. Average yearly precipitation estimates in different elevation zones of Aconcagua River at Chacabuquito catchment.

Table 4. Estimated annual precipitation of each method, and quotient between this estimation and real annual precipitation in the catchment case studies.

Month	Aconcagua at Chacabuquito		Juncal at Juncal	
	EAP (mm)	EAP/RAP	EAP (mm)	EAP/RAP
TP	544	0.67	628	0.55
KED	844	1.04	988	0.87
OIM	876	1.08	998	0.88

EAP: Estimated annual precipitation; RAP: Real annual precipitation; TP: Thiessen Polygons; KED: Kriging with External Drift; OIM: Optimal Interpolation Method.

(55%) are much lower than those obtained in Aconcagua River at Chacabuquito catchment (67%), suggest that the performance of these methods rapidly deteriorates at higher elevations. By contrast, the methods KED and OIM are both able to provide realistic estimates of long-term yearly areal precipitation in these mountain catchments. The ratios EAP/RAP obtained by KED and OIM in Aconcagua River at Chacabuquito catchment are very close to unity, implying the estimation of EAP is almost unbiased. In the more challenging case of Juncal River catchment, these ratios decrease to about 90%. This is not a surprising result, considering that the precipitation data used for the derivation of EAP values does not include any gauge that is actually located inside this latter catchment.

CONCLUSIONS

Cross-validation experiments revealed that the performance of all interpolation methods analyzed is similar in the case of the group of precipitation stations located in the valley zone of the study area. Considering that the external drift affecting precipitation in the study area becomes evident at higher elevations, it is not surprising that even the simplistic Thiessen Polygons (TP) method is able to provide reasonable estimations of monthly precipitation in the valley. By contrast, the Optimal Interpolation Method (OIM) notably outperforms TP and Kriging with External Drift (KED) for the interpolation of monthly precipitation in the case of the group of stations located in the Andes, producing precipitation estimates generally subjected to the smallest bias and showing a better agreement to gauge data in terms of root mean squared errors. Although KED achieves a much better agreement to mean monthly values than TP at high elevations, its performance is close to that of the simpler method in terms of root mean squared errors. Simplified long-term annual water balances did not provide evidence against the applicability of KED or the OIM. However, considering the performance of these methods in the cross-validation experiments, it can be concluded that the OIM is a more suitable than KED for the interpolation of monthly precipitation amounts in the study area.

ACKNOWLEDGEMENTS

This research was funded by Fondo Nacional de Desarrollo Científico y Tecnológico, FONDECYT, Research Project 1110279. The authors would like to express their gratitude to Dirección General de Aguas (DGA), Chile, División Andina de Codelco and Dirección Meteorológica de Chile (DMC) for providing the data used in this study.

LITERATURE CITED

Ahrens, B. 2006. Distance in spatial interpolation of daily rain gauge data. *Hydrology and Earth System Sciences* 10:197-208.

- Amani, A., and T. Lebel. 1997. Lagrangian kriging for the estimation of Sahelian rainfall at small time steps. *Journal of Hydrology* 192:125-157.
- Buytaert, W., R. Celleri, P. Willems, B. De Bièvre, and G. Wyseure. 2006. Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *Journal of Hydrology* 329:413-421.
- Chang, C.L., S.L. Lo, and S.L. Yu. 2005. Applying fuzzy theory and genetic algorithm to interpolate precipitation. *Journal of Hydrology* 314:92-104.
- Chen, M., P. Xie, and J.E. Janowiak. 2002. Global land precipitation: A 50-yr monthly analysis based on gauge observations. *Journal of Hydrometeorology* 3:249-266.
- Chilès, J.P., and P. Delfiner. 1999. *Geostatistics: Modeling spatial uncertainty*. Wiley Series in Probability and Statistics. 695 p. Wiley, New York, USA.
- Creutin, J.D., and C. Obléd. 1982. Objective analyses and mapping techniques for rainfall fields an objective comparison. *Water Resources Research* 18:413-431.
- Diodato, N., and M. Ceccarelli. 2005. Interpolation processes using multivariate geostatistics for mapping of climatological precipitation mean in the Sannio Mountains (southern Italy). *Earth Surface Processes and Landforms* 30:259-268.
- Émery, X. 2001. *Géostatistique linéaire*. 405 p. Centre de Géostatistique, École de Mines de Paris, Fontainebleau, France.
- Falvey, M., and R. Garreaud. 2007. Wintertime precipitation episodes in Central Chile: Associated meteorological conditions and orographic influences. *Journal of Hydrometeorology* 8:171-193.
- FAO. 2009. *Global map of yearly actual evapotranspiration-5 arc minutes*. Food and Agriculture Organization (FAO), Rome, Italy.
- Feki, H., M. Slimani, and C. Cudennec. 2012. Incorporating elevation in rainfall interpolation in Tunisia using geostatistical methods. *Hydrological Sciences Journal* 57:1294-1314.
- Garreaud, R., y J. Rutland. 1997. Precipitación estival en Los Andes de Chile Central: aspectos climatológicos. *Atmósfera* 10:191-211.
- Goovaerts, P. 2000. Geostatistical approaches for incorporating elevation into spatial interpolation of rainfall. *Journal of Hydrology* 228:113-129.
- Haberlandt, U., and G.W. Kite. 1998. Estimation of daily space-time precipitation series for macroscale hydrological modelling. *Hydrological Processes* 12:1419-1432.
- METI/NASA. 2009. *Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) Version 1*. Ministry of Economy, Trade and Industry of Japan (METI) and the National Aeronautics and Space Administration (NASA), USA.
- Ruelland, D., S. Ardoin-Bardin, G. Billen, and E. Servat. 2008. Sensitivity of a lumped and semi-distributed hydrological model to several methods of rainfall interpolation on a large basin in West Africa. *Journal of Hydrology* 361:96-117.
- Sen, Z. 2009. *Spatial modeling principles in Earth sciences*. 351 p. Springer, Dordrecht, Netherlands.
- Skaugen, T., and J. Andersen. 2010. Simulated precipitation fields with variance-consistent interpolation. *Hydrological Sciences Journal* 55:676-686.
- Subyani, A.M. 2004. Geostatistical study of annual and seasonal mean rainfall patterns in southwest Saudi Arabia. *Hydrological Sciences Journal* 49:803-817.
- Syed, K.H., D.C. Goodrich, D.E. Myers, and S. Sorooshian. 2003. Spatial characteristics of thunderstorm rainfall fields and their relation to runoff. *Journal of Hydrology* 271:1-21.
- Tabios, G., and J.D. Salas. 1985. A comparative analysis of techniques for spatial interpolation of rainfall. *Journal of the American Water Resources Association* 21:356-380.
- Viale, M., and M.N. Núñez. 2011. Climatology of winter orographic precipitation over the Subtropical Central Andes and associated synoptic and regional characteristics. *Journal of Hydrometeorology* 12:481-507.
- Xie, P., A. Yatagai, M. Chen, T. Hayasaka, and Y. Fukushima. 2007. A Gauge-based analysis of daily precipitation over East Asia. *Journal of Hydrometeorology* 8:607-626.