

ESTIMACIÓN DE LA LLUVIA MENSUAL EN SITIOS DE LA REGIÓN PAMPEANA ARGENTINA MEDIANTE MODELOS ESTADÍSTICOS: DESEMPEÑO Y APLICACIONES.

Juan Carlos Labraga, Laura Brandizi.
Centro Nacional Patagónico, CONICET.
laurabrandizi@gmail.com

INTRODUCCION

Los extensos períodos de sequía o exceso de lluvia que afectan recurrentemente la producción agrícola y ganadera en las provincias fitogeográficas de la Pampa y el Espinal en la República Argentina resaltan la importancia de investigar sus causas y desarrollar métodos de pronóstico climático eficientes que permitan mejorar la administración de su sistema agronómico. El objetivo de este trabajo es desarrollar un método para estimar la lluvia mensual en localidades distribuidas en estas regiones a partir del conocimiento del estado de la atmósfera sobre cada sitio e incorporando empíricamente los efectos de las características fisiográficas locales que condicionan la distribución espacial y temporal de la lluvia. El propósito de este método, conocido como *downscaling*, es obtener información más detallada y de utilidad práctica sobre las variaciones climáticas de la lluvia que actualmente son pronosticadas a escala regional mediante modelos globales de la circulación general de la atmósfera.

MATERIALES Y METODOS

Los modelos estadísticos utilizados para estimar la cantidad de lluvia en localidades de la Pampa y el Espinal (ver ubicación las localidades en Fig. 1) fueron desarrollados siguiendo la técnica de Regresión Lineal Múltiple (Wilks, 2006). El predictando es la anomalía del total de lluvia mensual y los predictores potenciales son seleccionados entre 23 variables atmosféricas relacionadas con los mecanismos productores de lluvia en estas regiones. Para discernir cuales de estas variables contribuyen significativamente a la variación de la respuesta se utilizó el procedimiento de regresión gradual por pasos hacia delante (forward stepwise), que incorpora sucesiva y selectivamente nuevas variables independientes hasta un máximo que, en este trabajo, comúnmente no supera las ocho variables. El criterio básico de selección es que una variable no debe ser incluida en el modelo si esta no contribuye en una cantidad significativa a la varianza explicada, para lo cual se aplica un test F estándar (Storch y Zwiers, 2001).

Las siguientes variables formaron parte del conjunto de predictores potenciales: flujo de calor latente (v3) y flujo de calor sensible (v4) en superficie, presión en superficie (v5), temperatura máxima (v6), mínima (v7) y media (v8) a 2m, cobertura total de nubes (v9), agua precipitable (v11), temperatura en 850 hPa (v13) y en 700 hPa (v14), altura geopotencial de 500 hPa (v15), espesor 500 -1000 hPa (v16), componentes zonal (v17) y meridional (v18) del viento en 850 hPa, presión reducida al nivel medio del mar (v19), vorticidad en 700 hPa (v20), 500 hPa (v21) y 200 hPa (v22), divergencia en 700 hPa (v23) y 200 hPa (v24), componente zonal del viento en 250 hPa (v25), advección de la relación de mezcla de vapor de agua en 850 hPa (v26) y convergencia del flujo de relación de mezcla de vapor de agua en 850 hPa (v27).

Las variables enumeradas se obtuvieron del conjunto de datos NCEP-DOE Atmospheric Model Intercomparison Project Reanalysis 2 (R2) (Kanamitsu et al., 2002). Los datos del nivel de superficie están disponibles en una grilla global Gaussiana de 1.9 ° longitud x 1.9 ° latitud aproximadamente. Los datos de altura están disponibles en una grilla uniforme de 2.5 ° longitud x 2.5 ° latitud y en 17 niveles estándar de presión. Los datos fueron interpolados

bilinealmente sobre las coordenadas de los sitios de medición de la lluvia usando los cuatro puntos de la grilla más próximos. Los valores medios mensuales así obtenidos permiten caracterizar el estado de la atmósfera sobre la vertical de cada estación pluviométrica.

Los registros de lluvia en la mayoría de las 30 estaciones seleccionadas se iniciaron a principios del siglo pasado o anteriormente. El conjunto de datos NCEP-DOE R2 comienza en el año 1979 y se extiende hasta la actualidad. El período de estudio, con datos simultáneos, se extiende desde 1979 hasta 2006 (28 años).

La regresión lineal múltiple fue calculada entre anomalías mensuales del predictando y de los predictores, entendiéndose por anomalía la desviación del valor mensual observado respecto del valor promedio de cada mes sobre el período 1979-2006. Se desarrollaron modelos de regresión para 28 localidades, 16 en la Pampa y 12 en el Espinal. La ecuación de regresión fue ajustada para cada estación del año, lo cual permitió modelar adecuadamente diferentes factores de lluvia y un tratamiento apropiado de datos dudosos (outliers).

La expresión general de los modelos estadísticos de regresión lineal múltiple es

$$P' = P - P_{MO} = B_0 + B_1X_1 + B_2X_2 + \dots + B_NX_N + \varepsilon$$

En donde P es la precipitación mensual, P_{MO} es la precipitación media mensual observada, B_0 es la constante de regresión, B_i son los coeficientes de regresión de las respectivas variables independientes X_i , ε es una variable aleatoria independiente con distribución normal, y N es el número de variables independientes o “predictores” considerados.

Los modelos fueron desarrollados o entrenados en el período 1987-2006 y verificados en el período parcialmente independiente 1979-1998, de igual longitud que el anterior y un solapamiento de 12 años. En unos pocos sitios con más de 20 y menos de 28 años de observaciones de lluvia disponibles, los modelos desarrollados, siempre en base a 20 años de observaciones, no fueron verificados.

Una de las medidas usadas más frecuentemente para ponderar la eficiencia con la cual la ecuación de regresión ajustada especifica valores de la variable dependiente P' dados los correspondientes valores de las variables independientes X_i es el coeficiente de determinación R^2 , interpretado como la proporción de la variación del predictando que es descrita por la ecuación de regresión (Wilks, 2006):

$$R^2 = 1 - SSE / SST$$

Siendo $SSE = \sum \varepsilon^2$ la suma de los cuadrados de los residuos de la regresión, y $SST = \sum (P' - \bar{P})^2$.

Para obtener una indicación de la bondad de las estimaciones de lluvia que cabe esperar en aplicaciones futuras del modelo, R^2 fue recalculado en el período de verificación.

Se define S^2 como la fracción de la variación de la lluvia mensual observada que da cuenta la suma de las anomalías de lluvia estimadas y el ciclo anual, calculado como el cuadrado del coeficiente de correlación de Pearson. S^2 permite ponderar la eficiencia de los pronósticos climático y por regresión combinados.

RESULTADOS

Desempeño de los modelos

Las Fig. 1 y 2 muestran valores de las dos medidas de eficiencia de los modelos, definidas anteriormente, obtenidos en la región de estudio en los períodos de desarrollo y de verificación.

El porcentaje de la varianza de las anomalías mensuales de lluvia que dan cuenta los modelos de regresión lineal múltiple, R^2 , en el período de desarrollo (Fig. 1a) crece uniformemente de

oeste a este con valores inferiores a 15% en el oeste del Espinal y más de 45% en el este de la Pampa.

En el Espinal, el desempeño de los modelos es pobre en el sector oeste menos lluvioso, con porcentajes de la varianza explicada inferiores a 20%, mientras que en el sector norte relativamente lluvioso los modelos exhiben una capacidad igual o mayor que en la Pampa, con porcentajes de la varianza explicada mayores que 40%.

La Fig. 1b muestra valores de S^2 , los cuales permiten ponderar la eficiencia ganada al incluir en el modelo la información sobre el ciclo anual.

El porcentaje de la varianza de las anomalías de lluvia que dan cuenta los modelos de regresión (R^2) es, en promedio, algo mayor en la Región Pampeana que en el Espinal. En cambio, la proporción de la varianza total (S^2) es levemente mayor en el Espinal. Lo primero puede deberse a la mayor variabilidad interanual de la lluvia en el Espinal que en la Pampa, y lo segundo a la mayor amplitud del ciclo anual en el Espinal que en la Pampa.

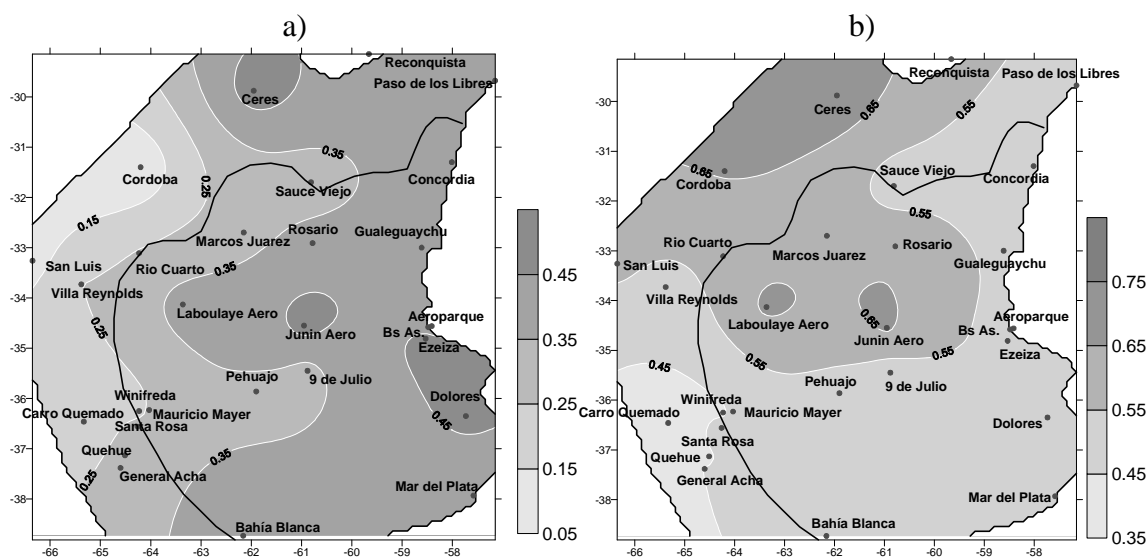


Figura 1 a) Porcentaje de la varianza de las anomalías mensuales de lluvia explicada por Regresión Lineal Múltiple, R^2 , en el periodo de desarrollo de los modelos. b) Porcentaje de la varianza total de la lluvia mensual modelada, S^2 , en el período de desarrollo de los modelos.

La línea negra continua delimita la Pampa (interior) del Espinal (exterior).

Cuando el valor de R^2 es mucho menor que el de S^2 , como en el borde oeste del Espinal, la varianza asociada al ciclo anual es mayor que la varianza de las anomalías de lluvia pronosticadas. Cuando la diferencia entre R^2 y S^2 es pequeña, como en el borde este de la Pampa, la varianza de las anomalías de lluvia pronosticadas es mucho mayor que la varianza del ciclo medio anual. Esto último ocurre en las estaciones Buenos Aires, Aeroparque, Ezeiza, Dolores, Mar del Plata, Concordia y Paso de los Libres. Estos sitios están ubicados al este de la llamada zona de transición entre el régimen subtropical continental de lluvias de verano y el régimen marítimo de lluvias invernales (Prohaska, 1976), cuyo régimen se caracteriza por la distribución uniforme de la lluvia durante el año. Las estaciones con mejor desempeño en la varianza total son Ceres, Laboulaye, Junín, Córdoba y Rosario.

Los valores del coeficiente de determinación obtenidos en el período de verificación (Fig. 2a) siguen un patrón muy similar al del período de desarrollo, pero disminuyen en promedio 15%. Esto sugiere que se debería utilizar un período mayor que 20 años si se desea desarrollar modelos más estables. Algunas características señaladas en el período de desarrollo son

válidas también en el período de verificación: los modelos de downscaling desarrollados son más eficientes en el centro y este de la región de estudio.

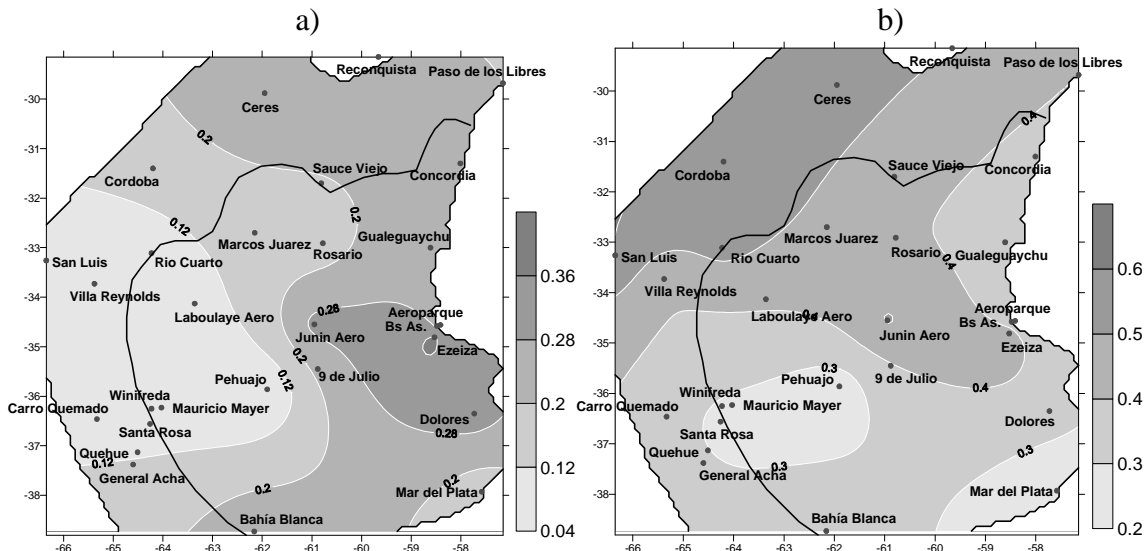


Figura 2 a) Porcentaje de la varianza de las anomalías mensuales de lluvia explicada por Regresión Lineal Múltiple, R^2 , en el periodo de verificación de los modelos. b) Porcentaje de la varianza total de la lluvia mensual modelada, S^2 , en el período de verificación de los modelos.

Para ponderar la bondad del procedimiento de downscaling a escala regional, se promediaron las series de anomalías de lluvia de los 28 sitios modelados y se calculó el coeficiente de correlación entre valores estimados y observados, obteniéndose un valor de 0.88.

Las variables independientes de los modelos de downscaling

La Fig. 3 muestra cuales son las variables predictoras que tuvieron participación más frecuentes en los modelos de regresión. El agua precipitable (v11) es la variable que tiene mayor participación, estando presente en 43% a 63% de los sitios modelados según la época del año. La variable vorticidad en 700 hPa (v20) integra los modelos de un 40% de los sitios, especialmente en el centro y sur de la región y en otoño e invierno. La variable componente meridional del viento en 850 hPa (v18) se destaca en verano en 27% de los sitios con predominio de la Pampa sobre el Espinal. La variable convergencia del flujo de vapor de agua en 850 hPa (v27) mantiene una distribución bastante pareja durante el año integrando solo un 23% de los modelos en el norte de la región de estudio. El flujo de calor latente en superficie (v3) participa en un 27% de los modelos en invierno y en menor proporción en las demás estaciones. El flujo de calor sensible en superficie (v4) se encuentra en un 24% de los modelos en primavera, verano y otoño decayendo en invierno. Otras variables tienen una participación menor en las ecuaciones de modelo.

El signo del coeficiente de regresión de la variable agua precipitable es siempre positivo. El agua precipitable es la profundidad de la lámina de agua líquida que resultaría de condensar todo el vapor de agua contenido en una columna de aire. Es un buen indicador de la cantidad de humedad potencialmente disponible para generar lluvia.

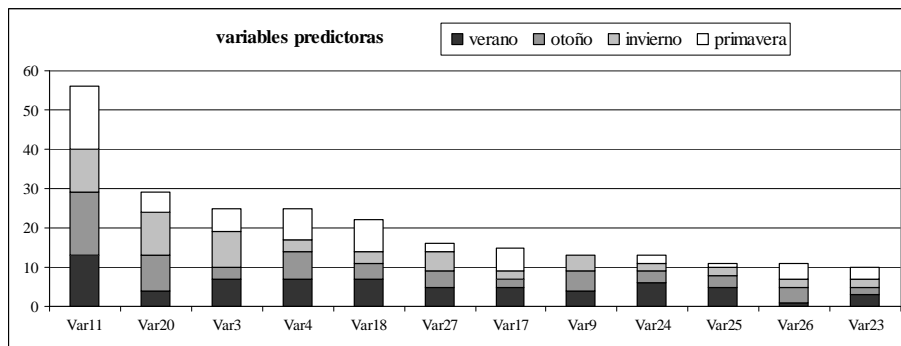


Figura 3. Variables predictoras que tuvieron mayor participación en los modelos de regresión.

La vorticidad en 850 hPa aparece más frecuentemente como variable predictora en otoño e invierno, época en la cual la presencia perturbaciones sinópticas transitorias es más frecuente en la región. Por su participación en el total de la varianza explicada puede ocupar el primero al cuarto lugar. El correspondiente coeficiente de regresión tiene signo negativo, lo que indica una relación consistente entre anomalías de vorticidad ciclónica y excesos de lluvia.

El signo del coeficiente de regresión de la variable flujo de calor sensible es siempre negativo. El transporte vertical de calor sensible cerca de la superficie depende de la diferencia de temperatura entre la superficie y las capas bajas de la atmósfera. También depende de la intensidad de la turbulencia, modulada por la estabilidad del aire y la velocidad del viento. Por convención, se considera que el flujo desde la superficie hacia la atmósfera es positivo.

El signo del coeficiente de regresión de la variable flujo de calor latente es predominantemente positivo. El transporte vertical de calor latente cerca de la superficie depende de la diferencia de humedad entre la superficie y las capas bajas de la atmósfera. También depende de la intensidad de la turbulencia, modulada por la estabilidad del aire y la velocidad del viento. Por convención, se considera que el flujo desde la superficie hacia la atmósfera es positivo. El resultado obtenido indica que cuando la superficie está en condiciones de suministrar abundante vapor de agua la lluvia supere los valores normales.

El signo del coeficiente de regresión de la variable componente meridional del viento en 850 hPa en verano, cuando es una variable dominante, es negativo. Esto indica que el predominio de los vientos del norte, que favorecen la advección de humedad hacia la región, esta asociado con anomalías positivas de lluvia.

La convergencia del flujo de vapor de agua en la región aparece asociado con excesos de lluvia, como lo indica el signo positivo del correspondiente coeficiente de regresión. Los signos de los coeficientes de regresión de las dos últimas variables son compatibles con la relación observada entre anomalías de circulación de gran escala y de lluvia en la región pampeana (Labraga et al., 2002).

Detección de Periodos Secos y Periodos Húmedos

Luego de aplicar un promedio móvil de tres meses a la serie del promedio regional de las anomalías de lluvia observadas desde 1979 hasta 2006, se detectaron periodos secos (anomalías negativas) y periodos húmedos (anomalías positivas), los cuales fueron comparados con los valores obtenidos por downscaling.

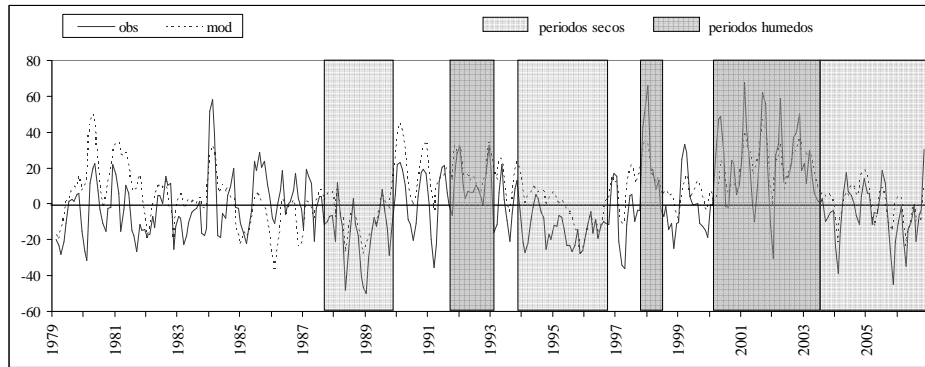


Figura 4. Promedio móvil de tres meses de las anomalías regionales de lluvia observadas (línea llena) y modeladas (línea de puntos). Periodos secos (gris claro) y húmedos (gris oscuro).

Se analizaron sólo aquellos períodos anómalos que presentaron mayor longitud y continuidad en el tiempo para ponderar la capacidad del modelo para reproducir el inicio y fin de estas rachas. Dentro de los extremos secos se distinguieron tres períodos; Septiembre de 1987 a Noviembre de 1989, Diciembre de 1993 a Octubre de 1996 y Septiembre de 2003 a Octubre de 2006. Los períodos húmedos son de Noviembre de 1991 a Enero de 1993, Noviembre de 1997 a Mayo de 1998 y Marzo de 2000 a Agosto de 2003. Puede verse en la Fig. 4 que el procedimiento de downscaling reproduce adecuadamente la duración de estos períodos anómalos.

También fueron comparados los patrones espaciales de estas anomalías. En la Fig.5 se puede observar la distribución espacial de las anomalías observadas y modeladas en el primer período seco. El modelo detecta los extremos secos en el noreste y este de la región de estudio, así como un sector con anomalías positivas pequeñas en el sudoeste. El mayor déficit se desarrolla en el este de la región pampeana. El coeficiente de correlación entre ambos campos es 0.60.

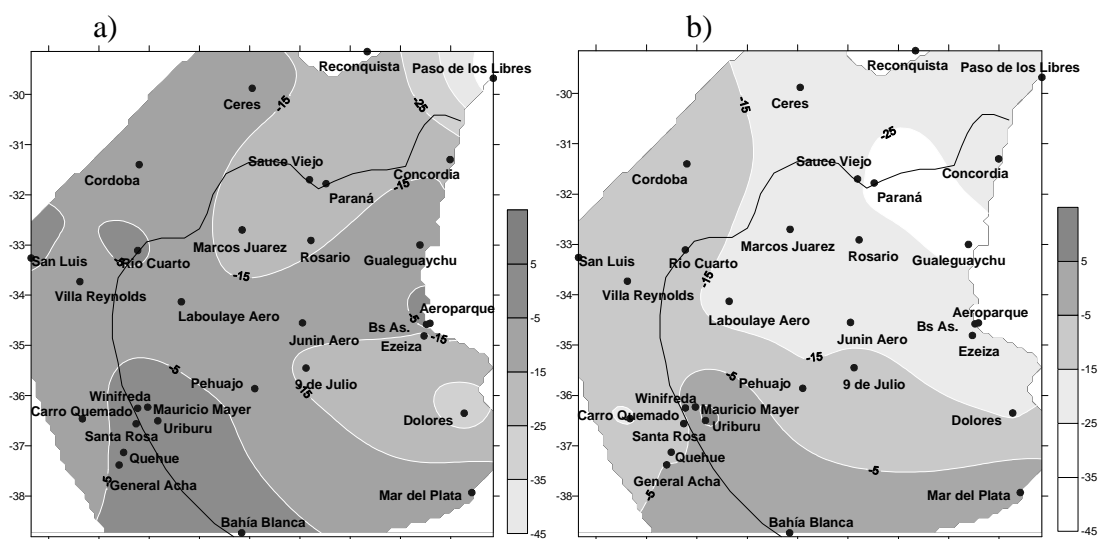


Figura 5. Campo de las anomalías de lluvia observadas (a) y modeladas (b), promediadas en

el período seco de Septiembre de 1987 a Noviembre de 1989.

CONCLUSIONES Y APLICACIONES FUTURAS

El modelo detecta satisfactoriamente las zonas en donde se producen los extremos secos y húmedos; en particular existe interés en los aspectos causantes de las sequías ya que en general producen mayor daño en las actividades agropecuarias que los extremos húmedos.

Los resultados obtenidos sugieren que esta técnica de downscaling podría ser aplicada provechosamente para mejorar los pronósticos locales de anomalías mensuales a trimestrales de lluvia en la región de la Pampa usando como variables predictoras las salidas de simulaciones con un modelo global de circulación general de la atmósfera.

BIBLIOGRAFIA

Kanamitsu M, Ebisuzaki W, Woollen J, Yang SK, Hnilo JJ, Fiorino M, Potter GL (2002) NCEP-DEO AMIP-II Reanalysis (R-2). *Bull. Amer. Met. Soc.*, 83: 1631-1643.

Labraga, JC, B. Sian y O. Frumento (2002) Anomalies in the atmospheric circulation associated with the rainfall excess or deficit in the Pampa Region in Argentina. *Jou. Geophys Res.- Atmosphere*, Vol 107, D23, I-15.

Prohaska, F. J., 1976 The climate of Argentina, Paraguay and Uruguay, in World Survey of Climatology, Climates of Central and South America, edited by W. Schwerdtfeger, pp. 13 – 112, Elsevier Sci., New York.

Von Storch H, F. W. Zwiers (2001). *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge, 484pp.

Wilks, D.S (2006). *Statistical Methods in the Atmospheric Sciences*. Academic Press, London, 627pp.