



Centro del Agua para Zonas  
Áridas y Semiáridas de  
América Latina y El Caribe



**US Army Corps  
of Engineers®**

institute for  
water resources



Organización  
de las Naciones Unidas  
para la Educación,  
la Ciencia y la Cultura



Programa Hidrológico Internacional



## **“PREPARACION DE DATOS PARA EL ANALISIS REGIONAL DE FRECUENCIAS: AREA PILOTO IV REGION DE COQUIMBO, CHILE”**

### **Informe Final**

**Proyecto realizado en el marco de UNESCO-PHI y del Gobierno de Flandes,  
Departamento de Ciencias e Innovaciones**

Desarrollado por:  
Centro del Agua para Zonas Áridas y Semiáridas de ALC – CAZALAC (Chile)  
En el marco del proyecto CAZALAC-IWR: Atlas de Sequía mediante Análisis Regional  
de Frecuencias en áreas piloto de Argentina, Chile y Perú

**La Serena, Chile 2008**



# Introducción

La disponibilidad de datos confiables es un requisito indispensable para cualquier tipo de análisis hidrológico, en particular, aquel relacionado con la determinación de probabilidades de ocurrencia de eventos extremos, en la que el error, presencia o ausencia de un dato en particular, puede tener fuertes consecuencias sobre las estimaciones de los cuantiles con altos periodos de retorno. Por lo mismo, los esfuerzos destinados a la recopilación y análisis de los datos, así como su corrección en la etapa previa a su incorporación en los procedimientos de análisis de frecuencias, se justifican plenamente al alcanzar estimaciones más razonables de los parámetros a utilizar como de las estimaciones de cuantiles y productos finales.

A continuación, se presenta al lector los resultados de la fase de Preparación de Datos, en conformidad con las recomendaciones del “Manual Breve de Preparación de Datos” como parte de las actividades del proyecto Atlas de Sequía, Area Piloto de la IV Región de Coquimbo.

## Etapa 3.1 del Manual Breve

### Objetivo de la etapa:

El objetivo de esta etapa, en la preparación de datos, es conocer estadísticamente y gráficamente los datos con los cuales uno cuenta. Se puede obtener información valiosa sobre promedios, variabilidad, distribución, existencia de un patrón común, presencia de valores que pueden causar sospechas de ser incorrectos, etc.

Para llevarla a cabo es posible recurrir a cualquier paquete estadístico que permita obtener estadísticas descriptivas y gráficas estadísticas.

La etapa 3.1, según se describe en el Manual Breve de Preparación de Datos, en adelante Manual Breve, se basa en el denominado Análisis Exploratorio de Datos y fue desarrollada siguiendo las pautas del mismo, así como de la bibliografía recomendada. La metodología, presentación de resultados y su análisis se discuten a continuación:

### **1.1 Metodología**

#### 1.1.1. Material:

##### 1.1.1.1. De los datos:

Se dispuso de registros de datos de precipitación diaria, en unidades de milímetros, obtenidos de la base de datos de la Dirección General de Aguas para estaciones meteorológicas de la IV Región. También, se dispuso de información sobre datos de precipitación mensual para 2 estaciones meteorológicas de la región, obtenidas del estudio CAZALAC (2006). En el Cuadro 1 se presentan algunas de las características de las estaciones y datos utilizados. Por otro lado, en la Figura 1 se presenta la ubicación de las estaciones meteorológicas consideradas en el análisis en el entorno regional. Hubo una serie de estaciones meteorológicas que, aunque registradas, fueron descartadas del proceso de análisis por no contar con el mínimo de 15 datos en el registro de la serie de precipitación anual o no tenerlas disponibles en el momento del estudio. Las mismas son indicadas en la Figura 1.

Cuadro 1. Variables descriptivas de las estaciones meteorológicas utilizadas en el análisis

Estacion	UTM N	UTM E	Altitud	Año inicial	Año final	Tipo de dato original	Fuente
ALMENDRAL	6681587	314896	370.00	1959	2006	PPD	DGA
CAREN	6585105	330669	740	1944	2006	PPD	DGA
COCHIGUAZ	6664665	364697	1560	1990	2006	PPD	DGA
COGOTI 18	6559446	313978	840	1943	2006	PPD	DGA
COGOTI EMBALSE	6567617	300887	740	1943	2006	PPD	DGA
COIRON	6469012	332538	840	1974	2006	PPD	DGA
COMBARBALA	6549382	309308	870	1977	2006	PPD	DGA
CUNCUMEN	6465733	347486	1100	1959	2006	PPD	DGA
EL TOME	6588881	311496	420	1966	2006	PPD	DGA
EL TRAPICHE	6748770	294413	300	1979	2006	PPD	DGA
HUANTA	6697229	366231	1240	1990	2006	PPD	DGA
HURTADO	6648182	336867	1100	1943	2006	PPD	DGA
JUNTAS	6683217	394411	2150	1990	2006	PPD	DGA
LA LAGUNA EMBALSE	6658175	399678	3160	1964	2006	PPD	DGA
LA ORTIGA	6658753	357334	1560	1979	2006	PPD	DGA
LA PLACILLA	6580441	279417	600	1990	2006	PPD	DGA
LA SERENA	6689401	282215	15	1972	2003	PPM	CAZALAC (2006)
LA TORRE	6610458	272404	120	1943	2006	PPD	DGA
LA TRANQUILLA	6469378	342019	1000	1966	2006	PPD	DGA
LAS BREAS	6639123	344987	1600	1944	1977	PPD	DGA
LAS BURRAS	6509830	327099	1150	1990	2006	PPD	DGA
LAS RAMADAS	6567266	348619	1380	1944	2006	PPD	DGA
LOS NICHOS	6663873	355688	1330	1977	2006	PPD	DGA
LIMAHUIDA	6485742	295198	300	1964	2006	PPD	DGA
MAL PASO	6484903	301139	375	1961	1988	PPD	DGA
MONTE GRANDE	6670313	356032	1120	1959	2006	PPD	DGA
OVALLE DGA	6612583	289042	220	1972	2006	PPD	DGA
PABELLON	6634584	350683	1920	1969	2006	PPD	DGA
PALOMA EMBALSE	6602287	304978	320	1944	2006	PPD	DGA
PEDREGAL	6585291	336698	880	1968	1988	PPD	DGA
PENA BLANCA	6578642	256232	360	1991	2006	PPD	DGA
PICHASCA	6636222	320546	725	1947	2006	PPD	DGA
PISCO ELQUI DMC	6666743	356133	1250	1977	2006	PPD	DGA
PUNITAQUI	6587267	284780	280	1962	2006	PPD	DGA
RAPEL	6599909	330009	870	1969	2006	PPD	DGA
RECOLETA EMBALSE	6623114	298493	350	1944	2006	PPD	DGA
RIVADAVIA	6682664	349330	820	1954	2006	PPD	DGA
SALAMANCA	6482674	313736	510	1974	2006	PPD	DGA
SAMO ALTO	6634199	313772	600	1970	1988	PPD	DGA
SAN AGUSTIN	6488102	326461	1050	1945	2006	PPD	DGA
SOTAQUI	6609375	296857	280	1955	2006	PPD	DGA
TASCADERO	6567461	340924	1230	1961	2006	PPD	DGA
TULAHUEN	6572676	331530	1020	1950	2006	PPD	DGA
VICUÑA	6673645	334506	730	1971	2003	PPM	CAZALAC (2006)



#### 1.1.1.2. De las herramientas de análisis:

Para el procesamiento de los datos se utilizaron el software Minitab v.15 en español, la planilla Excel <sup>TM</sup> , la planilla Excel Makesens para el análisis de tendencia (Salmi *et al*, 2002)

#### 1.1.2. Métodos:

##### 1.1.2.1. Generación de la serie hidrológica mensual:

A partir de los datos de precipitación diaria, mediante la herramienta de Tablas Dinámicas de la planilla Excel, se calcularon las precipitaciones medias mensuales para cada año y estación. Previo al cálculo se inspeccionó visualmente los registros a modo de constatar que no existieran datos ausentes de precipitación diaria dentro de un mes dado. En caso de presentar datos diarios ausentes, el mes correspondiente fue considerado como nulo y quedó calificado como dato ausente.

Con este procedimiento, se obtuvieron las series de datos mensuales originales y no rellenadas.

##### 1.1.2.2. Generación de la serie hidrológica anual:

Algunos análisis requirieron de la utilización de valores de precipitación media anual. Para tal efecto, se determinó, por suma de las series de datos mensuales, los valores de precipitación media anual. En caso de tener menos de 12 meses de registro para una estación y un mismo año, el valor de precipitación anual se consideró como dato ausente.

##### 1.1.2.2. Análisis exploratorio:

El análisis exploratorio consideró la generación de un conjunto de indicadores estadísticos descriptivos, así como gráficas de control. Ambas aproximaciones se utilizaron tanto para las series de precipitación mensual, para cada estación, como para las series anuales entre estaciones. En esta etapa, entonces, se obtuvieron indicadores de tendencia central y dispersión así como gráficas de caja (box plots), histogramas de frecuencia y gráficas de valores puntuales.

## 1.2 Resultados:

### 1.2.1. Estadísticas descriptivas:

En Anexo 1 se presentan las estadísticas descriptivas por estación, para la variable precipitación media mensual. En el Cuadro 2, como ejemplo, se presentan las estadísticas descriptivas de la Estación El Trapiche.

Cuadro 2. Estadísticas descriptivas Estación El Trapiche. IV Región.

Variable	Conteo			Media	Desv.Est.	Mínimo	Máximo	Sesgo	Kurtosis
	total	N	N*						
Enero	71	29	42	0.000000	0.000000	0.000000	0.000000	*	*
Febrero	71	29	42	0.000000	0.000000	0.000000	0.000000	*	*
Marzo	71	29	42	0.841	3.507	0.000	18.000	4.63	22.29
Abril	70	28	42	1.054	4.360	0.000	23.000	5.08	26.35
Mayo	70	28	42	6.07	11.97	0.00	41.00	2.07	3.27
Junio	70	28	42	10.97	24.21	0.00	98.00	2.62	6.53
Julio	70	28	42	18.01	28.93	0.00	112.50	2.02	3.82
Agosto	70	27	43	8.06	15.07	0.00	71.50	3.18	12.13
Septiembre	70	26	44	2.004	4.030	0.000	16.000	2.56	6.53
Octubre	70	27	43	1.074	4.212	0.000	20.500	4.28	19.01
Noviembre	70	28	42	0.000000	0.000000	0.000000	0.000000	*	*
Diciembre	70	28	42	0.000000	0.000000	0.000000	0.000000	*	*

La inspección de estas cifras permitió identificar las siguientes situaciones:

- Existen una serie de estaciones, incluidas la Estación El Trapiche, que presentan valores medios mensuales de precipitación igual a 0 en meses que van desde Diciembre a Febrero según el Cuadro 3 a continuación:

Cuadro 3. Distribución de estaciones con media mensual y desviación estándar igual a 0 según mes.

Mes	Nov	Dic	Ene	Feb
N° Estaciones	1	8	9	7
% Estaciones	2.3	18.2	20.5	15.9

Este fenómeno tiene implicancias desde el punto de vista del cálculo del sesgo y la curtosis, valores que son indefinidos frente a la presencia de una desviación estandar de 0. Los L-momentos asociados a estos meses en las respectivas estaciones no podrán ser calculados, afectando la determinación de las curvas de distribución de probabilidad para dichos meses. Para el resto de los meses y todas las estaciones, no se presentarían estos problemas de cálculo.

- En la totalidad de las estaciones, existe el siguiente patrón común: La media y la desviación estándar aumentan en los meses lluviosos, siendo mayor entre mayo y julio, mientras que el sesgo y la curtosis disminuyen (aunque siempre con valor

positivo), siendo menores entre mayo y julio. Se puede interpretar, entonces, que los montos de precipitación aumentan en invierno, así como su variabilidad interanual. Sin embargo, la distribución es más simétrica en la estación húmeda que en los meses secos, justamente por la mayor variabilidad de las precipitaciones, que aportan con una mayor cantidad de datos a la curva de distribución. En cambio, en los meses secos, prácticamente la media es cercana a 0, con mínima variabilidad, pero basta la ocurrencia de muy pocos eventos de precipitación en estos meses para que extiendan la cola de la distribución hacia la derecha, sesgando significativamente la curva de distribución de frecuencias. Se presume que un cálculo más detallado de los L-momentos en el proceso de regionalización, debiese considerar la separación de meses secos respecto de meses húmedos por la posibilidad de que las curvas de distribución de frecuencias sean distintas. Este mejor conocimiento obtenido al hacer el ajuste para distintas épocas se realiza en función de un mayor esfuerzo de cálculo.

### 1.2.2. Gráfica de caja:

En el Anexo 2 se presentan las gráficas de caja de las precipitaciones medias mensuales por estación. En la Figura 2 siguiente, se presenta el análisis para la Estación Cogotí 18, tomada como ejemplo.

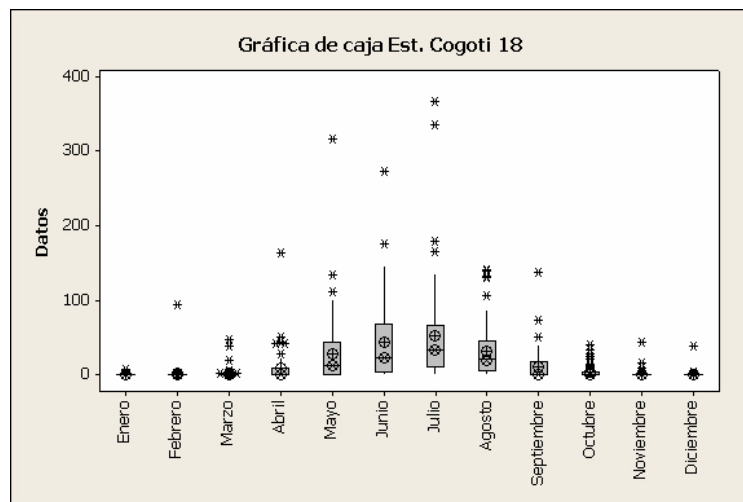


Figura 2. Gráfica de caja de la precipitación mensual Estación Cogotí 18. IV Región.

El patrón es concordante con los valores obtenidos en el punto 1.2.1. precedente. La media, así como la dispersión aumentan en los meses lluviosos, siendo junio o julio los meses más lluviosos en la Estación Cogotí 18 así como en la mayoría de las estaciones. Por otro lado, se hace más notorio en los meses lluviosos la presencia de valores por sobre 1.5 veces el rango intercuartil (valores que podrían ser considerados inicialmente como atípicos potenciales o outliers).



#### 1.2.4. Estadísticas descriptivas de la precipitación anual de la IV Región:

El Cuadro 4 presenta las estadísticas descriptivas de la precipitación anual incluidas todas las estaciones de la región. En tanto, en la Figura 4, se presentan las gráficas de frecuencia y caja de la misma variable.

Cuadro 4. Estadísticas descriptivas de precipitación anual IV Región de Coquimbo.

Conteo									
Variable	total	N	Media	Desv.Est.	Mínimo	Mediana	Máximo	Sesgo	Kurtosis
Media Anual	44	44	166.9	66.4	50.6	161.3	328.1	0.44	-0.57

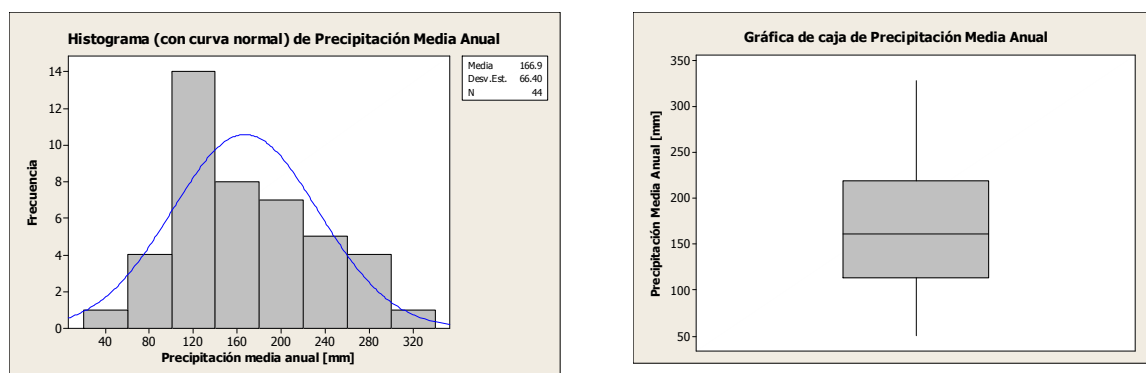


Figura 4. a) Histograma de frecuencias con curva normal de Precipitación Anual de la IV Región. B) Gráfica de caja de Precipitación Anual de la IV región.

De acuerdo con el Cuadro 4, la media de la precipitación anual de la IV Región, con base a la información proveniente de 44 estaciones, es de 166.9 mm, con una mediana ligeramente inferior, de 161.3 mm. Así también, el sesgo es positivo, lo que indica que la distribución presenta una asimetría hacia la derecha. Esta situación es confirmada por el histograma de frecuencias y la gráfica de caja de la Figura 4, donde, sin embargo, la distribución presenta una mayor simetría al compararla con los histogramas de frecuencias de las estaciones en forma individual. Sin embargo, debe tenerse presente que en este cálculo se han considerado estaciones con longitud muy variable de registros.

#### 1.2.5. Análisis de valores máximos:

Al analizar solo los valores máximos de las series de precipitación anual, es decir, la máxima precipitación anual registrada en cada estación, se obtienen los resultados presentados en el Cuadro 5.

Cuadro 5. Estadísticas descriptivas de las precipitaciones anuales máximas.

Media	Desv.Est.	Mínimo	Mediana	Máximo	Sesgo	Kurtosis
495.6	199.0	178.5	443.5	1056.5	0.83	0.19

Este análisis se realiza para mejorar el conocimiento sobre valores máximos de precipitación anual en todas las estaciones para ayudar en el criterio de considerar esos valores máximos como potencialmente atípicos o no. Los valores de sesgo y curtosis, así como un valor de la mediana menor a la media indican una distribución asimétrica y sesgada positivamente. La Figura 5 presenta la gráfica de caja de los valores máximos anuales para todas las estaciones.

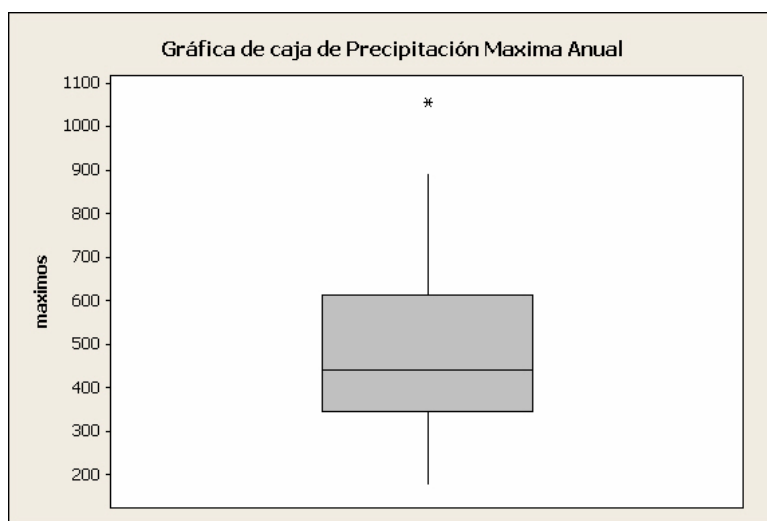


Figura 5. Gráfica de caja de Precipitación Máxima Anual para todas las estaciones.

Se aprecia la existencia de un valor atípico o outlier potencial, equivalente a 1056.5 mm correspondientes a la estación Coirón durante el año 1987. La eliminación de esta estación en el análisis contribuye a mejorar los indicadores de simetría de la distribución. Es decir, si consideramos solo las precipitaciones máximas, existe una estación cuya máxima precipitación anual registrada no se adecua a la distribución del resto de las estaciones. Esto sugiere que probablemente esta estación, así como las ubicadas en el extremo sur-este de la región pudieran presentar más homogeneidad con estaciones vecinas de la Quinta Región, de modo que su inclusión en el análisis regional de frecuencias se justifica en el análisis de los bordes, tal como lo indicara Wallis (2008).

### 1.2.5. Histograma de frecuencias:

En la Figura 6a. se presentan los histogramas de frecuencias de la precipitación anual para las estaciones Almendral a Las Burras, en tanto en la Figura 6b. se presentan las gráficas para las estaciones entre Las Ramadas y Vicuña. Se aprecia que las distribuciones en todas las estaciones tienen una distribución no normal, con un sesgo y curtosis positivas.

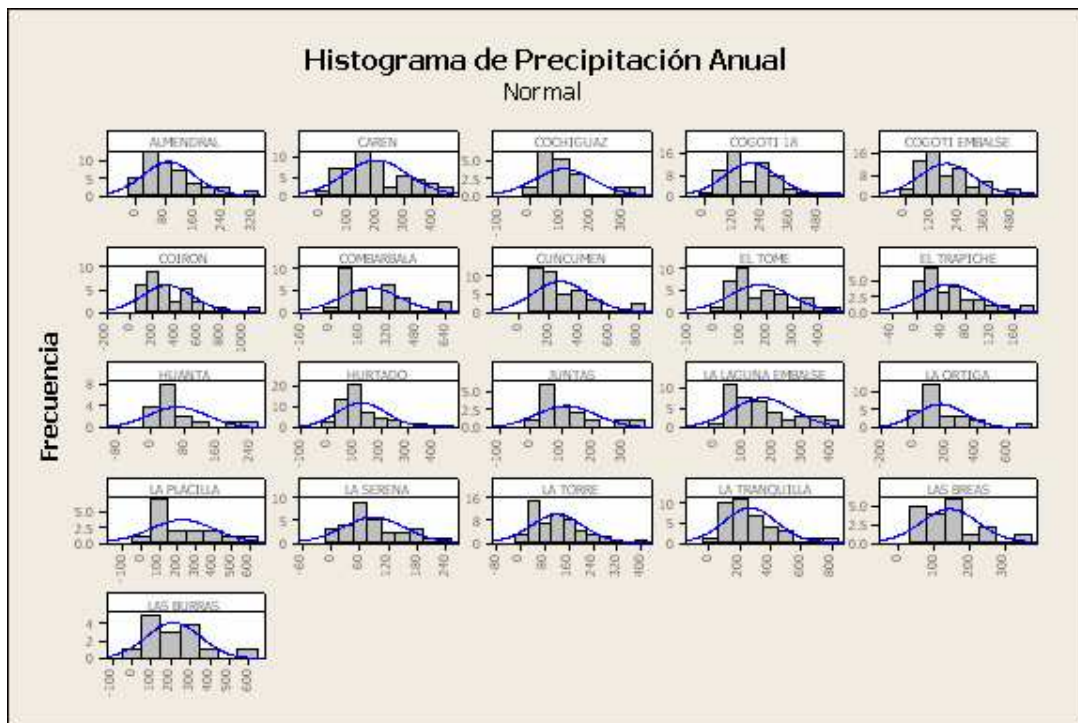


Figura 6a. Histograma de frecuencias de la precipitación anual según estación: Almendral a Las Burras.

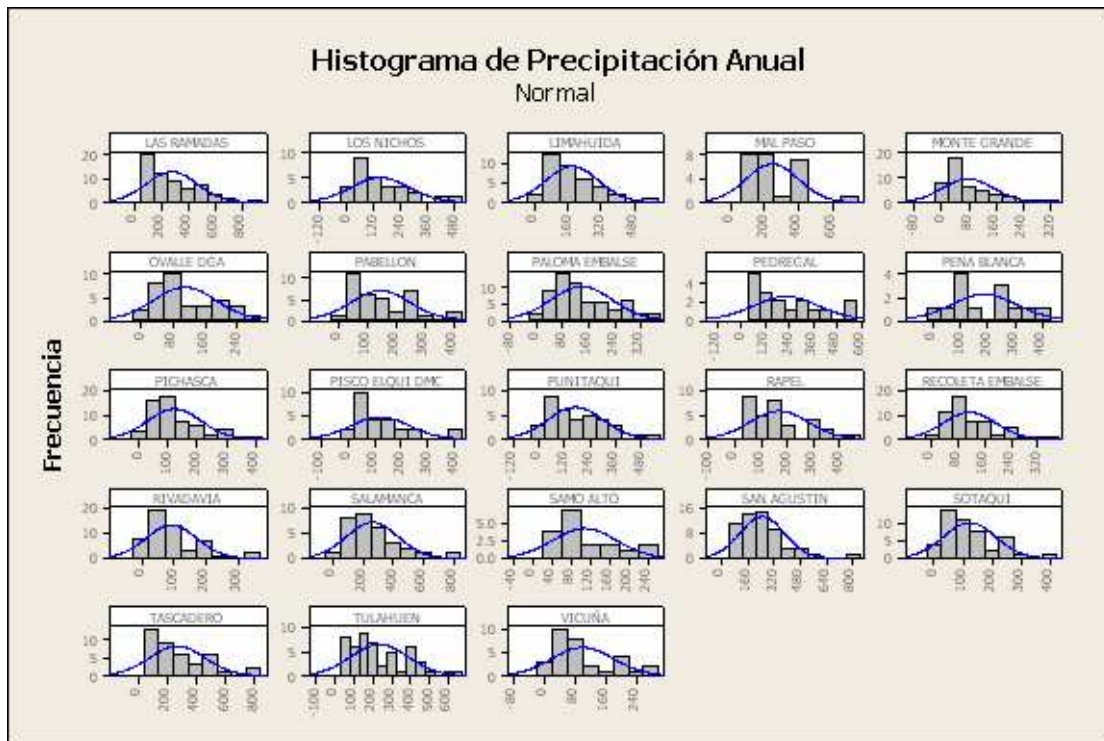


Figura 6b. Histograma de frecuencias de la precipitación anual según estación: Las Ramadas a Vicuña.

## 2

### Etapa 3.2 del Manual Breve

Objetivo de la etapa:

El objetivo de esta etapa es analizar la homogeneidad o consistencia de las series de datos de precipitación. Es decir, detectar los posibles errores en la adquisición de los datos que se van acumulando a través de los años, debido a que, para hacer un análisis adecuado, se requiere que las condiciones de medición sean constantes a través del tiempo. En caso contrario, se debieran corregir los valores para tomar en cuenta los posibles cambios que hayan afectado al instrumental de medición. Para ello se recurre al análisis de la Curva de Doble Masa o Curva Doble Acumulada, previo un agrupamiento inicial de estaciones vecinas.

La etapa 3.2, según se describe en el Manual Breve, se basa en el denominado Análisis de Consistencia y fue desarrollada siguiendo las pautas de la bibliografía recomendada. La metodología, presentación de resultados y su análisis se discuten a continuación:

#### 2.1. Metodología

##### 2.1.1. Material:

###### 2.1.1.1. De los datos:

Se dispuso de registros de datos de precipitación media anual, en unidades de milímetros, obtenidos como resultado de la etapa 3.1 precedente.

###### 2.1.1.2. De las herramientas de análisis:

Para el procesamiento de los datos se utilizaron el software Minitab v.15 en español y la planilla Excel <sup>TM</sup>.

##### 2.1.2. Métodos:

###### 2.1.2.1. Agrupamiento de estaciones:

El análisis de consistencia requiere la comparación de las series hidrológicas respecto a una estación patrón. Según la bibliografía, un criterio a utilizar en esta etapa es la cercanía de las estaciones a comparar. Por ello, y siguiendo las pautas presentadas en el Nacional Drought Atlas (IWR, 2008), en primer lugar se efectuó un análisis multivariante denominado Análisis de Conglomerados o Análisis Cluster. Este consiste en el agrupamiento no de las variables, sino que de objetos (en este caso estaciones meteorológicas) con base a medidas de similaridad que están en función de variables preconcebidas como aquellas que mejor caracterizan a la población en estudio. Las variables consideradas para realizar el análisis, basadas en IWR (2008) se presentan en el Cuadro 6.

Cuadro 6. Variables incluidas en el Análisis de Conglomerados

Variable	Unidad
Precipitación media anual	mm
Sesgo	Adimensional
Curtosis	Adimensional
Coordenadas UTM N	m
Coordenadas UTM E	m
Altitud	m

De la multiplicidad de opciones para realizar el análisis cluster se consideró la siguiente combinación de opciones de cálculo:

- Método de vinculación de Ward
- Medición de la distancia Euclidiano
- Estandarización de variables
- Número de conglomerados a obtener igual a 8

#### 2.1.2.2. Determinación de la estación patrón:

El criterio para obtener la estación patrón consiste en seleccionar aquellas de más largo registro y que se asuman confiables y generar una serie de valores promedio para estas estaciones. Sin embargo, en este estudio, dentro de cada conglomerado obtenido en la etapa previa, existía una amplia variabilidad en cuanto a la longitud de la serie en cada estación así como en la cantidad de datos ausentes en cada serie. Sin embargo, en todos los conglomerados fue posible obtener al menos una estación con una serie hidrológica de larga duración (mayor a 40 años) y con un mínimo de datos ausentes. Por ello, se optó por seleccionar una estación como patrón en cada conglomerado.

#### 1.1.2.2. Análisis de Consistencia:

El análisis de consistencia se realizó tal como se describe en el Manual Breve, siguiendo la metodología de la Curva de Doble Masa Acumulada. La metodología, requiere que, al menos, se rellene los datos ausentes en la estación a contrastar con la estación patrón. Para ello, se efectuaron análisis de regresión lineal entre cada estación y la patrón a fin de rellenar datos ausentes por medio de interpolación lineal. Los datos calculados para rellenar fueron mínimos, entre 3 a 4 datos como máximo en algunas estaciones a comparar con la estación patrón. Este relleno solo se realizó para facilitar la construcción de la Curva Doble Masa. En ningún otro análisis se efectuó relleno de datos y las series a ser sometidas al Análisis Regional de Frecuencias solo incluyen datos originales.

En caso que existieran datos ausentes en la estación patrón y no en la estación de contraste o dudosa, ese dato (año) se eliminó y se continuó calculando la CDA con los datos siguientes.

## 2.2. Resultados

### 2.2.1. Agrupamiento de estaciones:

En la Figura 7 se presentan los resultados del análisis por conglomerados o cluster a través del Dendrograma. El número de la observación en la Figura 7 corresponde al número de orden de la estación según el Cuadro 1. En el Cuadro 7, por su parte, se presentan las agrupaciones de las estaciones según su conglomerado.

De acuerdo con el análisis, los conglomerados se presentan en un gradiente noreste suroeste. Como es obvio, a partir de las variables utilizadas en el análisis, las estaciones se agrupan mayoritariamente en función de estaciones vecinas. Las regiones se agrupan mayoritariamente en el eje longitudinal, aunque existen separaciones transversales, separando los conglomerados de áreas costeras-zona interior respecto a los conglomerados del área de precordillera. La ausencia de regiones claramente definidas hacia la costa es producto de la falta de información meteorológica en esta área. La estación de La Serena presenta una similitud con las del interior del Valle, lo suficientemente grande como para ser incluida dentro del mismo conglomerado.

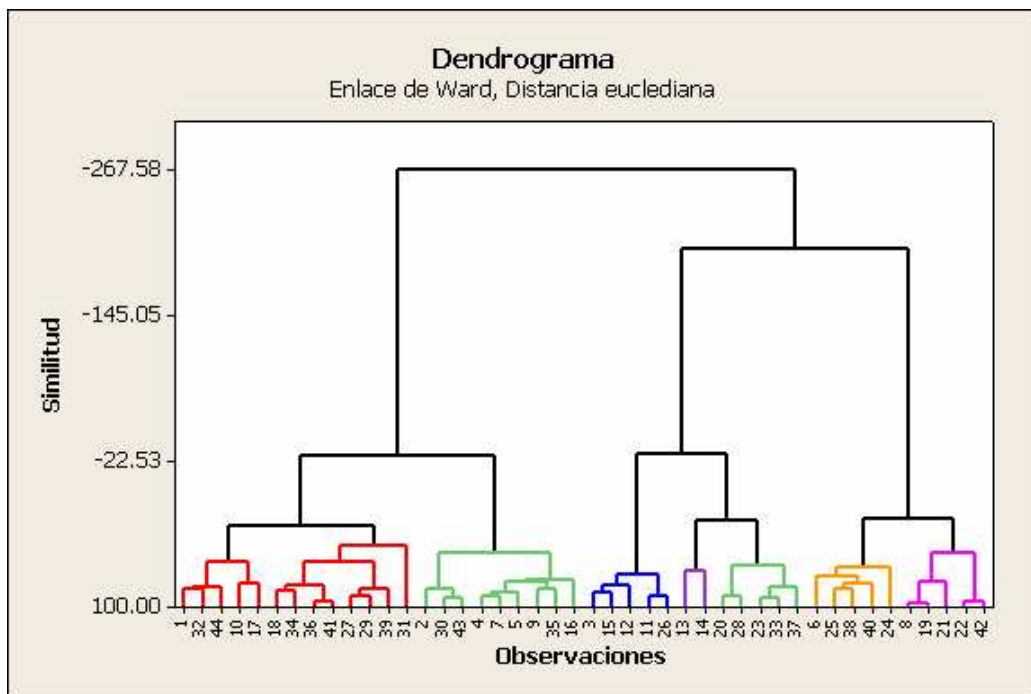


Figura 7. Dendrograma de la conglomeración de las estaciones meteorológicas. IV Región

Al comparar los conglomerados obtenidos en esta etapa, con información de regiones climáticas y tendencias de precipitación en la región (Bodini y Arraya, 2005; Morales *et al*, 2006, López-Cortés y López, 2004, Novoa y López, 2001; HABITERRA, 2005) proveniente de diversos estudios, existe una cierta correspondencia en la distribución geográfica de los cluster y las regiones climáticas. La mayor diferencia está en que el análisis de conglomerados no logró detectar agrupamientos en la zona costera, debido seguramente, a la falta de información meteorológica. Sin embargo, el criterio de agrupamiento es resuelto con base a indicadores de discordancia o mediante gráficas de visualización de L-Momentos, de modo que los conglomerados definidos por el análisis multivariante debe ser considerado como preliminar y sujeto a refinamientos posteriores.

Por otro lado, durante la etapa 3.3. siguiente, se realizó un refinamiento de los conglomerados basado en el análisis de las Curvas Dobles Acumuladas, lo que determinó una leve modificación de conglomerados. En el Cuadro 8 se presenta el agrupamiento post refinamiento. Esta mínima modificación se debió a que en primera instancia, el análisis de conglomerados agrupó la Est. Pichasca, en el valle de Limarí, con aquellas pertenecientes al valle de Elqui. Una situación similar ocurrió con la Est. Hurtado, lo que, además, se tradujo en falta de consistencia de esta estación al graficar las Curvas Dobles Acumuladas. Por ello se optó por agruparlas con sus vecinas dentro del mismo valle.

En la Figura 8 se presenta el mapa de la IV Región con los conglomerados definitivos, los que serán posteriormente sometidos al análisis de homogeneidad. Se aprecia, sin embargo, el traslape de algunas estaciones pertenecientes a distintos conglomerados. Por ejemplo, claramente el conglomerado 8 traspasa las barreras geográficas al incluir estaciones como Est. Rivadavia, en pleno Valle de Equi y Est. Las Breas y Est. Pabellón, pertenecientes al Valle de Limarí. Probablemente se requerirá de un refinamiento posterior a la hora de definir las regiones o subregiones homogéneas definitivas, para lo cual será necesario aplicar los test de Discordancia y Homogeneidad posteriores con el uso de los L-momentos.

Cuadro 7. Agrupamiento preliminar de las estaciones en conglomerados según el análisis cluster.

Conglomerado	Estación
1	ALMENDRAL
1	LA SERENA
1	VICUÑA
1	EL TRAPICHE
1	PICHASCA
2	LA TORRE
2	OVALLE DGA
2	PALOMA EMBALSE
2	PENA BLANCA
2	PUNITAQUI
2	RECOLETA EMBALSE
2	SAMO ALTO
2	SOTAQUI
3	CAREN
3	COGOTI 18
3	COGOTI EMBALSE
3	COMBARBALA
3	EL TOME
3	LA PLACILLA
3	PEDREGAL
3	RAPEL
3	TULAHUEN
4	COIRON
4	MAL PASO
4	SALAMANCA
4	SAN AGUSTIN
4	LIMAHUIDA
5	LAS RAMADAS
5	TASCADERO
5	CUNCUMEN
5	LA TRANQUILLA
5	LAS BURRAS
6	COCHIGUAZ
6	HUANTA
6	HURTADO
6	LA ORTIGA
6	MONTE GRANDE
7	JUNTAS
7	LA LAGUNA EMBALSE
8	LAS BREAS
8	LOS NICHOS
8	PABELLON
8	PISCO ELQUI DMC
8	RIVADAVIA

Cuadro 8. Agrupamiento post refinamiento de las estaciones en conglomerados según el análisis cluster. Se destacan las estaciones que cambian de conglomerado

1	ALMENDRAL
1	LA SERENA
1	VICUÑA
1	EL TRAPICHE
2	PICHASCA
2	LA TORRE
2	OVALLE DGA
2	PALOMA EMBALSE
2	PENA BLANCA
2	PUNITAQUI
2	RECOLETA EMBALSE
2	SAMO ALTO
2	SOTAQUI
2	HURTADO
3	CAREN
3	COGOTI 18
3	COGOTI EMBALSE
3	COMBARBALA
3	EL TOME
3	LA PLACILLA
3	PEDREGAL
3	RAPEL
3	TULAHUEN
4	COIRON
4	MAL PASO
4	SALAMANCA
4	SAN AGUSTIN
4	LIMAHUIDA
5	LAS RAMADAS
5	TASCADERO
5	CUNCUMEN
5	LA TRANQUILLA
5	LAS BURRAS
6	COCHIGUAZ
6	HUANTA
6	LA ORTIGA
6	MONTE GRANDE
7	JUNTAS
7	LA LAGUNA EMBALSE
8	LAS BREAS
8	LOS NICHOS
8	PABELLON
8	PISCO ELQUI DMC
8	RIVADAVIA

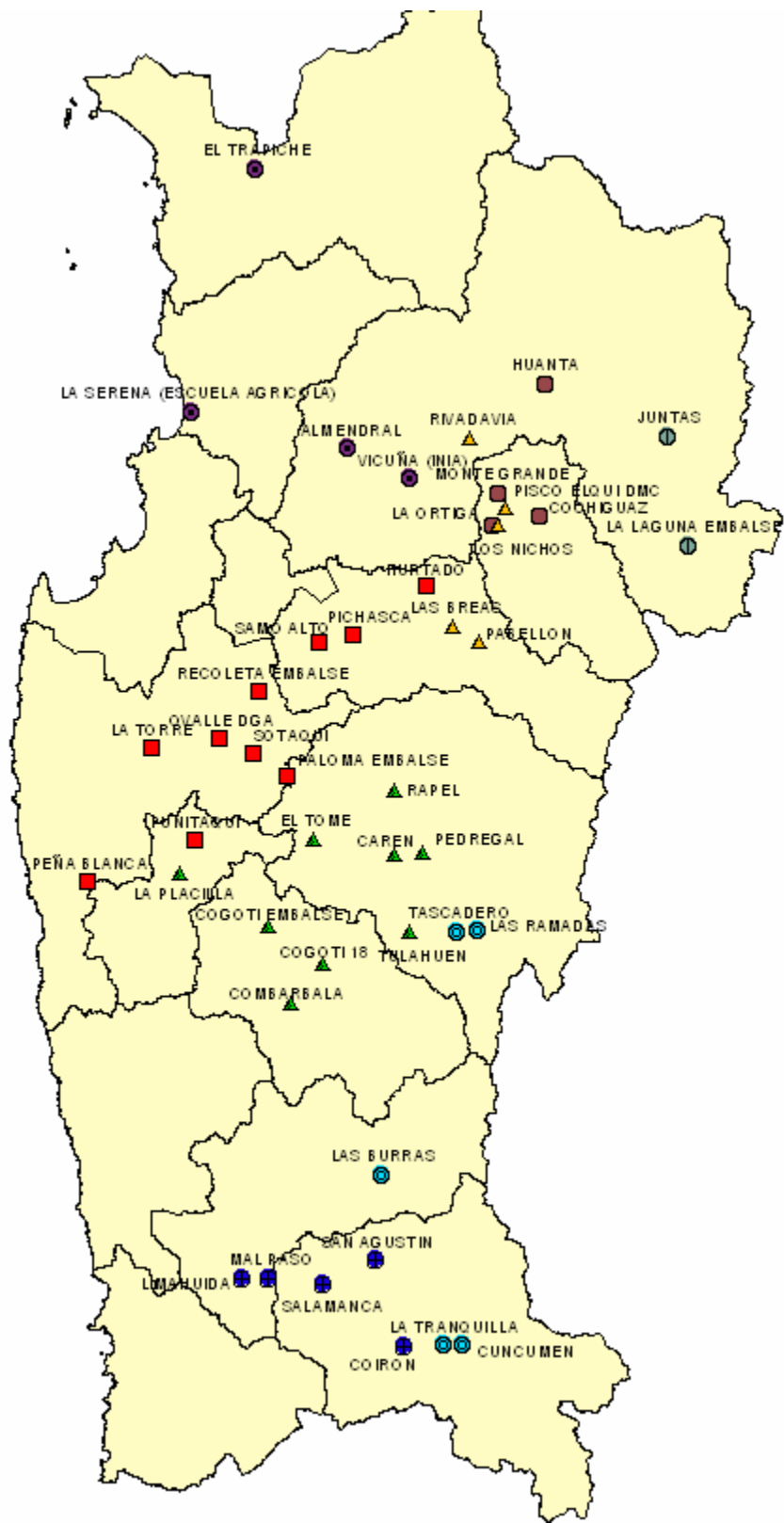


Figura 8. Conglomerados de estaciones meteorológicas. IV Región

### 2.2.1. Análisis de consistencia:

En el Anexo 3 se presentan las Curvas Dobles Acumuladas para cada estación respecto al patrón, en cada una de los 8 conglomerados propuestos. A modo de ejemplo, en la Figura 9 se presentan las CDA para la Est. Recoleta respecto a la Est. Paloma, tomada como patrón. Se aprecia un buen ajuste a una relación lineal, con un  $R^2$  de 0.99, lo que indica una alta consistencia de la precipitación anual respecto a la Est. Patrón. En 7 de las 8 regiones, se obtuvo inicialmente un buen nivel de consistencia interna, donde las CDA se ajustan notoriamente a una relación lineal, con coeficientes de determinación sobre 0.98.

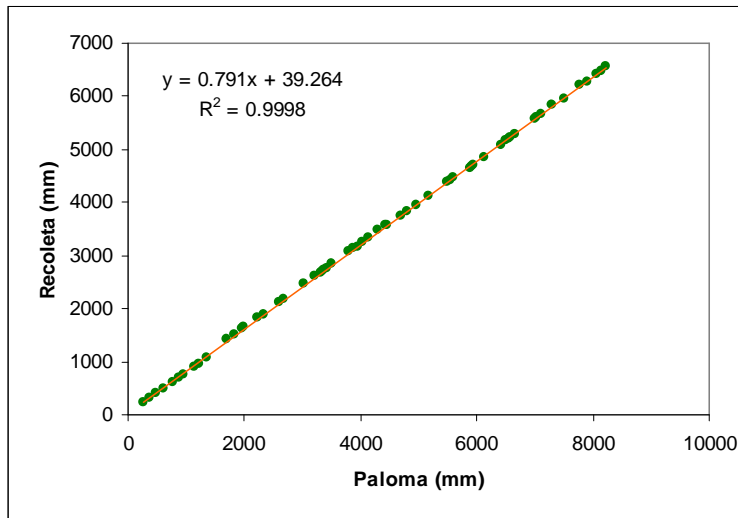


Figura 9. Curva Doble Acumulada de Est. Recoleta versus Est. Paloma

Un caso particular se detectó en el análisis del conglomerado 6. En él, tomando como estación patrón a la Est. Hurtado, por ser la que presentaba un registro de longitud mucho mayor al resto de las estaciones en el conglomerado, se obtuvieron relaciones que no se ajustaban visualmente a una relación lineal, en ninguna de las comparaciones de las estaciones respecto a la patrón. Al definir como patrón a la siguiente estación de mayor longitud de registros, se mejoró sustancialmente la regresión lineal de las CDA. La justificación a esto fue que, considerando la ubicación de la Est. Hurtado, respecto al resto de las estaciones del conglomerado, apoyado en coberturas SIG e imágenes satelitales, si bien desde el punto de vista de la ubicación y monto de la precipitación anual tiene similitud con el resto, geográficamente se ubica en un valle transversal distinto, siendo más cercana entonces, a la Est. Pichasca, con la cual presentó un buen ajuste lineal con la CDA. Sin embargo, la Est. Pichasca, si bien presentó una consistencia adecuada con el resto de las estaciones en su conglomerado (Conglomerado 1), también quedó ubicada en una región que geográficamente no presentan similitud. Por ello, se optó por agrupar las estaciones de Hurtado y Pichasca en el conglomerado 2 y recalculó las CDA en los conglomerados correspondientes (1, 2 y 6).

## 3

### Etapa 3.3 del Manual Breve

Objetivo de esta etapa:

El objetivo de esta etapa es confirmar dos de los supuestos básicos en el análisis estadístico de los datos: El primero, que las series de tiempo de las precipitaciones no presenten una tendencia a disminuir o a aumentar a través de los años, para el periodo de registro de los datos. Este supuesto se puede verificar, entre otros métodos, con el test de Mann-Kendall y con la estadística de pendiente de Sen. El segundo, que los datos, a través del tiempo, son independientes entre sí, es decir, no están autocorrelacionados. Este supuesto se puede verificar, entre otros, con el test de Durbin-Watson. El cumplimiento de estos supuestos es requisito para abordar los análisis estadísticos posteriores.

La etapa 3.3, según se describe en el Manual Breve, se basa en el denominado Análisis de Estacionariedad e Independencia Serial y fue desarrollada siguiendo las pautas de la bibliografía recomendada. La metodología, presentación de resultados y su análisis se discuten a continuación:

#### 3.1. Metodología

##### 3.1.1. Material:

###### 3.1.1.1. De los datos:

Se dispuso de registros de datos de precipitación media anual obtenidos en la etapa 3.1.

###### 3.1.1.2. De las herramientas de análisis:

Para el procesamiento de los datos se utilizaron el software Minitab v.15 en español, la planilla Excel <sup>TM</sup>, la planilla Excel Makesens para el análisis de tendencia.

##### 3.1.2. Métodos:

###### 3.1.2.1. Análisis de tendencia:

La serie de datos de precipitación media anual por estación fueron sometidos a un análisis de tendencia mediante el test de Mann-Kendall y la estadística de pendiente de Sen. Estos test son uno de los más recomendados para analizar tendencia en una serie de tiempo de datos hidrológicos. La planilla Excel usada en su cálculo está disponible en internet según se indica en el Manual Breve de Preparación de Datos. Se contrastó la hipótesis nula de tendencia igual a 0 respecto a la hipótesis alternativa de tendencia distinta de 0. La planilla de cálculo Makesens es de fácil uso y adaptada a valores faltantes y de cualquier longitud del registro.

### 3.1.2.2. Análisis de independencia serial:

El análisis de independencia serial o análisis de autocorrelación se verificó mediante el test de Durbin-Watson haciendo uso del paquete de cómputo Minitab v.15. Se utilizó en este caso, la versión de evaluación disponible en internet, que permite efectuar cálculos por un periodo de 30 días desde su instalación. Para cada estación se determinó la estadística D y se contrastó con los valores críticos superior e inferior de la distribución de Durbin-Watson. Se testó la autocorrelación positiva, para un nivel de significancia del 0.05.

## 3.2. Resultados

### 3.2.1. Análisis de tendencia o estacionariedad:

En el Cuadro 9a se presentan los resultados del análisis de tendencia de las precipitaciones medias anuales de las estaciones Almendral a Mal Paso con el test de Mann-Kendall y la estadística de pendiente de Sen (Sen's Slope). En el Cuadro 9b se presenta el análisis para las estaciones Monte Grande A Vicuña. La interpretación del cuadro es posible hacerla con base al tutorial que se adjunta a la planilla Makesens. En el caso del test de Mann-Kendall, indica con un \* si se rechaza la hipótesis nula de tendencia igual a cero con un 5% de significancia. Por otro lado, la estadística de pendiente de Sen establece los límites de confianza de la pendiente Q. De este modo, si los límites van entre un valor negativo y uno positivo, indica que la pendiente de la serie de datos no difiere significativamente de 0, es decir, constante y por lo tanto, consistente con la ausencia de una tendencia ya sea positiva o negativa en la serie de datos.

Cuadro 9a. Test de Man-Kendall y estadística de pendiente de Sen por estación: Almendral a Mal Paso

Time series	First year	Last Year	n	Test Z	Signific.	Q	Qmin99	Qmax99	Qmin95	Qmax95
ALMENDRAL	1959	2006	41	0.11	N.S.	0.080	-1.914	2.079	-1.455	1.609
CAREN	1944	2006	51	-0.39	N.S.	-0.364	-2.678	2.296	-2.013	1.496
COCHIGUAZ	1990	2006	17	0.41	N.S.	1.863	-9.307	11.110	-7.530	7.382
COGOTI 18	1943	2006	55	-0.68	N.S.	-0.427	-2.301	1.691	-1.900	1.196
COGOTI EMBALSE	1943	2006	59	-0.47	N.S.	-0.368	-2.238	1.956	-1.837	1.335
COIRON	1974	2006	32	-0.63	N.S.	-2.427	-11.623	9.333	-9.078	4.938
COMBARBALA	1977	2006	28	-0.77	N.S.	-2.115	-13.247	4.946	-9.550	3.275
CUNCUMEN	1959	2006	39	0.50	N.S.	0.810	-3.954	5.839	-2.828	4.512
EL TOME	1966	2006	36	1.46	N.S.	1.666	-2.063	7.040	-0.723	5.105
EL TRAPICHE	1979	2006	26	0.68	N.S.	0.750	-2.394	3.998	-1.536	2.998
HUANTA	1990	2006	17	-0.33	N.S.	-0.279	-7.419	4.407	-5.585	3.492
HURTADO	1943	2006	58	-0.46	N.S.	-0.187	-1.417	1.175	-1.078	0.678
JUNTAS	1990	2006	15	0.00	N.S.	-0.257	-15.885	19.230	-10.012	11.639
LA LAGUNA EMBALSE	1964	2006	41	0.28	N.S.	0.457	-2.991	3.581	-2.040	2.816
LA ORTIGA	1979	2006	26	-0.02	N.S.	-0.090	-11.629	6.397	-9.228	4.423
LA PLACILLA	1990	2006	16	0.86	N.S.	7.295	-24.415	38.779	-15.632	26.056
LA SERENA	1972	2003	29	0.99	N.S.	1.236	-1.991	6.148	-1.318	4.746
LA TORRE	1943	2006	54	-0.27	N.S.	-0.195	-1.608	1.402	-1.181	1.002
LA TRANQUILLA	1966	2006	38	0.45	N.S.	0.992	-4.962	7.197	-2.870	5.339
LAS BREAS	1944	1977	19	-1.54	N.S.	-2.684	-7.288	2.881	-5.330	0.791
LAS BURRAS	1990	2006	15	0.00	N.S.	-0.407	-24.166	23.214	-16.871	15.882
LAS RAMADAS	1944	2006	60	1.47	N.S.	1.582	-1.177	4.874	-0.459	3.957
LOS NICHOS	1977	2006	27	-0.63	N.S.	-1.541	-10.897	4.113	-7.694	2.967
LIMAHUIDA	1964	2006	36	0.75	N.S.	0.674	-3.688	5.325	-2.205	4.233
MAL PASO	1961	1988	25	-0.12	N.S.	-0.22	-8.73	10.10	-5.56	7.26

Cuadro 9b. Estadística de Man-Kendall y test de pendiente de Sen por estación: Monte Grande a Vicuña

Time series	First year	Last Year	n	Test S	Test Z	Signific.	Q	Qmin99	Qmax99	Qmin95	Qmax95
MONTE GRANDE	1959	2006	44		0.08	N.S.	0.053	-1.352	1.679	-1.000	1.198
OVALLE DGA	1972	2006	34		-0.44	N.S.	-0.429	-3.948	2.787	-3.051	2.004
PABELLON	1969	2006	35		2.03	*	2.350	-0.905	7.802	0.195	5.874
PALOMA EMBALSE	1944	2006	59		-0.60	N.S.	-0.350	-1.974	1.438	-1.529	0.921
PEDREGAL	1968	1988	17		1.36	N.S.	11.083	-7.544	28.630	-2.467	24.965
PENA BLANCA	1991	2006	13		-0.18	N.S.	-3.840	-37.373	26.260	-32.906	10.829
PICHASCA	1947	2006	58		-0.11	N.S.	-0.098	-1.436	1.465	-1.120	1.069
PISCO ELQUI DMC	1977	2006	26		-1.26	N.S.	-1.578	-8.286	3.254	-6.790	1.862
PUNITAQUI	1962	2006	36		-0.23	N.S.	-0.260	-4.198	4.085	-3.454	2.805
RAPEL	1969	2006	33		0.79	N.S.	1.128	-2.861	7.256	-1.884	4.875
RECOLETA EMBALSE	1944	2006	55		-0.85	N.S.	-0.376	-1.675	1.076	-1.396	0.585
RIVADAVIA	1954	2006	53		0.08	N.S.	0.043	-1.193	1.471	-0.888	1.123
SALAMANCA	1974	2006	31		-0.20	N.S.	-0.504	-8.663	6.964	-6.368	5.402
SAMO ALTO	1970	1988	18		1.18	N.S.	2.680	-4.900	12.554	-2.871	10.100
SAN AGUSTIN	1945	2006	57		-0.32	N.S.	-0.339	-2.788	2.446	-2.151	1.697
SOTAQUI	1955	2006	47		0.08	N.S.	0.041	-1.999	1.966	-1.396	1.425
TASCADERO	1961	2006	40		0.44	N.S.	0.737	-3.900	6.801	-2.834	5.268
TULAHUEN	1950	2006	50		-0.12	N.S.	-0.150	-3.296	3.924	-2.582	2.588
VICUÑA	1971	2003	31		0.24	N.S.	0.453	-2.492	5.820	-1.731	4.464

Se aprecia que en todas las estaciones, no fue posible rechazar la hipótesis nula de ausencia de tendencia, con un nivel de significancia del 5%. Por ello, es posible concluir que, bajo ese nivel de significancia, la precipitación media anual es de tipo estacionaria, cumpliendo uno de los supuestos básicos para el análisis regional de frecuencias.

### 3.2.1. Análisis de la independencia serial:

En el Cuadro 10 se presentan los indicadores de autocorrelación de Durbin-Watson para un nivel de significancia de 0.05. De acuerdo con el Cuadro 10, no es posible rechazar la hipótesis nula, de modo que se concluye que no existe autocorrelación positiva en la serie de datos en cada una de las estaciones. De esta manera, se cumple el supuesto de autocorrelación requerido en la etapa de preparación de datos, de que los valores de precipitación son independientes entre sí a través del tiempo, con miras a los análisis de frecuencia con el método de L-momentos posteriores.

Cuadro 10. Estadístico de Durbin-Watson por estación para test de autocorrelación positiva.

Estación	N	alfa	d	dL	dU	Significancia
ALMENDRAL	41	0.05	2.32	1.44	1.54	N.S.
CAREN	51	0.05	2.19	1.5	1.59	N.S.
COCHIGUAZ	17	0.05	2.42	1.12	1.37	N.S.
COGOTI 18	55	0.05	2.61	1.53	1.61	N.S.
COGOTI EMBALSE	59	0.05	2.46	1.55	1.62	N.S.
COIRON	32	0.05	2.44	1.35	1.49	N.S.
COMBARBALA	28	0.05	2.54	1.34	1.48	N.S.
CUNCUMEN	39	0.05	2.66	1.44	1.54	N.S.
EL TOME	36	0.05	2.42	1.39	1.53	N.S.
EL TRAPICHE	26	0.05	2.37	1.29	1.45	N.S.
HUANTA	17	0.05	2.38	1.12	1.38	N.S.
HURTADO	58	0.05	1.84	1.54	1.62	N.S.
JUNTAS	15	0.05	2.45	1.08	1.36	N.S.
LA LAGUNA EMBALSE	41	0.05	2.58	1.44	1.54	N.S.
LA ORTIGA	26	0.05	2.31	1.29	1.45	N.S.
LA PLACILLA	16	0.05	2.24	1.08	1.36	N.S.
LA SERENA	30	0.05	2.1	1.35	1.49	N.S.
LA TORRE	54	0.05	2.38	1.53	1.61	N.S.
LA TRANQUILLA	38	0.05	2.34	1.43	1.54	N.S.
LAS BREAS	19	0.05	2.22	1.19	1.41	N.S.
LAS BURRAS	15	0.05	2.58	1.08	1.36	N.S.
LAS RAMADAS	15	0.05	2.5	1.08	1.36	N.S.
LOS NICHOS	27	0.05	2.48	1.33	1.48	N.S.
LIMAHUIDA	36	0.05	2.32	1.39	1.53	N.S.
MAL PASO	25	0.05	2.46	1.29	1.45	N.S.
MONTE GRANDE	44	0.05	2.15	1.47	1.56	N.S.
OVALLE DGA	34	0.05	2.36	1.4	1.52	N.S.
PABELLON	35	0.05	2.43	1.39	1.53	N.S.
PALOMA EMBALSE	59	0.05	2.19	1.55	1.62	N.S.
PEDREGAL	17	0.05	2.52	1.12	1.37	N.S.
PENA BLANCA	13	0.05	2.51	1	1.3	N.S.
PICHASCA	58	0.05	2.35	1.54	1.62	N.S.
PISCO ELQUI DMC	26	0.05	2.41	1.29	1.45	N.S.
PUNITAQUI	36	0.05	2.69	1.39	1.53	N.S.
RAPEL	33	0.05	2.28	1.38	1.51	N.S.
RECOLETA EMBALSE	55	0.05	2.12	1.53	1.61	N.S.
RIVADAVIA	53	0.05	2.26	1.53	1.61	N.S.
SALAMANCA	31	0.05	2.32	1.35	1.49	N.S.
SAMO ALTO	18	0.05	2.45	1.12	1.38	N.S.
SAN AGUSTIN	57	0.05	2.43	1.54	1.62	N.S.
SOTAQUI	47	0.05	2.39	1.48	1.57	N.S.
TASCADERO	40	0.05	2.43	1.44	1.54	N.S.
TULAHUEN	50	0.05	2.38	1.5	1.59	N.S.
VICUÑA	31	0.05	2.18	1.35	1.49	N.S.

## 4

### Conclusiones

En análisis de los datos para su preparación y verificación de supuestos, con miras a su utilización en las etapas siguientes del Análisis Regional de Frecuencias, contribuyó a determinar los siguientes aspectos:

1. Las series hidrológicas cumplen con los supuestos básicos para su utilización en el análisis regional de frecuencias, como son: presentar adecuada consistencia u homogeneidad, ser estacionarias, es decir, no presentar tendencia positiva o negativa de la serie de datos a través del tiempo y no presentar autocorrelación o que los datos sean independientes entre sí a través de los años.
2. El análisis cluster permitió definir 8 conglomerados, los que, luego de un refinamiento, son considerados como base para el análisis de regionalización y su testeo posterior mediante medidas gráficas y analíticas de homogeneidad. En adelante, y hasta que los análisis de discordancia no sugieran otra cosa, estos conglomerados pasarán a denominarse Regiones Propuestas.
3. El análisis exploratorio permitió esclarecer patrones de variabilidad de la precipitación en cuanto a su magnitud y dispersión estacional, así como ajustar criterios de detección de valores atípicos.

## Bibliografía

- CAZALAC, 2006. Aplicación de metodologías para determinar la eficiencia de uso del agua. Estudio de caso de la región de Coquimbo. CAZALAC-RODHOS. Informe de Estudios y Proyectos.
- Bodini, H. y Arraya, F. 2005. La Región de Coquimbo Espacios y Recursos para un Desarrollo Sustentable: Visión Geográfica Global. Universidad de La Serena, Centro de Estudios Regionales. 41 pp
- HABITERRA, 2005. Plan Intercomunal Costero de la Region de Coquimbo. Capítulo 1.4 Medio Físico Natural. Disponible en internet: <http://www.planregional.cl/ivregion/>
- IWR, 2008. National Drought Atlas. Institute for Eater resources. US Army Corps of Engineers. Disponible en internet: [http://www.iwr.usace.army.mil/iwr/atlas/Atlasintro.htm]
- López-Cortés, F. y López, D. 2004. Capítulo 2. Antecedentes bioclimáticos del Parque Nacional Bosque Frey Jorge. En: Squeo, F.; Gutierrez, J. y Hernandez, I. 2004. Historia Natural del Parque Nacional Bosque Frey Jorge. Ediciones Universidad de La Serena.
- Morales L.; Canessa, F.; Mattar, C.; Orrego, R. y Matus, F. Caracterización y zonificación edáfica y climática de la Región de Coquimbo, Chile. J. Soil Sc. Plant. Nutr. 6(3) 2006(52-74). Disponible en internet: [www.scielo.cl/pdf/rcsuelo/v6n3/art05.pdf](http://www.scielo.cl/pdf/rcsuelo/v6n3/art05.pdf)
- Novoa, J. y López, D. 2001. Capítulo 2. IV Región: El escenario geográfico físico. En: Squeo, F.; Arancio, G. y Gutierrez, J. (Eds.), 2001. Libro Rojo de la Flora Nativa y de los Sitios Prioritarios para su Conservación: Región de Coquimbo. Ediciones Universidad de La Serena
- Salmi, T; Määttä, A.; Anttila, P.; Ruoho-Airola, T. and Amnell, t. 2002. Detecting trends of annual values of atmospheric pollutants by the Mann-Kendall test and Sen's slope estimates -the excel template application Makesens. Publications of air quality n° 31. Disponible en internet: [http://www.fmi.fi/kuvat/MAKESENS\\_MANUAL.pdf](http://www.fmi.fi/kuvat/MAKESENS_MANUAL.pdf)