

**PREDICCIÓN DE INDICADORES DE CAMBIO CLIMÁTICO PARA
ARGENTINA DURANTE EL SIGLO XXI**

TESIS DE LICENCIATURA



Diciembre de 2008

Tesistas

Luciano Leggieri y Juan Aizpuru

Directores

Dr. Jean-Philippe Boulanger y Dr. Enrique Carlos Segura

Jurados

Dra. Ana Ruedin y Dra. Mónica Marino

Contenidos

Resumen	4
Introducción	5
1 Control de Calidad	7
1.1 Introducción	7
1.2 APACH: Parte 1: Control de Calidad y aplicación a las estaciones del Servicio Meteorológico Argentino	7
1.2.1 Resumen del artículo	7
1.2.2 Introducción	8
1.2.3 Datos	10
1.2.4 Exámenes y árbol de decisión para los registros diarios de temperatura	11
1.2.5 Pruebas y árboles de decisión para precipitación	15
1.2.6 Aplicación al Servicio Meteorológico Nacional Argentino	18
1.2.7 Conclusiones y perspectivas	21
1.3 Implementación de la solución	23
1.3.1 Requerimientos de la implementación	23
1.3.2 Arquitectura	23
1.4 Conclusiones	25
1.5 Bibliografía	26
2 Estimación de variograma	27
2.1 Resumen	27
2.2 Introducción al problema	27
2.2.1 Variograma	28
2.3 Conjunto de datos	28
2.4 Modelos estadísticos de variograma	30
2.5 Modelo de variograma basado en MLP	31
2.5.1 Arquitectura	31
2.5.2 Variantes	32

2.5.3	Optimización y evaluación de redes.....	33
2.6	Comparación entre modelos / Resultados	36
2.6.1	Análisis de las diferencias entre modelos MLP	38
2.7	Conclusiones. Hacia la interpolación... ..	40
2.8	Bibliografía	42
3	Proyección del los modelos de predicción del clima respecto a datos verídicos	43
3.1	Introducción	43
3.1.1	Grupo Intergubernamental de Expertos sobre el Cambio Climático.....	43
3.1.2	Reporte Especial sobre Escenarios de Emisiones	43
3.1.3	Familias de Escenarios	44
3.2	Modelos.....	44
3.2.1	Datos utilizados (temperatura y precipitación).....	44
3.3	Análisis de los datos de los modelos	45
3.3.1	Determinación temporal	45
3.3.2	Pre-procesamiento de los resultados	48
3.4	Obtención de clases a partir de los modelos.....	49
3.4.1	Aplicación de redes neuronales.....	49
3.4.2	Clasificación de los resultados obtenidos.....	49
3.5	Proyección de los modelos respecto a los valores observados.....	54
3.6	Conclusiones	61
3.7	Trabajo futuro.....	62
3.7.1	Clasificación a nivel global	62
3.7.2	Proyección de los modelos en otras regiones	63
3.8	Bibliografía	63
	Conclusiones generales	64
	Apéndice	65

Resumen

Se desarrollaron tres herramientas que forman parte de un sistema integral de proyección de indicadores de cambio climático. El objetivo final de este sistema es la proyección de indicadores de clima para el siglo XXI, utilizando modelos pre-existentes de cambio climático y registros del siglo XX de estaciones meteorológicas.

La primera herramienta es APACH, un procedimiento automatizado de control de calidad para datos históricos de temperatura y precipitación. Este se basa en dos árboles de decisión (uno para temperatura, el otro para precipitación), que procesan las series diarias de cada estación, y elaboran conclusiones sobre la calidad de cada dato, tanto en referencia a la serie en la que ocurre (tests intra-estación) como el mismo contexto espacial (tests inter-estación). Cada árbol como resultado genera una descripción para cada valor de temperatura o precipitación, donde se registra cualquier situación anómala encontrada (o la ausencia de estas), así como un indicador numérico (de 0 a 3) del nivel de confianza que se tiene sobre el valor.

APACH se utilizó para controlar y depurar el registro histórico del Servicio Meteorológico Nacional Argentino para el período 1960-2005, que a continuación fue utilizado para generar una grilla regular de datos diarios de temperatura, con precisión de 0.5°, sobre el territorio de la república Argentina.

Este proceso de interpolación está compuesto por dos etapas, la primera consiste en un estimador de variograma basado en redes neuronales supervisadas de tipo Perceptrón multicapa (MLP). La segunda etapa utiliza los valores estimados de variograma e interpola la base de datos de temperatura utilizando Kriging.

El modelo MLP de variograma fue justamente la segunda herramienta desarrollada como parte de este trabajo, mientras que la interpolación por Kriging no se trata en este informe, pero fue desarrollada dentro del mismo grupo de trabajo y se presenta en Pausader (2008).

El variograma es una función que describe el grado de dependencia espacial de un proceso estocástico $Z(x)$, y se define como el incremento esperado (al cuadrado) de los valores entre un punto X y un punto Y. Kriging es un interpolador muy conocido en geo-estadística, y se puede aplicar a procesos estocásticos geográficos utilizando un modelo de variograma.

La segunda herramienta, combinada con el trabajo de interpolación, genera información de temperatura diaria, presentada en una grilla regular con resolución de 0.5°.

Este nuevo conjunto de datos fue justamente utilizado por la tercera herramienta para evaluar la aplicabilidad de distintos modelos de cambio climático al territorio de la República Argentina. Como parte de este trabajo se estudiaron las mediciones de temperatura de modelos climáticos del IPCC (Intergovernmental Panel on Climate Change / Panel intergubernamental del Cambio Climático) utilizando Mapas Auto-Organizados (SOM, por sus siglas en inglés) cuando éstos predicen un aumento de 2 grados respecto al período de referencia del siglo 20. De este análisis se hizo foco en Sudamérica para obtener regiones (clases) en las cuales analizar la veracidad en la predicción de los hechos climatológicos del siglo 20 realizada por cada modelo sobre Argentina.

Introducción

En este trabajo se presentará una nueva metodología para estimar proyecciones de cambio climático durante el siglo XXI, así como su implementación en un software que realiza los procedimientos asociados de forma automática.

Si bien la metodología presentada es aplicable a cualquier zona geográfica para la que se cuente con la información necesaria, durante este trabajo se ejemplifica con la aplicación realizada sobre la República Argentina.

Esta metodología comprendió tres grandes procedimientos:

1. Control de calidad: Desarrollo y aplicación de un sistema automático de control de calidad, para depurar y garantizar la confiabilidad del registro meteorológico a utilizar.
2. Interpolación de series de temperatura: Creación de una metodología para interpolar las series de temperaturas diarias a una grilla regular de precisión arbitraria.
3. Evaluación de los modelos preexistentes, para el territorio Argentino utilizando las series interpoladas.

Control de Calidad

Como parte de este trabajo se desarrolló un sistema automatizado de control de calidad de series diarias de temperatura y precipitación provenientes de estaciones meteorológicas. Estos controles utilizan información estadística de la propia estación, así como de la vecindad, para juzgar los valores de cada registro. Por estar basado en técnicas estadísticas no paramétricas, estos controles aplican a cualquier zona geográfica y período histórico, sin necesidad de calibrar parámetros.

La precipitación y las temperaturas mínimas y máximas de cada registro se evalúan utilizando sendos árboles de decisión. Durante este proceso se tiene en cuenta información interna del registro (comparación de las distintas variables entre sí), información estadística de la estación a la que pertenece e información estadística y diaria de las estaciones vecinas a la evaluada.

Al final del control se genera información detallada de la confianza asociada a cada registro, para sus valores de temperatura y precipitación. Esta información podría luego ser procesada automáticamente, o estudiada manualmente para un control más fino.

El sistema de control fue aplicado durante su desarrollo al registro histórico del Servicio Meteorológico Nacional. Éste comprende un conjunto de 250 estaciones de tiempo a lo largo de 40 años de historia, para un total de más de 3 millones de registros diarios y casi 10 millones de valores de tiempo.

Interpolación de series de temperatura

La información obtenida del control de calidad fue utilizada como filtro, y resultó en una base calificada de series diarias de temperatura y precipitación.

Como ya se mencionó, este proceso fue realizado para evaluar modelos climáticos, que usualmente presentan la información en una grilla geográfica regular. Por lo tanto, un paso intermedio imprescindible fue generar series equivalentes para los puntos de la grilla a evaluar, interpolando la información de las estaciones.

Para esto, se utilizó el concepto de variograma, fuertemente ligado a la correlación: Se desarrolló e implementó un algoritmo de predicción de variograma entre dos puntos, basado en redes neuronales de tipo MLP. Éste utiliza, durante el entrenamiento, indicadores obtenidos de la base de datos, y genera un modelo de variograma aplicable a cualquier par de puntos de la región en cuestión.

El variograma predicho se utilizaría luego para interpolar series de temperatura. Si bien este último trabajo no fue realizado por los autores del presente estudio, se hizo en colaboración, y está disponible en un paper que será referenciado más adelante.

Evaluación de los modelos preexistentes utilizando las series interpoladas

Debido a la creciente preocupación por el aumento de temperatura que ocurre en el mundo, las comunidades científicas han generado diversos modelos de predicción del clima basados en formulas matemáticas que emulan los factores humanos y naturales que ocurren en el planeta.

Para cada modelo se toman sus datos estimados de temperatura de todo el mundo promediados en el período en el que el cada uno indica un aumento de dos grados a nivel mundial con respecto al período de referencia. Con estos datos, tomando aquellos que cubren Sudamérica, se los procesa utilizando un tipo de red neuronal artificial conocida como Mapas auto-organizados (SOM) y se organizan los resultados del proceso en una cantidad especificada de clases, a través de un análisis de cluster jerárquico.

Estas clases representan a distintos sectores de la geografía sudamericana, y este trabajo se especializa en aquellas que pisan territorio argentino. Para cada región obtenida de la Argentina, se comparan los datos de cada modelo, tomando en cuenta las diferencias entre lo que predicen para el siglo 21 y lo que modelaron para el siglo 20, y las diferencias entre lo que modelaron para el siglo 20 contra los datos reales observados del mismo período. Debido a la modalidad de trabajo utilizada, fue necesario contar con los valores del clima de Argentina organizados en una grilla regular.

De la comparación de los datos se buscan patrones de comportamiento y se observan aquellos modelos que se destacan en los distintos sectores del país. Por último se muestra como el análisis de modelos utilizando redes neuronales y su posterior clasificación pueden ser útiles para un trabajo futuro de análisis global o de otras regiones del planeta.

1 Control de Calidad

1.1 Introducción

En esta sección se presentará el sistema automatizado de control de calidad de series meteorológicas diarias “APACH” (por sus siglas en inglés).

En la sección 1.2 se incorpora el trabajo “*A Procedure for Automated Quality Control and Homogenization of historical daily temperature and precipitation data (APACH)*” [1], co-escrito por los autores del presente.

En la sección 1.3 se presentará la implementación de la metodología descrita. En la sección 1.4 se presenta la conclusión del trabajo en Control de Calidad, con hincapié en el futuro del mismo y comentando el uso de esta herramienta en el marco del presente proyecto. En la sección 1.5 se presentará la bibliografía y en el apéndice se describen con mayor detalle algunos conceptos introducidos a lo largo del trabajo.

1.2 APACH: Parte 1: Control de Calidad y aplicación a las estaciones del Servicio Meteorológico Argentino

1.2.1 Resumen del artículo

El presente artículo describe el componente de control de calidad de un procedimiento automático (APACH: A Procedure for Automated Quality Control and Homogenization of Weather Station Data) desarrollado para controlar la calidad y homogenizar los datos diarios históricos de temperatura y precipitación obtenidos de estaciones meteorológicas. El método de control de calidad se encuentra basado en un conjunto de algoritmos del estilo de árboles de decisión, que analizan separadamente la precipitación y temperaturas mínimas y máximas. Todos los exámenes son no-paramétricos, y ofrecen por lo tanto un campo espacial de aplicaciones mucho mayor (universal).

Se aplicó el método a los datos diarios históricos comprendidos entre los años 1959 a 2005 de la base de datos del Servicio Meteorológico Nacional. Los resultados obtenidos son coherentes con la historia del Servicio y más específicamente con la historia de la implementación de procesos sistematizados de control de calidad.

Con respecto a la temperatura, el método detectó un número mayor de valores sospechosos antes del año 1967 (época anterior al control de calidad) y luego de 1997 (cuando se dejó un control de calidad de tiempo real). Para la precipitación, la detección de errores de precipitaciones extremas es compleja, pero el método claramente detectó una fuerte disminución en el número de valores atípicos potenciales luego de 1976, cuando el Servicio Meteorológico Nacional fue militarizado y su red fuertemente reducida, enfocándose más en aquellas estaciones meteorológicas ubicadas en aeropuertos. También en precipitación se analizó en detalle las secuencias de sequías largas y se fue capaz de identificar potenciales secuencias largas erróneas. Esto es importante para el uso de los datos para estudios de impactos hidrológicos o agrícolas.

Finalmente, todos los datos fueron marcados con códigos que representan el camino seguido por los algoritmos de árboles de decisión. Mientras que cada código se asocia con una categoría (“Useful”, “Need-Check”, “Doubtful” o “Suspect” correspondientes a “útil”, “a comprobar”,

“dudoso” y “sin confianza” respectivamente), el usuario final es libre de redefinir cada asignación de las categorías.

1.2.2 Introducción

En el contexto del cambio climático, rescatar los datos climatológicos diarios para luego controlar su calidad y homogeneizarlos se ha convertido en una tarea crucial para los equipos de investigación del clima que se encuentran alrededor del mundo (Alexandersson y Moberg, 1997; Moberg y Alexandersson, 1997; Wijngaard et al., 2003; Caussinus y Mestre, 2004; Rusticucci y Barrucand, 2004; Brandsma y Können, 2006; Rusticucci y Renom, 2007). Entender la variabilidad diaria del clima pasado, la relación entre eventos extremos (temporadas templadas o frías, inundaciones, sequías, etc.) y la variabilidad a gran escala es muy importante para evaluar a los futuros escenarios de cambio climático, y más específicamente, sus impactos en aquellos componentes vulnerables de la sociedad como la agricultura, hidrología o la salud.

En el proyecto CLARIS, varios equipos performaron sus propios controles de calidad en los datos diarios históricos de la Argentina (Rusticucci y Barrucand, 2004), Uruguay (Rusticucci y Renom, 2007), Brasil o Chile. En algunos casos, los métodos de homogeneización se aplicaron además para corregir tales datos (Rusticucci y Renom, 2007). Aquí, el primer objetivo es utilizar tales conocimientos preexistentes en la forma de un procedimiento automático de manera tal de controlar la calidad de los datos diarios de temperatura y precipitación de estaciones meteorológicas. Este objetivo es actualmente un paso importante hacia el desarrollo de una base de datos meteorológica extendida, componiendo a cientos de estaciones tal como se encuentra agendado en el proyecto europeo CLARIS LPB (La Plata Basin, Cuenca del Plata) FP7 (2008-2012). Un procedimiento de homogeneización será pronto implementado para proveer datos diarios homogeneizados para tendencias seculares y análisis de eventos extremos. En vista de la importancia de datos diarios históricos para cada país, es crucial trabajar conjuntamente con el Servicio Meteorológico Nacional Argentino en la aplicación y evaluación de tal procedimiento. Esta colaboración permite acceso a la base de datos más completa disponible en el país y provee al Servicio Meteorológico Nacional de una lista con datos potencialmente sospechosos en su base histórica. A largo plazo, esta colaboración llevará a una base de datos nacional homogeneizada, crucial para evaluar los cambios en tendencias seculares y extremos eventos pasados y en la posible evolución en el contexto de escenarios de cambio climático.

Este artículo se organiza de la siguiente manera: la sección 1.2.3 presenta el conjunto de datos diarios históricos usados para probar al procedimiento automático; las secciones 1.2.4 y 1.2.5 describen a los exámenes aplicados a los datos de temperatura y precipitación respectivamente y se discuten los árboles de decisión usados en el procedimiento automático; la sección 1.2.6 presenta los resultados del control de calidad y se analizan en relación a la historia del Servicio Meteorológica Nacional; finalmente, la sección 1.2.7 describe el método y muestra algunas perspectivas más amplias del mismo.

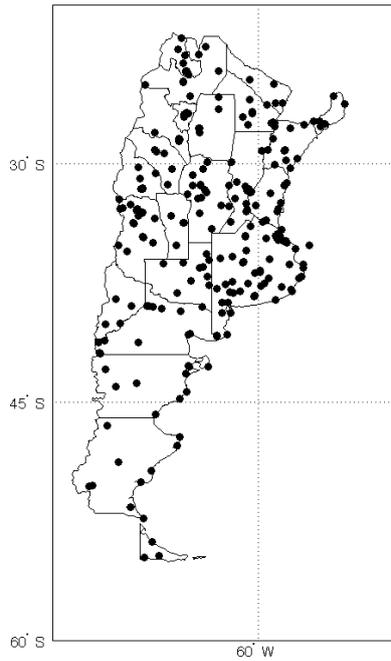


Figura 1 – Ubicación espacial de las estaciones meteorológicas que proveen de datos diarios de precipitación y temperatura para todo o parte del período 1959-2005.

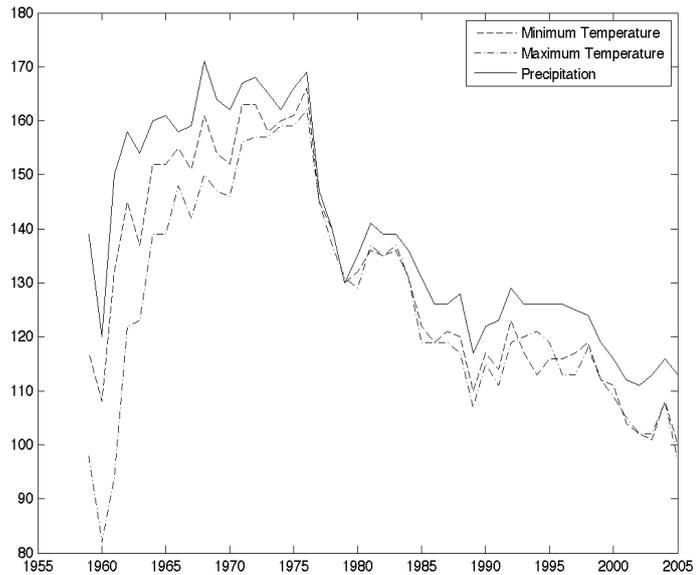


Figura 2 – Número de estaciones meteorológicas con menos del 10% de datos faltantes en la red. Mostrado por variable por año, para el período 1959-2005.

El conjunto de datos diarios cubre al período 1959-2005 para las estaciones ubicadas sobre toda la Argentina. Sin embargo, como se muestra claramente en la Figura 1, la densidad de las estaciones es pobre al sur de los 40°S (Patagonia). La mejor densidad de estaciones se encuentra en la Provincia de Buenos Aires y en las regiones de actividades agrícolas importantes. Varias estaciones se encuentran ubicadas en aeropuertos. Si se hace foco en las estaciones con menos del 10% de datos faltantes cada año (Figura 2), el número incrementa del año 1959 hacia el 1973-1975, pero desde entonces (Golpe Militar en 1976), el número de estaciones con menos del 10% de datos faltantes ha decrecido sostenidamente (en paralelo con la disminución en el número total de estaciones pertenecientes a la red). Es más, durante el primer gobierno democrático (1983-1989), la hiperinflación afectó fuertemente a la economía. Aún más reducciones en la red ocurrieron desde 1999.

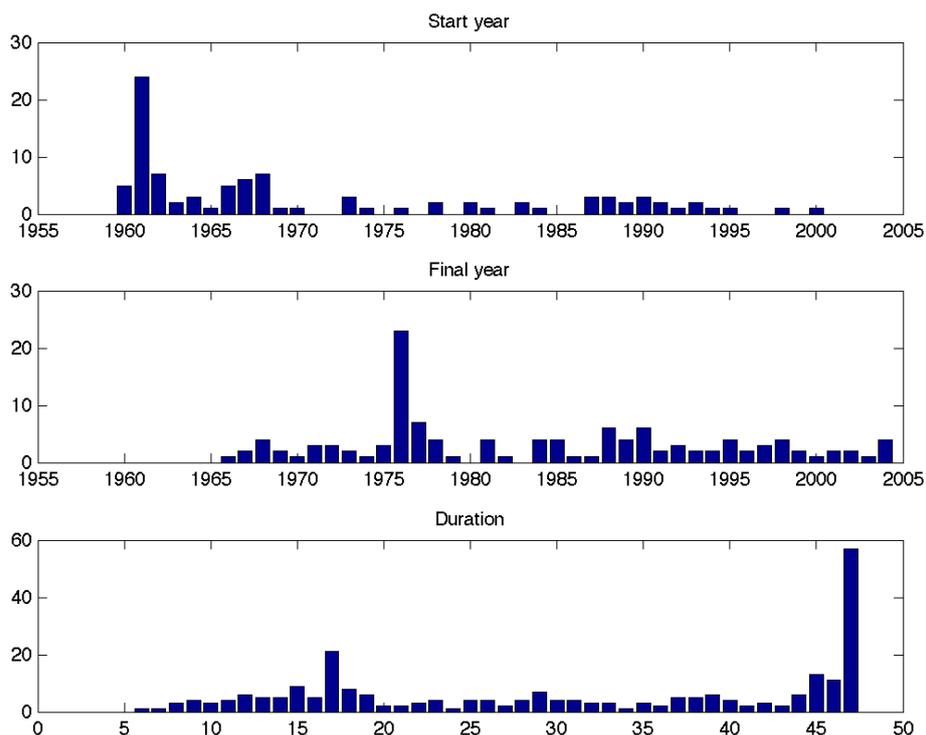


Figura 3 – Número de estaciones por año de iniciación (arriba), año de finalización (medio) y duración (abajo). La gran mayoría de las estaciones con 47 años de registros se encuentran todavía operativas.

1.2.3 Datos

Este resultado es confirmado por los histogramas mostrados en la Figura 3, los cuales muestran que la mayoría de las estaciones fueron creadas alrededor de 1960 o del año 1990. Y la mayoría que pararon de registrar datos, lo hicieron durante los 70s, 90s o 2000. Luego del Golpe Militar de 1976, la mayoría de las instituciones argentinas fueron militarizadas, incluyendo al Servicio Meteorológico Nacional (SMN). Gran cantidad de los observadores de las estaciones del SMN fueron reemplazados por personal militar, el cual carecía de entrenamiento y experiencia adecuados. En general, más del 40% de la red fue perdida desde 1970. Afortunadamente, más de 70 estaciones tienen registros que datan desde hace más de 40 años.

Finalmente, el porcentaje de disponibilidad de los datos en toda la base muestra que mucha de la información faltante se encontraba disponible antes del fin de los 70s. En efecto, en este período, el SMN empezó a sistematizar su colección de datos, y a copiar datos de tarjetas a cintas. Desafortunadamente, muchas de las tarjetas de papel estaban dañadas por el polvo y la humedad, y varios de sus datos no pudieron ser leídos correctamente. Aunque tales datos pudieron ser recuperados de registros en papel, la falta de recursos humanos (especialmente en la época de la digitalización en 1978 y luego de 1985) imposibilitó el llenado de los datos faltantes.

Los datos actuales fueron contralados manualmente en dos pasos: primero, un operador los controlaba en tiempo real. Luego, al final de cada década, fueron comprobados manualmente, mirando a coherencias temporales y espaciales. Cuando se accedió a la información, fue en el proceso de control de calidad, por lo que sólo el período que va desde octubre de 1967 a 1997

ha sido controlado dos veces. Estos datos (período 1960-2000) han sido usados para detección de eventos extremos en Vincent et al., (2005) y en Haylock et al. (2006).

1.2.4 Exámenes y árbol de decisión para los registros diarios de temperatura

Los métodos de control de calidad aplicados a los datos diarios de temperatura utilizan dos grupos fundamentalmente diferentes: (1) métodos de una sola estación donde los datos de una única estación son analizados y (2) métodos espaciales donde los datos de varias estaciones vecinas son analizados. Los métodos espaciales son mucho más confiables en detectar o confirmar errores potenciales porque pueden estimar la probabilidad de detección de un error basándose en un conjunto bien definido de estaciones vecinas. Luego, los exámenes espaciales son necesarios para un marcado más acertado de las observaciones (Vejen et al., 20-2).

El procedimiento de control de calidad se la temperatura diaria mínima y máxima se base en tres clases de comprobaciones:

1. Comprobaciones de consistencia (una estación)
2. Comprobaciones de rangos y escalonadas basadas en la distribución de la estación (una estación)
3. Comprobaciones espaciales basadas en una vecindad de estaciones (espacial)

Al final del procedimiento, cada dato diario tiene marcas que describen los resultados de cada prueba y se da un valor de confianza adecuado:

- Useful (útil): el valor es ciertamente correcto.
- NeedCheck (a comprobar): el valor es probablemente correcto.
- Doubtful (dudoso): el valor es probablemente erróneo, pero en casos inusuales puede ser correcto.
- Suspect (sin confianza): el valor es ciertamente erróneo, pero en casos excepcionales, puede ser correcto.

Comprobaciones de consistencia

Se aplicaron dos pruebas de consistencia. La primera verifica el rango diario (si la temperatura mínima es más alta que la temperatura máxima). La segunda comprueba si la temperatura se mantiene igual por varios días.

- Diferencia diaria entre las temperaturas mínimas y máximas. Mientras el rango de temperatura diaria se usa a menudo para detectar puntos de corte, sólo se usaron los rangos diarios de valores para detectar días en los que la temperatura mínima era más fuerte que la temperatura máxima. Tales condiciones pueden ocurrir en los datos argentinos por dos razones:
 - Algunas observaciones históricas de las temperaturas mínimas y máximas fueron hechas en diferentes momentos del día. La temperatura mínima era el mínimo absoluto entre las 21 horas y las 9 de la mañana, mientras que la temperatura

máxima era el valor máximo absoluto entre las 9 de la mañana y las 21. Como los períodos eran diferentes, era factible que un registro diario de una temperatura mínima fuera mayor al de la máxima. Afortunadamente tales casos desaparecieron a medida que los procedimientos de observación fueron sistematizados.

- El registro estaba mal o se leyeron mal los datos de las tarjetas perforadas. La mayoría de los errores encontrados en el conjunto de datos parecen relacionados con esto último.

Esta comprobación se aplicó primero. Para no marcar ambos valores de temperaturas, el resultado se guardó y se comparó con los resultados de otras pruebas, antes de decidir si sólo uno o ambos valores de temperatura debían ser marcados como erróneos (Suspect).

- Series de temperatura constante: una marca especial se creó cuando los valores de temperatura son iguales por dos días o más. El código usado para esta marca es SL para series largas de 3 o más días consecutivos, o SS para series cortas de 2 días consecutivos.

Pruebas de rango y escalonadas

El objetivo era diseñar un procedimiento tan universalmente aplicable como fuera posible, así que se decidió no utilizar las pruebas tradicionales relacionadas con valores absolutos en umbrales, las cuales suelen ser apropiadas para regiones específicas. Se re-escalaron los datos en percentiles para computar las distancias al percentil 25^{to} (para valores bajos) o al 75^{to} (para valores altos) de la siguiente manera:

$$\text{Per}(X) = (X - P_{25}) / (P_{50} - P_{25}) \text{ if } X \leq P_{25} \text{ (25}^{\text{to}} \text{ percentil)}$$

$$\text{Per}(X) = (X - P_{75}) / (P_{75} - P_{50}) \text{ if } X \geq P_{75} \text{ (75}^{\text{to}} \text{ percentil)}$$

$$\text{Per}(X) = 0 \text{ de otro modo}$$

Otros percentiles podrían ser usados sin modificar la detección actual de valores periféricos en las distribuciones. La elección se encuentra basada en el método más usado en los paquetes de software matemáticos para dibujar esquemas de caja y bigotes: cualquier valor cuya distancia sea mayor que 1.5 la distancia entre la mediana y el tercer (o primer) cuartil, se considera un valor periférico. Los percentiles se computan para cada mes del año (de enero a diciembre) basándose en los datos diarios de cada mes. El método está luego basado en Periféricos y Depresión: Dos pruebas aplicadas directamente a las series temporales reescaladas de temperaturas.

- **Prueba de Periféricos (Outlier Test):** Los métodos de control de calidad utilizan pruebas de periféricos para identificar valores muy grandes (más grandes que 3 o 4 desviaciones estándar de la media). Aquí se utiliza la distancia no dimensional del percentil 25^{to} al 75^{to}: $\text{OUT}(T_i) = \text{Per}(T_i)$ en donde T es la temperatura mínima o la máxima. Esta prueba se usa de manera complementaria, cuando la Prueba de Depresión (ver más abajo) no puede ser computada. Si la prueba de periféricos es mayor que 1.5, un código relacionado con su amplitud se asocia con el dato: OM (para valores entre 1.5 y 3), OK (para valores entre 3 y 5) y OX (para valores mayores que 5). La elección de estos umbrales fue basada en la función matemática de caja y bigotes y en el análisis de los resultados. En cualquier caso,

los resultados finales no son significativamente afectados por pequeños cambios en estos umbrales, ya que la prueba más importante es la Prueba de Vecindad (comprobación espacial, ver más abajo). Si la prueba de periféricos no puede ser computada (por no haber datos suficientes), se asocia el código ON.

- **Prueba de escalonamiento (Step Test):** Esta prueba mide cuán grande es la diferencia entre dos días consecutivos, de la siguiente manera: $STEP(T_t) = \text{Per}(T_t - T_{t-1})$, donde T es la temperatura mínima o máxima. Esta prueba no es utilizada directamente por el sistema, pero conceptualmente es una base para la Prueba de Depresión.
- **Prueba de Depresión (Dip Test):** Una prueba equivalente a la descrita en Vejen et al., 2002, se desarrolló según la siguiente definición:

$$DIP(T_t) = -STEP(T_t) * STEP(T_{t+1}) \text{ si } (T_t - T_{t-1}) * (T_{t+1} - T_t) < 0$$

$$DIP(T_t) = 0 \text{ de otro modo}$$

Básicamente, la Prueba de Depresión permite detectar picos muy largos de un día (mayores o menores) que contrastan con la variabilidad de su entorno. De acuerdo con la definición de un dato periférico (distancia mayor a 1.5), el valor más bajo distinto de cero de un posible dato periférico es el cuadrado de 1.5 (2.25). Por ende, si la prueba de depresión es mayor a 2.25, el siguiente código de amplitud es asociado con el dato:

- DN para valores con datos faltantes en el día anterior o el posterior.
- DM para valores medios entre 2.25 y 3.5
- DL para valores mayores entre 3.5 y 5.5
- DX para valores mayores a 5.5

En el caso de un valor bajo (menor a 2.25), primero se comprueba si el valor es mayor o menor a los valores diarios observados un día antes y uno después del dato a comprobar. Luego se marca el valor si:

- LE si es un mínimo o máximo local
- LN en otro caso

La elección de tres umbrales fue basada en el análisis de los resultados. Finalmente, si la Prueba de Depresión no puede ser computada (por no haber datos suficientes), se calcula la Prueba de Periféricos.

Prueba Espacial

Los eventos extremos pueden ser detectados como errores potenciales utilizando cualquier clase de pruebas, basándose solamente en la distribución de la estación. Por lo tanto es crucial para cualquier procedimiento de control de calidad comparar estaciones con sus estaciones vecinas.

La prueba espacial usa los valores de temperatura de estaciones vecinas calificadas del mismo día (mas el posterior y el anterior) de un evento como marco de referencia para el mismo. Esta información puede confirmar o rechazar dudas sobre un extremo evento, o aún sembrar duda sobre algún valor de temperatura que de otro modo hubiera sido confiable.

La prueba evalúa cada estación vecina dentro de un perímetro de la estación en cuestión. Empieza preseleccionando estaciones vecinas en base a cuán bien se encuentran relacionadas con la estación, utilizando la prueba de Pearson y una regresión lineal de las series.

Los valores de temperaturas para un día dado en los vecinos elegidos son interpolados, de acuerdo a la correlación entre cada vecino y la estación. Esta interpolación resulta en un valor esperado de temperatura para la estación, y en un nivel de confianza para la propia vecindad.

Finalmente, el valor interpolado de temperatura se compara con el valor observado en la estación, resultando en una brecha en la predicción. De acuerdo a ambos la magnitud de la brecha y el nivel de confianza de la vecindad, se marca el valor de temperatura de la estación con un Código de Prueba Espacial.

Este código podría ser N0 si el valor de temperatura concuerda con la estimación de los vecinos, o desde N1 a N4 dependiendo de la brecha en la predicción y el nivel de confianza de la vecindad. N4 implica una brecha amplia y un gran nivel de confianza en la predicción de la vecindad.

En el Apéndice 1 (APACH: Prueba Espacial) se muestra con más detalle la Prueba Espacial.

Finalmente, el algoritmo de árbol de decisión para el control de calidad de los valores de temperatura se puede resumir en lo siguiente (ver Figura 4):

- Computar la Prueba de Depresión y asignar el código resultante al dato. Será uno de DX, DL, DM, LE, LN o DN dependiendo del valor de la prueba.
- SI el código de la Prueba de Depresión es DN, se aplica la Prueba de Periféricos (código OX, OL, OM u O0 de acuerdo al resultado).
- Si la prueba de Depresión es cero o no es computable, se aplica la Prueba de Series Constantes y se mide su longitud, marcando a los valores idénticos de acuerdo a su longitud (SL para series largas de 3 días consecutivos o más, o SS para series cortas de 2 días consecutivos). Si una serie constante es detectada, el valor de temperatura es marcado con su código aún si hay además algún código O0, OM, OL o LN. Los códigos OX tienen prioridad sobre los códigos de series constantes.
- Se aplica una prueba de vecindad, llevando a uno de los códigos NN (sin vecinos, “no neighbour”), N0 a N4 (relacionados con la distancia y el ángulo de las estaciones vecinas).
- Si el valor tienen un código de Depresión LE, la misma prueba de depresión se aplica al valor interpolado según la vecindad, y el resultado simple Ni de la prueba se complementa con códigos informativos. NEO: valor de vecindad extremo de signo opuesto; NEI: valor de vecindad extremo de idéntico signo; NEZ: valor de vecindad no extremo y NEN significando que el valor de vecindad no pudo ser computado.

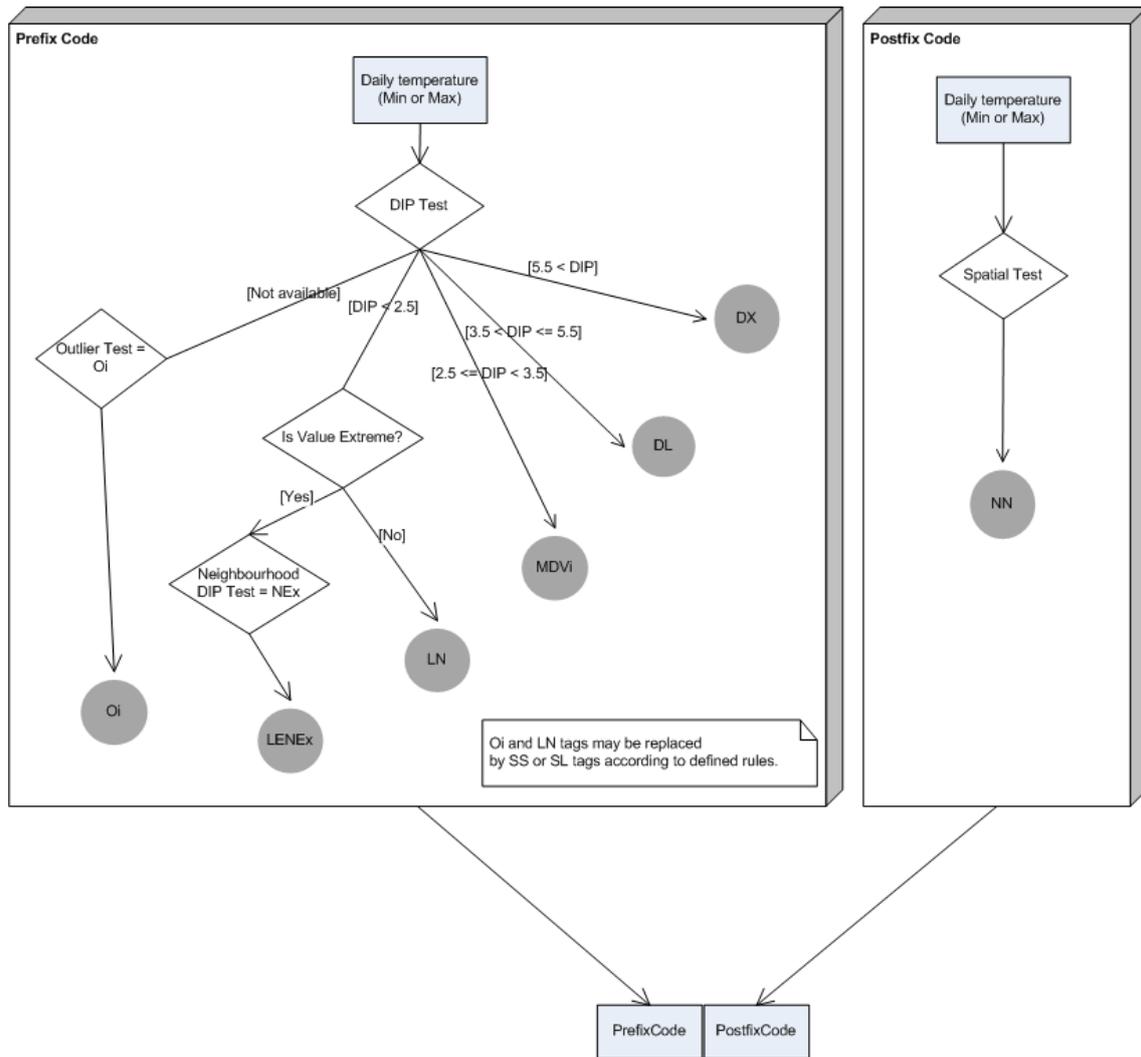


Figura 4 – Árbol de decisión para las temperaturas diarias mínimas y máximas (ver Apéndices 1 y 2 para las expresiones en las hojas)

Aquí se deben hacer dos comentarios importantes: Primero, la prueba de vecindad es una prueba muy importante. Y en casos donde no hay datos de vecindad disponibles, el nivel de confianza va a ser más que nada Useful (útil) o Need-Check (a comprobar). Esto puede ser considerado como una medida precavida para balancear en contra de la falta de datos. Segundo, La salida principal del método es un código describiendo el camino del dato diario observado en el algoritmo del árbol de decisión.

El Apéndice 1 (APACH: Prueba Espacial) describe cada código de prueba utilizado en las mismas, y muestra la tabla de confianza por defecto (nivel de confianza para cada código de control de calidad de la temperatura). El usuario puede reemplazar esto con su propia tabla de confianza.

1.2.5 Pruebas y árboles de decisión para precipitación

El procedimiento para controlar la calidad de los datos de precipitación se basa en dos clases de pruebas:

1. Precipitaciones extremas diarias
2. Secuencias extremas de sequías

Prueba de precipitaciones extremas diarias

Considerando la distribución, forma y amplitud de la precipitación diaria, puede ser relativamente difícil asegurar si un dato diario de precipitación está mal o es siquiera dudoso. Por lo tanto, el algoritmo fue mas bien diseñado para identificar cuan “extremo” es un valor diario de precipitación, dejando al usuario decidir si considerar a tal dato en su análisis. El algoritmo está basado a grandes rasgos en los siguientes pasos:

1. Para cada mes del año (de enero a diciembre), la amplitud de la precipitación se transforma en la distancia al percentil 75% de la distribución diaria del mes. SI la distancia es mayor a 1.5 veces la diferencia de amplitud entre los percentiles 75% y 50%, se marca al día como un potencial valor periférico.
2. Empezando desde el menor valor potencial periférico, se computa si la diferencia de amplitud con el valor periférico siguiente es mayor que el 50% de su amplitud. Si es menor, se descarta al día como periférico. Este método permite desmarcar a potenciales periféricos, los valores de aquellos que son continuos con otros valores de la distribución.
3. Si la distribución de precipitación diaria del mes tienen suficientes observaciones (30 como mínimo), se calcula una distribución paramétrica y se ajusta a la distribución existente (el mejor ajuste entre una Gamma y una Weibull). Luego, se obtiene y remueve el mayor valor de la distribución y se realiza un nuevo ajuste. Si las dos distribuciones ajustadas son significativamente diferentes (de un 5% de acuerdo a la prueba de Kolmogorov-Smirnov), se marca el día una segunda vez. Este método se aplica a todos los potenciales periféricos removiendo la cola de la distribución en un caso y la cola y su valor consecutivo en el segundo caso. Si la prueba da negativa para todos los potenciales periféricos, son todos desmarcados. Sin embargo, si un potencial periférico es marcado por segunda vez, también se marcan todos aquellos con mayor amplitud.

Considerando que muy pocos casos son marcados, un se ha desarrollado una prueba espacial automática, y se deja a un procedimiento de control manual la decisión de aceptar o no al valor. Finalmente, también se computa si el valor marcado fue observado un lunes, pues puede representar a un valor acumulado (domingo más lunes).

Pruebas de secuencias extremas de sequías

Esta prueba se estructura de la siguiente manera:

- a. **Elegir las sequías a ser comprobadas:** Primero, se identifican todas las secuencias de días secos y se computa la distribución de secuencias secas (de la más corta a la más larga). Todas las secuencias secas de más de 10 días y aquellas mayores al 50% de la duración de la secuencia inmediatamente anterior, son marcadas. Tal comprobación es bastante conservadora ya que el objetivo es reducir el tiempo de computación al filtrar secuencias cortas sin perderse de cualquier secuencia seca dudosa. Todas las secuencias de más de un año se codifican LLD (Long Length Drought: sequía de larga longitud) y se consideran sin confianza (valor Suspect).

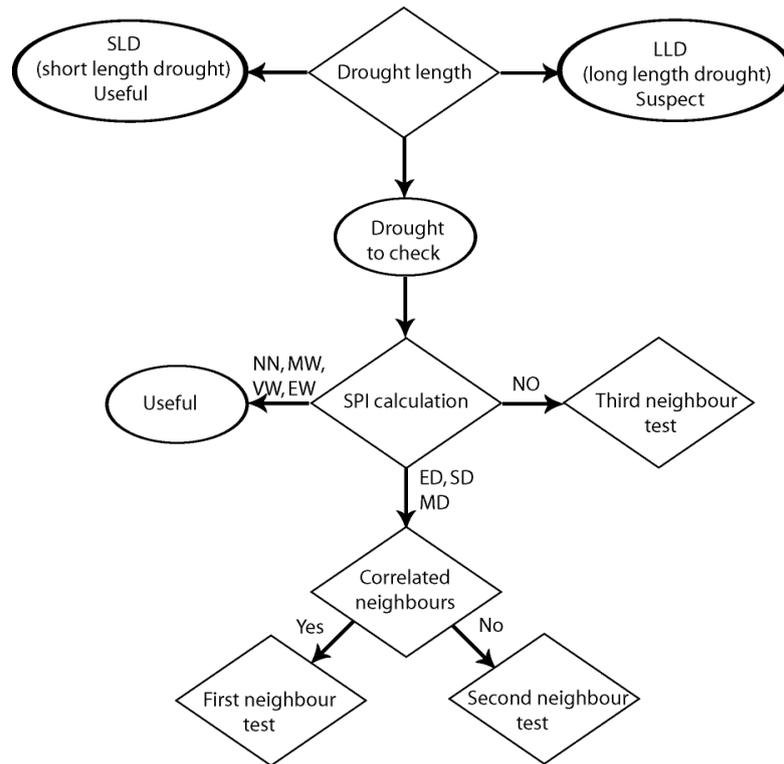


Figura 5 – Árbol de decisión de sequías

- b. **Pruebas de la distribución de la estación:** Para cualquier secuencia seca previamente marcada, se computa la precipitación media sobre el mismo período de tiempo (ejemplo, del 15 de enero al 10 de abril) durante todos los años registrados en la estación. Primera se comprueba si hubo alguna otra sequía en la misma ventana de tiempo en otro año. De ser así, se considera a la sequía como factible, y se marca a la sequía con un valor igual al número de otras secuencias existentes.
- c. **Pruebas de vecindad de estaciones:** Se desarrollaron tres pruebas de vecinos: En cada una, primero se seleccionan todas las estaciones a menos de 200km de distancia. Luego se computa en la estación actual la precipitación media caída en la misma ventana de tiempo que la secuencia a ser comprobada (ejemplo, del 15 de enero al 10 de abril) para cada año de la estación. En todos los casos, si existe una sequía existente en el mismo período en una estación vecina, la secuencia se marca como útil (Useful, marca ND). En caso contrario, se procede con una de las tres pruebas de vecinos tal como se muestra en el árbol de decisión (Figura 5). En la primera prueba de vecino, se calcula el rango de correlación entre las series de media de precipitación (media durante la ventana de tiempo equivalente al de la secuencia que está siendo comprobada) entre la estación de referencia y cualquier estación vecina. Si es menor a 0.8 o no llega a un nivel de significancia del 99%, se aplica la segunda prueba de vecinos. La tercera prueba se aplica cuando las series en la estación de referencia son menores a 20 años. Si no se encuentran vecinos disponibles, se marca con N. El apéndice 3 provee de información más detallada sobre las pruebas de vecinos así como de las tablas que relacionan las marcas resultantes de las pruebas con los niveles de confianza.

1.2.6 Aplicación al Servicio Meteorológico Nacional Argentino

1.2.6.1 Resultados en temperatura

Primero de todo, es importante notar que más del 90% de los datos de temperaturas mínimas y máximas fueron considerados como útiles (Useful) luego de aplicar el árbol de decisión. Entre el 5% y el 10% de los datos marcados fueron hechos como “a comprobar” (Need-Check, ver Figura 6), mostrando una tendencia positiva debido a la tendencia negativa en la densidad de la red. Menos del 0.1% de los datos (aprox. 3000 registros) fueron marcados como “dudosos” (Doubtful) y alrededor del 0.05% como “erróneos” (Suspect, aprox. 1500 registros).

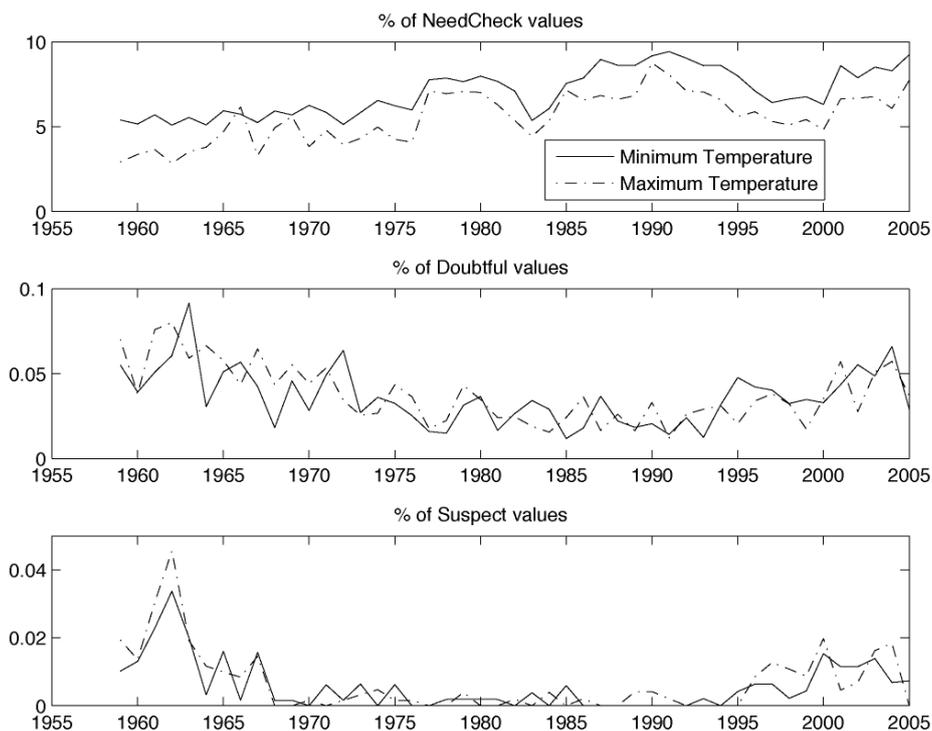


Figura 6 – Porcentajes de observaciones de temperatura mínima (línea continua) y máxima (línea punteada) clasificadas como NeedCheck (panel superior), Doubtful (panel al medio) y Suspect (panel inferior)

Esta figura muestra que las dos curvas (dudosas y erróneas) para las temperaturas mínimas y máximas son similares. Ambas tienen valores altos al principio de la serie temporal, cuando la mayoría de los datos fueron registrados y archivados en tarjetas de papel, las cuales, en algunos casos, fueron dañadas por el polvo y la humedad. En efecto, antes de octubre de 1967, no se contralaba la calidad de los datos. Luego de esa fecha, hubo un período largo de valores muy bajos en el nivel de errores, correspondiente a 1967-1997, cuando un control de calidad manual era realizado por el Servicio Meteorológico Argentino. También se aprecia un incremento hacia el final del período de los valores erróneos, cuando sólo se aplicaron controles de calidad en tiempo real. Estos resultados demuestran que los controles de calidad ejecutados por el Servicio eran buenos, pero dejaron algunos potenciales errores identificados por estas nuevas pruebas. Como consecuencia, la serie temporal de marcas muestra una variabilidad consistente con la historia del Servicio Meteorológico y sus procedimientos de control de calidad.

Finalmente, la distribución espacial de los errores evidencia claramente (Figura 7) que para las estaciones al sur de los 40°S, sin vecinos, es imposible computar las pruebas de estaciones de vecindad para la temperatura mínima, las cuales son cruciales para marcar un dato como erróneo de manera inequívoca. Sin embargo, para la temperatura máxima (Figura 8), la prueba se pudo aplicar a algunas estaciones permitiendo la detección de valores dudosos y erróneos. En regiones de alta densidad espacial de estaciones, se detectaron mayores proporciones de datos erróneos en ambos conjuntos de datos (temperatura mínima y máxima). Esta diferencia se explica por el hecho que la variabilidad en la temperatura mínima es más local (menos correlacionada espacialmente) que la variabilidad de la temperatura máxima.

Símbolos utilizados en las Figuras 7 y 8:

- Figuras 7 y 8, a la izquierda: Círculos pintados representan a un radio de valores marcados como “a comprobar” (Need Check) mayores al 15%, círculos huecos representan a un radio de entre 5% y 15%, y cruces representan a un radio menor al 5%.
- Figuras 7 y 8, al medio: Círculos pintados representan a un radio de valores marcados como “dudosos” (Doubtful) mayores al 0.1%, círculos huecos representan a un radio de entre 0.04% y 0.1%, y cruces representan a un radio menor al 0.04%.
- Figuras 7 y 8, a la derecha: Círculos pintados representan a un radio de valores marcados como “erróneos” (Suspect) mayores al 0.05%, círculos huecos representan a un radio de entre 0.01% y 0.05%, y cruces representan a un radio menor al 0.01%.

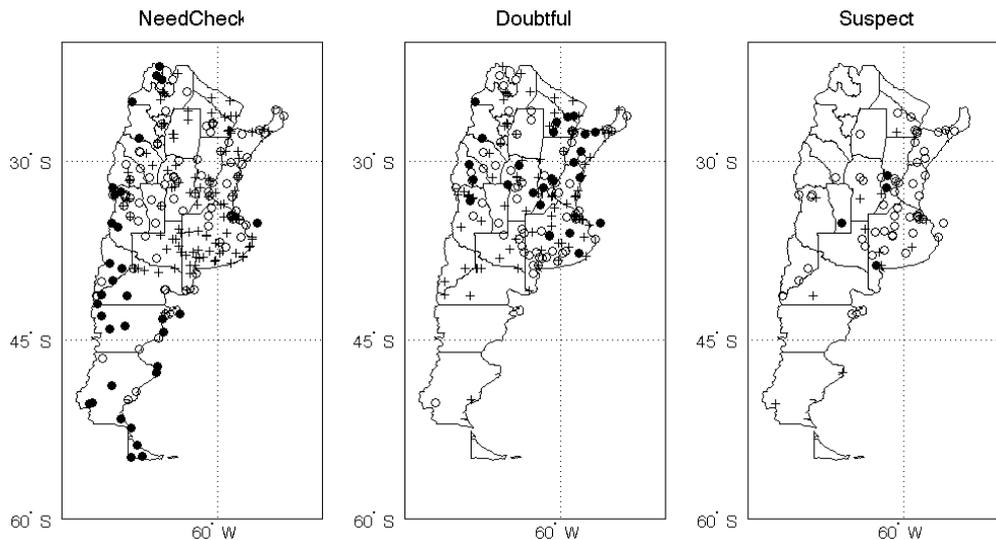


Figura 7 - % de problemas en la temperatura mínima, mostrados por estación y severidad del problema

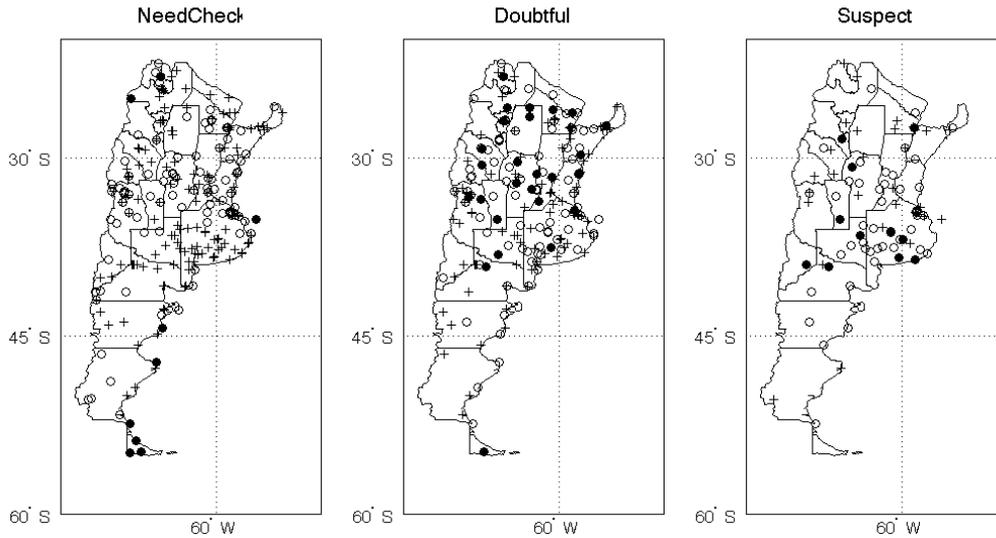


Figura 8 - % de problemas en la temperatura máxima, mostrados por estación y severidad del problema

Este resultado sugiere que algunos datos erróneos probablemente existan en registros de estaciones aisladas, pero no pueden ser clasificados como tales utilizando solamente la distribución de las estaciones como se hace en las pruebas. También sugiere que el desarrollo de la red en el futuro debería tomar en cuenta la pobre densidad de la red en las partes sureñas del país. Es posible que la comparación con estaciones automátatas pertenecientes a redes privadas pueda completar parcialmente a la red del Servicio Meteorológico Nacional y contribuir a una mejora en la detección de errores en el futuro.

1.2.6.2 Resultados de precipitación

Con las pruebas de Precipitaciones Diarias Extremas acerca de 80 valores potencialmente periféricos por año fueron identificados antes de 1976, y acerca de 20 a 40 luego. Esta diferencia puede tener dos explicaciones: La primera, previo a 1967, no se aplicaron controles de calidad a los datos, lo cual puede explicar por qué se detectaron mayores casos periféricos. La segunda, la diferencia entre los períodos antes y después de 1976 puede resultar del hecho de que luego del Golpe Militar, muchas estaciones, especialmente en el interior del país, fueron clausuradas.

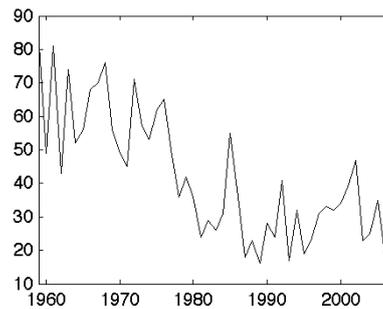


Figura 9 – Variabilidad interanual del número de potenciales periféricos en la precipitación diaria

Estas estaciones se ubicaban en regiones más secas que aquellas ubicadas en la Pampa Húmeda. Más aún, una gran proporción de las estaciones restantes se encuentra en aeropuertos y tienen un valor estratégico, el cual ciertamente favorece a un mejor control de calidad en tiempo real. Como se explicó previamente, también se comprobaron si los valores de precipitación

potencialmente periféricos fueron observados un lunes. Justamente se encontró que así fue el caso para un 17% de tales valores. Aunque este dato es mayor que un proceso estocástico (14.3%), se encuentra bastante cerca. En esta etapa, se cree que un “Control Humano” es requerido para verificar los valores de precipitación marcados por el método.

La prueba de Secuencias Secas Extremas (Figura 10) muestra un gran pico durante el período invernal, alcanzando el doble de números de días secos “a comprobar” (Need-Check) que durante el resto del año. Mientras que la variabilidad interanual muestra varios picos (Figura 11), también evidencia (Figura 12) un número mucho mayor de valores no útiles (aprox. Un 4% de todos los días secos) al principio de las series temporales, antes de los controles de calidad del SMN. Más aún, se observa un mínimo antes del Golpe Militar en 1976. Desde esa fecha, una pequeña tendencia positiva se ve en el porcentaje de valores no útiles.

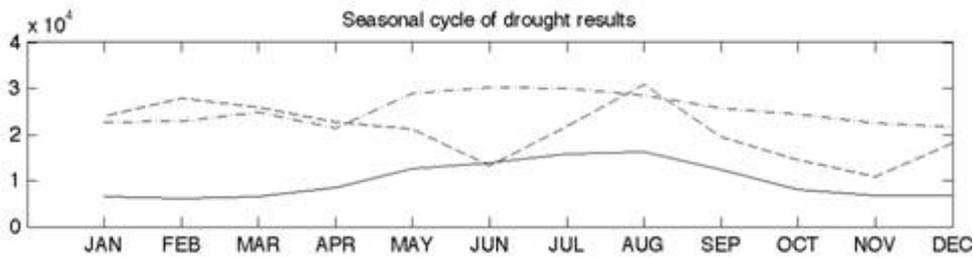


Figura 10 – Número de días secos clasificados como NeedCheck (línea continua), Doubtful (guiones) y Suspect (guiones y puntos) por mes. Los casos Doubtful y Suspect cases fueron multiplicados por 100

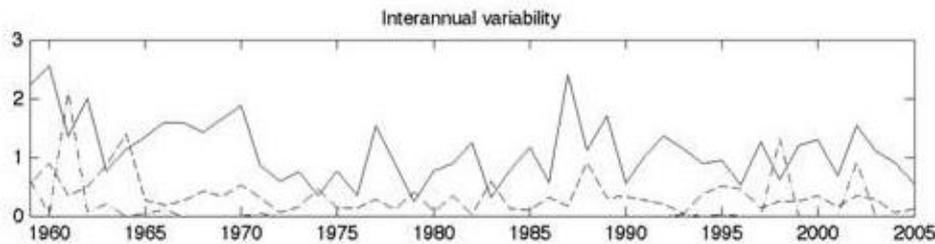


Figura 11 – Variabilidad interanual del porcentaje de casos de NeedCheck (línea continua), Doubtful (guiones) y Suspect (guiones y puntos)

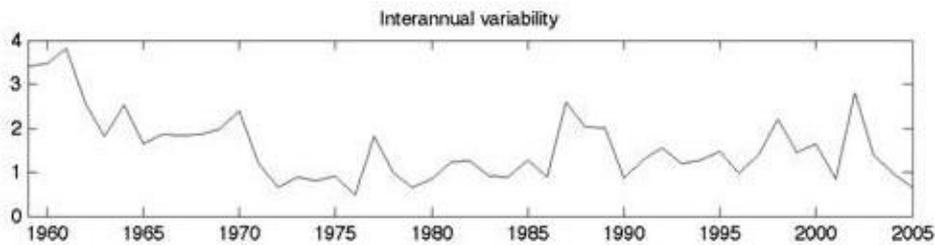


Figura 12 – Variabilidad interanual del porcentaje total de casos de NeedCheck, Doubtful y Suspect

1.2.7 Conclusiones y perspectivas

Uno de los tres objetivos estratégicos del Proyecto CLARIS era iniciar el armado de una base de datos climática diaria de alta calidad para la temperatura y la precipitación, pues sería de gran valor para validar y evaluar la habilidad de los modelos en simular tendencias climáticas y cambios frecuentes en los eventos extremos. Distintos grupos han contribuido a objetivos similares usando sus propios controles de calidad de los datos (a menudo manuales o visuales).

El proyecto CLARIS LPB (7th EC Framework Programme, 2008-2012) ahora apunta a la obtención de cientos de registros de estaciones oficiales y privadas, tanto de precipitación como de temperaturas mínimas y máximas. Es por lo tanto crucial desarrollar un sistema automatizado para ejecutar el control de calidad. El presente trabajo fue el primer paso hacia tal objetivo.

En el presente estudio, la precipitación diaria y las temperaturas mínimas y máximas fueron comprobadas separadamente. Una mejora potencial del método (especialmente en regiones sin estaciones vecinas) sería comprobar diferentes tipos de datos diarios. El primer árbol de decisión se aplica sobre la temperatura mínima o sobre la máxima. La prueba central se llama Prueba de Depresión (Dip Test) y mide cuán alto es un pico positivo o negativo (cuando ocurre) en una secuencia de 3 días. Esta prueba es completada con una Prueba Espacial confirmando si el dato específico puede ser marcado como “útil” (Useful), “a comprobar” (NeedCheck), “dudoso” (Doubtful) o “erróneo” (Suspect). La categoría “a comprobar” a menudo incluye datos que fueron marcados pero para los cuales no hay suficiente información complementaria (como estaciones vecinas) disponible. Es importante notar que cada dato diario se codifica con una cadena de caracteres que sigue el camino que atravesó en el árbol de decisión. La marca final (Useful, NeedCheck, Doubtful o Suspect) es una sugerencia y puede ser modificada por los usuarios a su discreción.

El control de calidad de los datos diarios de precipitación es mucho más complejo. Por lo tanto se han aplicado dos árboles de decisión distintos, uno para las precipitaciones extremas, y otro para las sequías largas. La prueba de precipitación extrema apuntaba a detectar valores muy altos de precipitación. Los datos marcados fueron mucho más numerosos antes de 1976 que luego de tal momento. Este resultado sugiere que, aunque es difícil detectarlos, algunos valores altos y erróneos de precipitación pueden existir en la base de datos, requiriendo de un control de calidad manual de los datos marcados. Este proceso será hecho en el futuro por el Servicio Meteorológico Nacional. La prueba de sequías identificó secuencias muy largas, particularmente utilizando el *Standardized Precipitation Index* (SPI, McKee et al., 1993). La prueba de sequías fuertemente necesita de una vecindad de estaciones. Como en el caso de la temperatura, mas sequías (marcadas como NeedCheck, Doubtful o Suspect) fueron encontradas al principio del período. Un mínimo se observó a comienzos de los '70 antes del Golpe Militar (1976). Desde entonces, se ha detectado una débil tendencia positiva en el número de días secos (especialmente NeedCheck). Es posible que la disminución en el número de estaciones activas desde 1976 haya reducido el número de estaciones vecinas en las pruebas, y por ende aumentado el número de marcas NeedCheck.

Los resultados también sugieren que hay algunos datos erróneos (Suspect) en los registros de estaciones aisladas, pero no pueden ser clasificados como tales utilizando solamente la distribución de la estación. Esto sugiere que el desarrollo de la red en el futuro debería tomar en cuenta la pobre densidad en las regiones sureñas del país. Es posible que con una colección más amplia de estaciones automáticas de datos pertenecientes a redes privadas puedan completar parcialmente a la red del Servicio Meteorológico y contribuir a una mayor detección de errores en el futuro. El método se aplicará en el entorno del proyecto CLARIS LPB, el cual apunta a obtener un gran número de estaciones tanto de redes oficiales como privadas en la Cuenca del Plata para ofrecer un mejor entendimiento en la variabilidad diaria regional del clima y crear un enrejado de productos diarios, útiles para validar modelos globales y regionales como se ha hecho en el ENSEMBLE EC Project (6th FP; <http://ensembles-eu.metoffice.com/>).

1.3 Implementación de la solución

En esta sección se describe la implementación de APACH [1].

En la sección 1.3.1 revisaremos los requerimientos de la solución, las razones que condujeron las decisiones tomadas, y en la sección 1.3.2 se mostrará la arquitectura que se eligió para su desarrollo.

En los apéndices se podrá encontrar más detalle sobre este desarrollo, tanto desde el punto de vista del usuario (apéndice 4, Referencia para el usuario de APACH) como desde el punto de vista del desarrollador (apéndice 5, Implementación de APACH).

1.3.1 Requerimientos de la implementación

A la hora de definir la implementación de APACH, se tomaron las siguientes decisiones:

- **Portabilidad:** La implementación debería ser portable entre distintos sistemas operativos. Esto por una parte restringió la decisión a un subconjunto tecnologías, en todos los rubros abarcados, y por otra parte facilitó el desarrollo, que se realizó en maquinas de distinta naturaleza.
- **Presentación:** La implementación se debería poder correr en una PC de escritorio, no requerir conexión a Internet, y ser utilizable por un usuario sin conocimientos de desarrollo de Software. Esto moldeó la arquitectura de la aplicación, ya que se requiere poder instalar la misma de forma automática. También implica un grado alto de interacción con el usuario, mayor al que podría requerir una aplicación web.
- **Eficiencia:** Si bien no se manejaron requerimientos específicos de eficiencia, siempre fueron un foco de atención durante el desarrollo. Los procesos largos deberían mantenerse al mínimo necesario y en todo caso deberían ser predecibles y anulables.

En base a esto se decidió que la solución se desarrollaría bajo tecnología Java, en particular versión 5 SE compatible, utilizando soporte en base de datos transaccional.

Para facilitar la interacción entre la aplicación Java y la base de datos fue muy importante el uso de Hibernate, que permite a la aplicación Java independizarse de la tecnología subyacente de base de datos, con lo que se pudo, por ejemplo, utilizar como soportes dos tecnologías SQL sumamente distintas: MySQL de Sun durante el desarrollo, y Apache Derby hacia el final del desarrollo. La primera ofrece excelente eficiencia y soporte para varias plataformas (como Windows, Linux y Mac), mientras que la segunda, al embeberse en la aplicación Java permite un proceso de instalación y posterior administración totalmente automático, además de depender solamente del entorno Java, desvinculándose absolutamente del sistema operativo.

Al incorporar Apache Derby, la instalación del programa pasa a incluir el motor de base de datos, los datos y el programa de control de calidad en un solo paquete.

1.3.2 Arquitectura

A grandes rasgos, la arquitectura de la aplicación resultante se muestra en la figura 13.

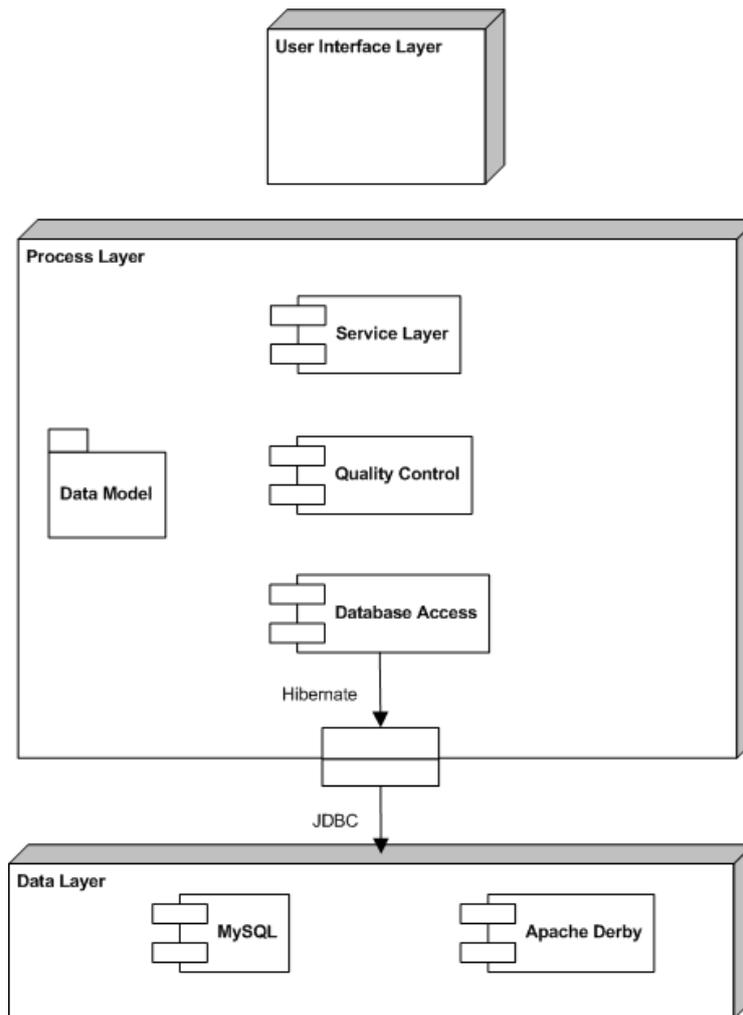


Figura 13 – Arquitectura de la aplicación

Interfaz de usuario

Al menos por esta primera instancia del desarrollo se decidió que el software debería ser presentado como una aplicación de escritorio, con la que un usuario final pudiera interactuar con los procesos de control de calidad desarrollados. Esta necesidad de interacción con un usuario final implicó en la práctica ajustar muchos aspectos del sistema, tales como optimización temporal, dentro de lo posible, de los procesos más importantes del sistema, o estructurar las implementaciones de los cálculos de forma tal que se pudiera ofrecer permanentemente feedback sobre el estado actual de un proceso.

Control de calidad y modelo de datos

Esta capa expone ciertos puntos de entrada a la interfaz de usuario e implementa toda la lógica de control de calidad y modelo de datos.

Se incluye en esta capa la conexión a la base de datos a través de Hibernate. El componente juega un rol importante en este aspecto, ya que es responsable (justamente utilizando Hibernate) de administrar la estructura de la base de datos.

Capa de datos

Esta capa está implementada por una base de datos relacional (SQL). Si bien utilizar un motor de base de datos en principio puede parecer un overhead importante para el sistema, por otro lado se espera que la aplicación soporte grandes volúmenes de datos, y una base de datos relacional es lo ideal para manejar los tipos de datos que usa esta aplicación. Se utilizaron índices en algunas tablas seleccionadas como para aumentar el rendimiento de consultas que se realizaban frecuentemente. Al finalizar el desarrollo se cuenta con una implementación MySQL y una Apache Derby.

1.4 Conclusiones

Se presentó el trabajo realizado en control de calidad de series meteorológicas. Este trabajo resultó en una herramienta terminada, multi-plataforma y sencilla de utilizar por un usuario final. Aquí se implementaron controles de calidad desarrollados para la ocasión, pero reutilizables en cualquier contexto y marco geográfico.

Esta herramienta fue utilizada más adelante como parte del trabajo, para depurar la base de datos del SMN. Luego se valió de tal base en la interpolación presentada en la sección siguiente: “Estimación de variograma”.

El desarrollo de APACH tuvo la peculiar característica de definir sus requerimientos sobre la marcha, por lo cual no siguió ninguna metodología convencional de desarrollo de software. Esto también implicó código fuente muy inestable durante todo el desarrollo (con más del 50% del código desechado en el camino) y un gran énfasis en los datos y resultados por sobre la generalidad de los procesos. Sin embargo se obtuvo una base de código estable, reutilizable y funcional, con potenciales direcciones a seguir y una fundación a utilizar para ocasionarlo.

El segundo paso natural de esta herramienta, manejado durante la etapa de desarrollo, fue otorgarle una interfaz web y capacidades multi-usuario. Esto cambiaría las premisas de la aplicación actual, desechando la interfaz gráfica de escritorio y la restricción de concurrencia a un solo usuario en simultáneo. Sin embargo estos cambios fueron contemplados durante el desarrollo, y los pasos a seguir están claros. Vamos a mencionar los grandes pilares que implicaría un cambio así:

- Cambio del cliente de escritorio por un cliente web: Esto implicaría re-escribir todo el cliente, manteniendo la capa de servicios y persistencia en un servidor web. La idea a priori para escribir el cliente sería utilizar JSP, GWT o cualquier herramienta Java de desarrollo web para una fácil integración con la aplicación actual. El código que habría que escribir no tendría mayores desafíos de interacción con el resto de la aplicación, ya que las APIs de los servicios son lo suficientemente simples para invocarse aún desde una interfaz gráfica.
- Transaccionalidad de la capa de servicios: En un entorno multi-usuario, sería conveniente que los servicios sean transaccionales. Es decir que el código de un servicio este embebido en marco de conexión a la base de datos. Hay herramientas que hacen esto automáticamente, por ejemplo, Spring Framework, que si bien no fue utilizada en la aplicación, fue contemplada todo el tiempo para tener una fácil integración. Este cambio quitaría a los servicios la responsabilidad de manejar transacciones.

- **Multi-threading:** La gran mayoría del código de la aplicación tendría que estar preparado para un uso multi-hilo. Esto está parcialmente contemplado. Aquí de nuevo Spring soporta hilos múltiples naturalmente, y sería de gran asistencia para dar soporte a esto en la capa de servicios. Hibernate (y el uso que se hace de este) hace lo propio en el estado actual de la aplicación, por lo que las conexiones a bases no tendrían mayores problemas.
- **Multi-usuario:** Sí habría que contemplar la posibilidad de muchos usuarios buscando acceder a los mismos datos. Para esto, el concepto de dataset provee un camino de extensión natural, manteniendo separadas distintas bases. El aspecto multi-usuario sería especialmente delicado ya que un proceso sobre un dataset podría tardar horas enteras, durante las cuales no debería poder utilizarse el mismo. Se puede resolver esto con una solución de bloqueado sencilla a nivel dataset.
- **Depuración de código:** Para poder ofrecer un servicio multi-usuario estable y eficiente haría falta depurar y testear el código existente más allá de lo hecho hasta el momento. Esto no tiene ningún riesgo asociado, y nuevamente hay herramientas que podrían ser de gran utilidad. Spring provee naturalmente una gran sinergia con el testing unitario en general y JUnit en particular, bajando el acoplamiento de las entidades de la aplicación y ofreciendo infraestructura propia para asistir en el testing. Herramientas como Selenium permiten testear automáticamente una aplicación web, asegurando que la navegación y la ejecución de los casos de uso se puede realizar según lo planeado.

1.5 Bibliografía

- [1] J. P. Boulanger, J. Aizpuru, L. Leggieri, M. Marino, 2008: A Procedure for Automated quality Control and Homogenization of historical daily temperature and precipitation data (APACH). Part 1: Quality Control and Application to the Argentine Weather Service Stations. Publication Pending
- [2] Haylock M., T. C. Peterson, L. M. Alves, T. Ambrizzi, Y. M. T. Anunciação, J. Baez, V. R. Barros, M. A. Berlato, M. Bidegain, et. al.: Trends in total and extreme South American rainfall 1960-2000 and links with sea surface temperature. *Journal of Climate*, 19, 1490-1512, 2006.
- [3] McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. *Proc. Eighth Conf. of Applied Climatology*, Anaheim, CA, Amer. Meteor. Soc., 179–184.
- [4] Rusticucci, Matilde y M. Renom, 2007 “Variability and trends in indices of quality controlled daily temperature extremes in Uruguay.” *Int. J. of Climatol.* DOI: 10.1002/joc.1607
- [5] Rusticucci, Matilde, y M. Barrucand, 2004 "Observed trends and changes in Temperature Extremes over Argentina". *J. Climate* vol 17, No. 20, 4099-4107.
- [6] Vejen, F., C. Jacobsson, U. Fredriksson, M. Moe, L. Andresen, E. Hellsten, P. Rissanen, B. Palsdottir, and B. Arason, Quality Control of Meteorological Observations : Automatic Methods Used in the Nordic Countries, Report 08/2002, KLIMA, pp. 111, 2002.
- [7] Vincent, L. A. , T. C. Peterson, V. R. Barros, M. B. Marino, M. Rusticucci, G. Carrasco, E. Ramirez, L. M. Alves, T. Ambrizzi, M. A. Berlato, et. al., Observed Trends in Indices of Daily Temperature Extremes in South America 1960–2000, *Journal of Climate*, 18, 5011-5023, 2005.

2 Estimación de variograma

2.1 Resumen

El objetivo de este subproyecto fue desarrollar un modelo de variograma para series diarias de temperatura mínima y máxima localizable para regiones geográficas acotadas. Este modelo fue basado sobre redes neuronales de tipo MLP, ya que ofrecen un poder de estimación más preciso que los modelos estadísticos usados en la literatura. El principal fin de este modelo será la interpolación de temperatura.

Describiremos entonces un modelo de variograma de series de temperatura basado en redes neuronales, y lo compararemos a dos modelos estadísticos utilizando como base de datos de referencia el registro histórico del Servicio Meteorológico Nacional, depurado a partir del control de calidad con APACH. Este registro se presentará acotado al período 1961-1990, período de referencia para los modelos climáticos que analizaremos en secciones subsiguientes.

En la sección 2.2 se describirá el concepto de variograma y su uso en la geoestadística. En la sección 2.3 se presentará el conjunto de datos utilizado, en la sección 2.4 se presentarán dos modelos estadísticos de variograma que se tomarán como referencia para presentar la metodología desarrollada, que será a su vez ampliada en la sección 2.5.

En la sección 2.6 se mostrarán los resultados obtenidos y finalmente en la sección 2.7 se hará un cierre del trabajo en estimación de variograma, con vistas al futuro, y presentando la utilización del mismo dentro del marco general del estudio de cambio climático.

2.2 Introducción al problema

En el interés de evaluar modelos de clima, es necesario contar con mapas de temperaturas (mínima y máxima) diarias que presenten valores en una grilla regular de precisión arbitraria, ya que este es el formato tradicional en el que se presentan datos meteorológicos generados a partir de los modelos climáticos.

El presente trabajo es parte de un esfuerzo por interpolar temperaturas de una base de datos de estaciones meteorológicas con el fin de obtener esta información presentada en forma de grilla regular.

En Haylock et al., se referencia al Kriging como uno de los métodos más robustos para la interpolación espacial de datos climáticos. Kriging es un conjunto de técnicas geoestadísticas para interpolación de procesos estocásticos en puntos sin observaciones, a partir de observaciones de su valor en ubicaciones vecinas.

Kriging computa el mejor estimador no sesgado $\hat{Z}(x_0)$, basándose en un modelo estocástico de la dependencia espacial, cuantificada por el Variograma $\gamma(x,y)$ o por la esperanza $\mu(x) = E[Z(x)]$ y la función de covarianza $c(x,y)$ del proceso estocástico.

En el marco de este proyecto se desarrolló un procedimiento para interpolación de datos diarios de temperatura, que fue separado en 2 pasos según los conceptos de Kriging.

El primer paso consistió en un modelo de variograma para datos de temperatura; el resto de este trabajo hará foco sobre tal modelo de variograma.

El segundo paso continuó a partir de lo presentado aquí para desarrollar un interpolador de temperatura utilizando el variograma estimado. Dicho interpolador se encuentra fuera del alcance de este informe y fue presentado en Pausader (2008).

2.2.1 Variograma

El variograma es una función que describe el grado de dependencia espacial de un proceso estocástico $Z(x)$, y se define como el incremento esperado (al cuadrado) de los valores entre un punto X y un punto Y .

$$2\gamma(x, y) = E(|Z(x) - Z(y)|^2)$$

Si existe una función de covarianza para este proceso, el variograma se caracteriza como

$$2\gamma(x, y) = C(x, x) + C(y, y) - 2C(x, y)$$

Existen en la literatura distintos modelos de aproximación del variograma, basados en distribuciones estadísticas y que se alimentan exclusivamente la distancia entre los puntos de los pares $\langle X_1, X_2 \rangle$. En esta sección se estudian estos modelos, y se presenta uno nuevo basado en redes neuronales de tipo MLP, que permite modelar relaciones complejas sin conocer su distribución subyacente, y en este caso permite incorporar al modelo información adicional a la distancia del par.

Dado que buscamos cubrir con un modelo espacios geográficos arbitrariamente grandes, es natural que estos comprendan zonas con distintos comportamientos climáticos. Vamos a disminuir los efectos estacionales que puedan darse durante el año acotando cada modelo de variograma a un mes especial del año. Así, para una zona geográfica determinada se puede hablar de 24 escenarios distintos: Uno por cada mes del año y cada tipo de medición de temperatura (mínima y máxima).

2.3 Conjunto de datos

Como mencionamos previamente, para este trabajo se utilizó el conjunto de estaciones del Servicio Meteorológico Nacional de Argentina, para el período 1961-1990.

Antes de utilizarlo, se aplicó el control de calidad APACH, referido en la primera parte de este informe. Esto arrojó como resultado un 2% de registros diarios con niveles de confianza de tipo 1 (Need Check), 2 (Doubtful) o 3 (Error), que fueron removidos de la base.

Luego se seleccionaron, para cada tipo de temperatura por separado, las estaciones que contaran con al menos 75% de registros diarios con información para el período en cuestión. Al final de este proceso, obtuvimos un conjunto de datos con 94 a 108 estaciones.

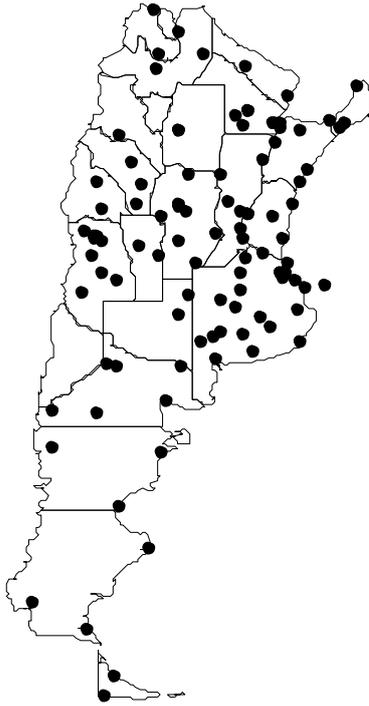


Figura 13 - Distribución de estaciones para el mes de Enero (Mínima + Máxima)

En este gráfico se puede ver la distribución de las estaciones del conjunto de datos, acumulada para ambas variables, para el mes de Enero. Si bien hay pequeñas variantes en la completitud según la variable y mes, la distribución geográfica de estaciones de este gráfico se repite en todos los casos. Se puede observar la gran concentración de estaciones en la Pampa húmeda y el litoral, con menos estaciones hacia el noroeste y una densidad muy baja sobre la cordillera y en la Patagonia.

Esto causa que en la provincia de Buenos Aires veamos estaciones meteorológicas con 11 estaciones vecinas en un radio de 100km, lo que contrasta notoriamente con el sur, donde contamos con 10 estaciones en un área que representa casi el 40% del territorio nacional.

A partir de estas diferencias de densidad, es natural esperar resultados de predicción de variograma mucho mejores en la Pampa húmeda que en la Patagonia. De todas formas acá deberían ser apreciables las ventajas del aprendizaje con redes neuronales: La capacidad de localizar la predicción (cuando se cuente con información de posición) y al mismo tiempo reutilizar lo aprendido en regiones con mucha densidad sobre zonas con menos información, en cuanto sirva para optimizar el modelo.

Sobre el registro de temperaturas para las estaciones mencionadas, se tomaron mes por mes y variable por variable, todos los pares de estaciones con separación de 500km máximo. Para cada par resultante se guardaron las siguientes variables: Latitudes, Longitudes y Alturas de las estaciones del par, mas variograma empírico del par.

El dataset a utilizar está compuesto entonces por 24 subconjuntos (doce meses, dos variables) de entre 1600 y 2100 pares de estaciones cada uno.

2.4 Modelos estadísticos de variograma

Dado que no contamos con información de la distribución de temperatura en cada punto de la grilla, nos interesa modelar el variograma entre cada par de puntos, utilizando la información de las estaciones más cercanas en cada caso.

En la literatura se suelen utilizar para esto diversos estimadores del variograma, basados exclusivamente en la distancia entre un par de puntos. Esta asunción es un poco fuerte ya que implican que el proceso es isotrópico y homogéneo a lo largo de todo el dominio. Sin embargo son estimadores muy aceptados. Dos funciones de la distancia h que normalmente se utilizan son (Chiles & Delfiner 1999, Cressie 1993):

Exponencial

$$\gamma(h) = (s - n) \left(1 - e^{-\frac{h}{ra}} \right) + n1_{(0,\infty)}(h)$$

Esférica

$$\gamma(h) = (s - n) \left(\left(\frac{3h}{2r} - \frac{h^3}{2r^3} \right) 1_{(0,r)}(h) + 1_{[r,\infty)}(h) \right) + n1_{(0,\infty)}(h)$$

Para distancia h y parámetros:

- *Nugget* n : La altura del salto del variograma en la discontinuidad en el origen. Permite representar una variación espacial a distancias de órdenes de magnitud inferior a aquellas entre las estaciones de la red.
- *Sill* s : Limite del variograma en distancias infinitas. Describe la correlación esperada entre dos estaciones lejanas.
- *Range* r : Se caracteriza como la distancia a la cual la diferencia del variograma con el parámetro s es despreciable. Es una medida de cuan rápidamente la correlación se reduce con la distancia.
- a : Tiene diferentes valores según la referencia, debido a la ambigüedad en la definición del rango. Ej.: $a = 1/3$ es el valor usado en (Chiles & Delfiner 1999).

La función $1_A(h)$ es 1 si $h \in A$ y 0 en caso contrario.

Para definir en cada conjunto de datos y cada función, los valores de los parámetros, se utilizó regresión no lineal por mínimos cuadrados. En ambos casos, se tomaron los siguientes valores iniciales, extraídos de la literatura:

- $n = 0$
- $s = \max(v)$
- $r = 500$
- $a = 1/3$

Donde \mathbf{v} son los valores de variograma empírico obtenidos entre los pares de estaciones del conjunto de datos, y \mathbf{r} es el umbral utilizado como distancia máxima entre 2 puntos (en kilómetros) para los que se desea calcular variograma. Este valor de 500km fue utilizado consistentemente a lo largo de todo el trabajo.

La función de error a minimizar durante la regresión es la siguiente:

$$E = \sqrt{\text{mean}((y_p - y)^2)}$$

donde “ y_p ” es la predicción de temperatura dada por la función gamma e “ y ” es el valor empírico.

Se tomó como indicador de eficiencia final de los modelos, el error obtenido luego de estabilizarse la regresión.

2.5 Modelo de variograma basado en MLP

Los modelos de variograma de la sección anterior son algunos de los que comúnmente se refieren en la literatura (ver Haylock, 2008).

En esta sección presentamos un modelo de variograma alternativo, basado en redes neuronales de tipo MLP, que ofrecen un mayor poder de expresión. Esta ventaja puede referirse a 3 características de los MLP:

1. Flexibles a cualquier distribución de datos de entrada: Las redes neuronales durante su ciclo de aprendizaje se adaptan al conjunto de datos que se le presente, por lo que pueden ser localizadas a cualquier región o clima sin necesidad de ningún proceso de ajuste manual de parámetros. Por naturaleza, todo ajuste de parámetros internos se realiza durante el entrenamiento.
2. Cuentan con una representación interna arbitrariamente compleja: Un MLP se compone por una cantidad arbitraria de “capas escondidas” y de neuronas en cada una de estas capas. La adición de cualquiera de estas dos dimensiones de la red permite a la misma representar funciones más complejas.
3. Aceptan una cantidad arbitraria de información de entrada: La arquitectura del MLP es extensible por naturaleza. Es sumamente sencillo ingresar nuevas variables al modelo, que serán utilizadas por la red en cuanto aporten información para generar estimaciones de mayor precisión.

2.5.1 Arquitectura

Se utilizaron para las redes MLP arquitecturas con las siguientes características:

- Una neurona en la capa de entrada por cada variable de entrada.
- Una capa escondida con cantidad variable de neuronas, según se defina conveniente.

- Una capa de salida de una neurona para la temperatura.
- Sinapsis completa entre la capa de entrada y la escondida, y sinapsis completa entre la capa escondida y la capa de salida.
- Función de activación lineal para la capa de salida.

La decisión más significativa es la de la cantidad de capas escondidas de la red. Se decidió que fuera capa única ya que este tipo de redes, aun siendo el más simple, permite funciones de órdenes de complejidad mucho mayores a los modelos estadísticos de referencia, por lo que debería servir perfectamente para una comparación entre ambas metodologías. Al mismo tiempo, esta elección permite tiempos de entrenamiento de las redes muy inferiores a las que demandaría una red de más capas.

2.5.2 Variantes

Una de las ventajas de resolver la estimación de variograma con redes neuronales es la posibilidad de agregar información adicional al modelo sin tener que describir funciones proporcionalmente complejas. La arquitectura de un perceptrón multicapa puede representar funciones arbitrariamente complejas según la cantidad de neuronas con las que cuente, siendo, en cualquier caso, similares los cálculos necesarios.

Durante este trabajo se incorporaron al modelo de variograma distintas variables que podrían aportar información para predecirlo cuando se presentaran casos desconocidos. Ya que los modelos estadísticos de referencia se basan exclusivamente en la distancia entre los dos puntos para los que se quiere calcular el variograma, agregar información adicional podría mejorar la calidad de la predicción.

Se desarrollaron según lo comentado, cuatro redes neuronales, diferenciadas entre sí por el vector de entrada aceptado y la consiguiente diferencia en la complejidad de las capas escondidas. Ya que cada tipo de vector de entrada refleja distribuciones sobre espacios más grandes, y por ende más complejos, se variaron con estos la cantidad de neuronas en esta capa así como distintos parámetros del entrenamiento de la red.

Estos son entonces los vectores de características utilizados en las cuatro variantes:

1. Distancia
2. Distancia + $|\Delta$ Altura
3. Distancia + $|\Delta$ Altura + Altura media
4. Distancia + $|\Delta$ Altura + Altura media + Latitud media + Longitud media

El primer vector es equivalente al utilizado por los modelos estadísticos de referencia: Solo utiliza la distancia entre los puntos. El segundo vector incorpora la diferencia de altura entre ambos puntos (en valor absoluto), con la esperanza de que este dato permita un modelo más preciso. El tercer vector incorpora la altura promedio de los puntos. El cuarto vector incorpora la posición geográfica del promedio de los puntos del par, descompuesto por ejes. Esto debería ser complementario de la distancia entre ambos puntos y permitirle al MLP tener en cuenta la ubicación geográfica para estimar el variograma.

Se notará que todos los valores utilizados son simétricos en cuanto al orden de los puntos geográficos de cada par. Esto es porque el variograma en si es una relación simétrica y expresar los atributos de esta forma permite una descripción más compacta de la información.

Utilizar representaciones simétricas implica una reducción de $2^{\#Attr}$ en el tamaño del dominio que deberá procesar la red. Por ejemplo: Si no tomamos en cuenta la simetría del variograma, para un área con altura máxima de 3000mts necesitaríamos describir la diferencia de altura entre dos estaciones por un valor en el rango [-3000; 3000]. Teniendo en cuenta la simetría solo hace falta tomar en cuenta el modulo de este valor, con rango [0; 3000].

2.5.3 Optimización y evaluación de redes

Cada tipo de red de las cuatro mencionadas en el punto anterior fue puesta a prueba sobre los conjuntos de datos mensuales descriptos en la sección 2.3. En este punto se describe el proceso de optimización y evaluación de redes.

Un MLP es una red neuronal supervisada. Es decir que intenta aprender una función objetivo (no explícita) a través de la exposición de casos conocidos (*conjunto de entrenamiento*). Una vez aprendida la función se puede evaluar la eficacia de la red sobre un nuevo conjunto de datos, disjunto del anterior, llamado *de evaluación*.

Si la función fue bien aprendida, es de esperar que la red ya entrenada pueda estimar con precisión aceptable el variograma para pares nuevos de puntos geográficos. Dos problemas importantes durante el aprendizaje son:

- Falta de información, ya sea por cantidad de ejemplos presentados o por falta de atributos con los cuales la red pueda aproximar el atributo objetivo.
- El sobre-aprendizaje, por el cual la red se ajusta demasiado a detalles de los ejemplos que le fueron presentados perdiendo capacidad de generalización. Después de un sobre-aprendizaje, la red puede predecir a los ejemplos presentados con muy alta precisión pero no es capaz de hacer lo mismo con casos que no fueron expuestos durante el entrenamiento.

Es común ver que mientras se sobreentrena una red, su eficacia sobre el conjunto de evaluación empeora. Por esto es especialmente delicada la evaluación de eficiencia de una red: Hay que separar muy estrictamente los ejemplos presentados durante su entrenamiento de aquellos con los que se la evalúa, para poder contar con un conjunto de evaluación que sirva realmente de referencia.

Otro factor que complica el ciclo de desarrollo – evaluación – optimización es el azar: Los perceptrones multicapa suelen inicializar parámetros internos con valores elegidos al azar, lo que si bien es muy positivo para la eficiencia de la red, también hace que no sea tan sencillo evaluar si se desempeño bien o mal. Es decir: La eficiencia de una red deja de ser una función del conjunto de entrenamiento y su configuración, pasando a ser más bien una distribución.

Volviendo al caso concreto: Se cuenta con un conjunto de estaciones meteorológicas, y para cualquier par de estaciones se puede calcular variograma, por lo cual contamos con un número de ejemplos de entrenamiento y evaluación en el orden de n^2 (donde n es la cantidad de

estaciones), que en la práctica se reducen más de 50% por restricciones concretas de distancia entre las estaciones y otro 50% por usar relaciones simétricas.

Se decidió que para evaluar una arquitectura concreta, sobre un vector de características fijado, se desarrollaran:

- Distintas particiones del conjunto de datos en entrenamiento y evaluación (6 particiones por cada conjunto, compuestas por 1/6 de evaluación y 5/6 de entrenamiento).
- El conjunto de evaluación de cada partición está compuesto por estaciones exclusivas, no compartidas con el conjunto de entrenamiento, de forma que ninguna estación del conjunto haya sido presentada como parte de un par durante el entrenamiento.
- Cada partición será utilizada para entrenar 5 redes con la misma arquitectura, pero con inicialización independiente de sus estructuras internas. De esta forma se intentan capturar las variantes que pueden involucrar la inicialización aleatoria dentro de cada arquitectura.

Este doble nivel de diversificación de las evaluaciones hace que cada arquitectura sea representada por 30 redes neuronales, que en conjunto nos permiten evaluar la eficiencia para predecir el variograma, para un vector de características dado. Adicionalmente, al estar el conjunto de datos partido por mes y por variable, se debería evaluar los resultados para 24 conjuntos de datos distintos. Por suerte los resultados muestran que la eficiencia relativa de cada arquitectura no varía en general con el mes ni con la variable de temperatura utilizada.

En el marco de este informe se trabajó con 4 vectores de características dados, que fueron en particular los que mejor resultado mostraron para sus respectivos niveles de complejidad de dominio. No se pretende con esta selección ser concluyentes con respecto a qué información es útil para este modelo y que información no lo es. Lo que si se intenta es evaluar la utilidad de incorporar información adicional al modelo. Los resultados obtenidos serán presentados en la sección 2.6.

Para decidir que arquitectura representa mejor a un conjunto de datos, se evaluaron entonces distintas combinaciones de **# neuronas en la capa escondida**, y **# iteraciones del ciclo de aprendizaje**. Cada combinación de neuronas e iteraciones tiene asociado como vimos, un conjunto de 40 redes entrenadas sobre el dataset.

Algunos conceptos utilizados durante este trabajo de optimización:

- Una **instancia** es un par de estaciones, con su vector de datos asociado y su variograma real calculado.
- Un **conjunto simple de redes** es un grupo de 6 redes, tales que para cada estación **e** hay exactamente una red del conjunto que no la presenta en ninguna instancia de su conjunto de entrenamiento. Cada instancia está presente por lo tanto en los conjuntos de evaluación de exactamente dos redes dentro de un conjunto simple de redes (recordemos que cada estación de la instancia está asociado con exactamente 1 conjunto de evaluación).
- Un **conjunto compuesto de redes**, es la agrupación de 5 conjuntos simples de redes inicializados con distintas semillas de azar, y representa a la muestra con la cual se evalúan las arquitecturas.

- La **predicción consensuada de un conjunto para una instancia** es el promedio de valores predichos para esta instancia por las redes idóneas (aquellas que contaban con la instancia en su conjunto de evaluación). Lo interesante de este concepto es que la predicción consensuada suele ser mejor que todas las predicciones individuales que la componen.
- Finalmente, el **error de consenso** e sobre un conjunto de datos E se define como

$$e = \sqrt{\text{media}((y_i' - y_i)^2)}, i \in E$$

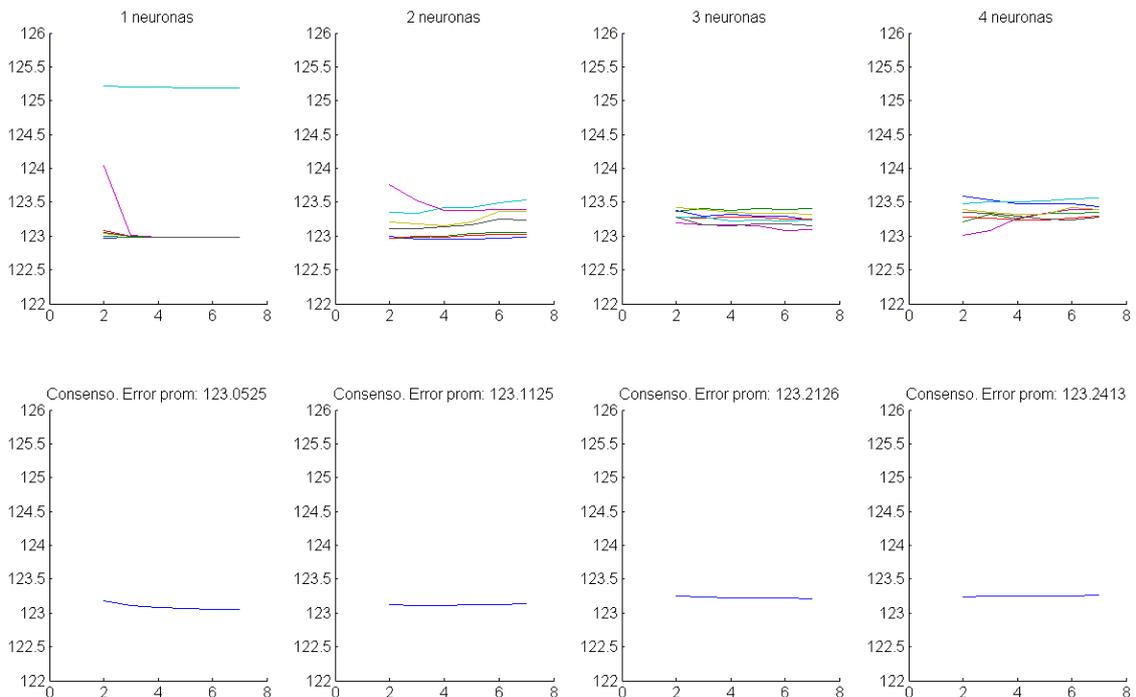
Donde y_i' es la predicción (de variograma) consensuada para la instancia i , e y_i es el valor real (de variograma) para esta instancia.

El error de consenso fue el principal indicador de eficiencia de una arquitectura sobre un conjunto de datos.

A modo de ejemplo, se presenta a continuación la tabla de resultados para el 1^{er} modelo: Aquel formado exclusivamente por la distancia entre los puntos. Esto es sobre el conjunto de datos correspondiente al mes de Julio, para temperatura máxima.

En la fila superior se ven los resultados de los conjuntos simples de redes, mientras que en la fila inferior se ve el resultado del conjunto compuesto de redes. Cada columna representa una arquitectura (cantidad de neuronas en la capa escondida).

Cada cuadro muestra el error de consenso según va progresando con la cantidad de iteraciones (x100) del ciclo de aprendizaje.



Se puede ver que agregar neuronas por sobre la 1^{era} no mejora el resultado final (incluso lo empeora), y que entrenar las redes más allá de los 600 ciclos tampoco mejora el resultado.

Viendo resultados similares en distintos meses, tanto para temperatura mínima como máxima, se decidió que para este modelo concreto se utilizaría una arquitectura de una neurona en la capa escondida y las redes se serian entrenadas por 600 ciclos.

2.6 Comparación entre modelos / Resultados

Se aplicó la metodología descrita en el punto anterior para elegir una arquitectura óptima para cada vector de datos utilizados. Entre los distintos vectores evaluados, se seleccionaron 4 que sobresalían por su relación complejidad-capacidad de estimación. Durante esta sección presentaremos y analizaremos los resultados de los dos modelos estadísticos descriptos al principio de esta sección y los cuatro modelos referidos, basados en redes MLP.

El análisis a realizar en esta sección lo vamos a referir al conjunto de datos del mes de Julio para temperatura máxima. Si bien hay cambios significativos entre los distintos conjuntos de datos, los resultados de los distintos modelos mantienen siempre sus particularidades y sus diferencias relativas.

Dados los 4 modelos MLP

1. Distancia
2. Distancia + $|\Delta$ Altura
3. Distancia + $|\Delta$ Altura + Altura media
4. Distancia + $|\Delta$ Altura + Altura media + Latitud media + Longitud media

A continuación, las configuraciones y resultados:

Configuración y resultados para los 4 modelos MLP				
Modelo	M#1	M#2	M#3	M#4
# atributos	1	2	3	5
# capa escondida	1	2	5	9
Iteraciones	400	800	1800	8100
Error promedio	123.6	114	109	100
Error de consenso	123.5	113.5	107.8	95
STDV error	0.08	0.2	1	2
T° de training	<1min	2min	8min	45min

El error promedio es el promedio de errores de consenso de cada conjunto simple de redes, mientras que error de consenso corresponde al conjunto compuesto.

A modo de comparación, los resultados para los modelos estadísticos son los siguientes:

Esférico	123.6
Exponencial	123.6

De los modelos 3 sobre distancia (estadísticos mas modelo MLP #1) se puede concluir que no hay diferencias importantes de eficiencia, ya que los 3 cometen error en rango 123.5 – 123.6.

Esto valida de alguna forma el modelo MLP, ya que muestra como equipara a los modelos tradicionales, que son el estándar usado en la literatura.

Las dos familias de modelos tuvieron una instancia de ajuste de parámetros: Optimización no lineal (por mínimos cuadrados) en el caso de los modelos estadísticos y entrenamiento en el caso del MLP, y si bien los modelos estadísticos se ajustan sobre los mismos valores que pretenden representar (lo que podríamos llamar ventaja competitiva por sobre las redes), el MLP más sencillo tiene un poder mucho mayor de representación de una función que los modelos estadísticos.

En la figura 15 a continuación, se puede ver la predicción de variograma de cada modelo en función de la distancia entre los puntos del par. Si bien los 3 modelos tuvieron resultados numéricamente similares, se puede ver una diferencia de complejidad entre las funciones representadas por el modelo MLP y los otros dos.

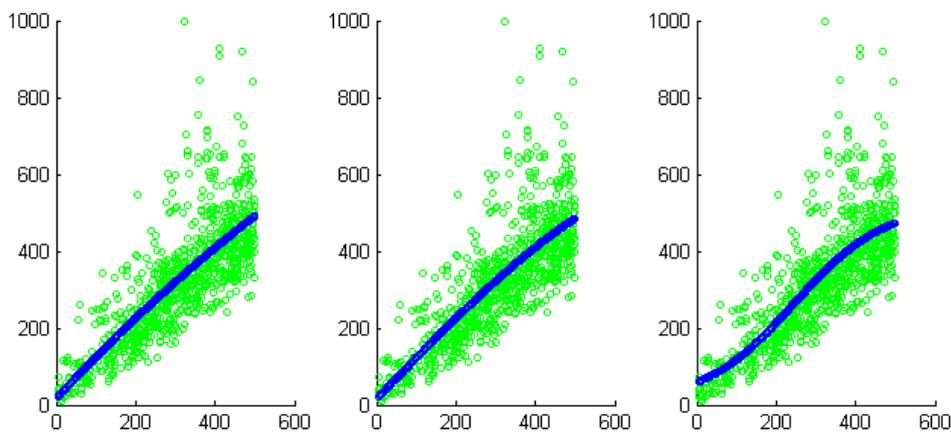


Figura 14 - Comparación de los 3 modelos sobre distancia. Modelos Exponencial, Esferico y MLP en orden de izquierda a derecha.

Pasando a la comparación entre modelos MLP, se puede observar cómo la complejidad del vector de características influye sobre la complejidad de la arquitectura que mejor lo representa. El modelo 1 se puede presentar naturalmente por una función de 1 parámetro, monótona creciente. Internamente esto comprende una red de una neurona de entrada, una neurona escondida y tiene asociado un entrenamiento de 400 ciclos.

Al agregar información sobre la diferencia de altura del par (modelo #2), y luego sobre promedio de altura del par (modelo #3), en cada paso se duplica la cantidad de neuronas y se duplica la duración del entrenamiento (en # ciclos).

El modelo 4 incorpora información sobre la posición de cada par. Esta información es más compleja ya que no se puede generalizar ni reutilizar fuera del contexto actual. Se espera que la red pueda acomodar esta información con éxito hasta cierto punto, a cambio de un aumento importante de complejidad. En la práctica vemos que este modelo duplica la cantidad de neuronas del paso anterior y requiere un entrenamiento 4,5 veces más largo (en # ciclos).

En cuanto a la eficiencia, se ven diferencias de 30% entre el mejor y peor estimador MLP. Más aún, el uso de MLP y la incorporación de información al modelo ofrece un 30% de mejora sobre los modelos estadísticos.

2.6.1 Análisis de las diferencias entre modelos MLP

A continuación se presentan detalles de los resultados de los modelos MLP, con el fin de entender el impacto de cada variante en el vector de características sobre la representación que la red resultante genera sobre el dominio.

Las figuras a continuación muestran la función de variograma según los 4 modelos generados aplicada a 5 puntos del territorio argentino y sus respectivas vecindades. En el centro de cada cuadrado se encuentra el punto analizado, y a su alrededor, los contornos de la función de variograma con cada punto de la vecindad, en colores desde azul (mínimo variograma, correlación perfecta) a rojo (máximo variograma, baja correlación)

Esto permite ver para cada punto, como los diferentes modelos generan distintas funciones, así como analizar las razones en cada caso.

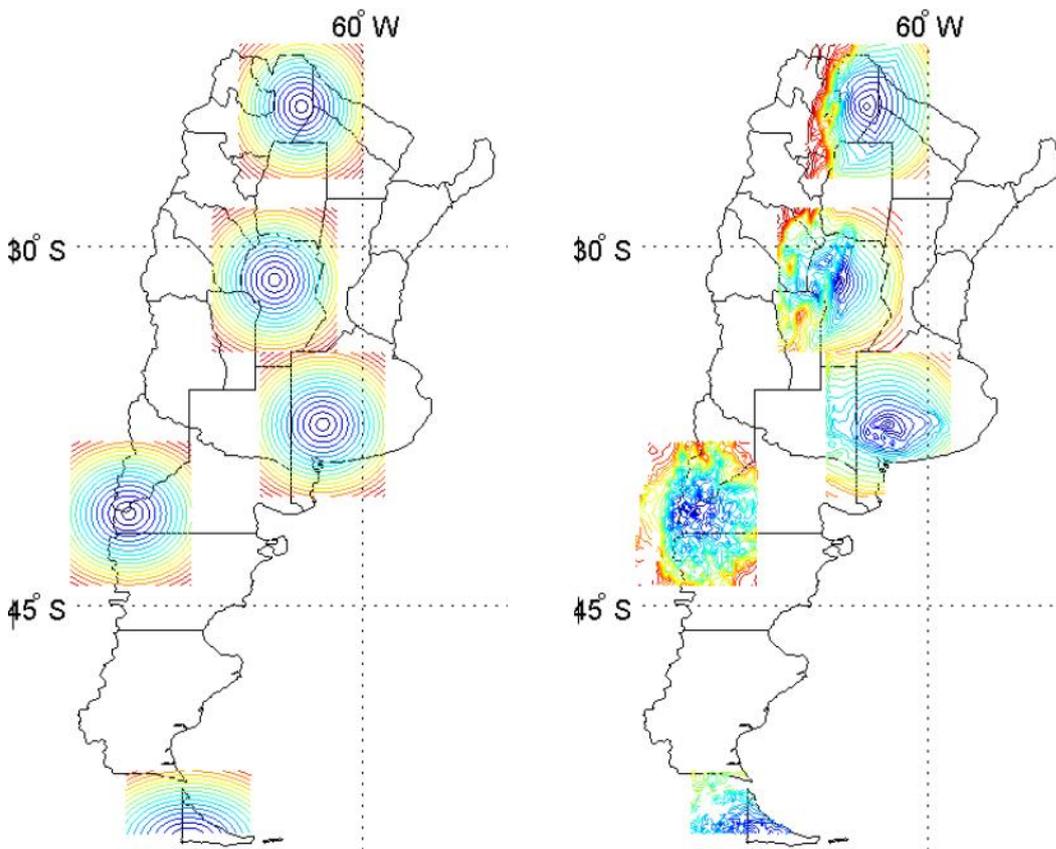


Figura 15 - Variograma para 5 puntos del territorio argentino, según modelos MLP #1 y #2

En el 1er mapa se puede ver claramente que el variograma depende exclusivamente de la distancia. Cada punto presenta contornos circulares perfectos que decaen con la distancia.

En el 2do mapa se ven contornos más irregulares, que castigan fuertemente las diferencias significativas de altura. Esto se puede ver claramente en los 3 puntos de la izquierda, donde quedan demarcadas claramente en rojo las diferencias de altura de la cordillera. Se puede ver a su vez que en los puntos ubicados en Buenos Aires y Tierra del Fuego (sin mayores saltos de altura), los valores de variograma decaen menos con la distancia que en el primer modelo.

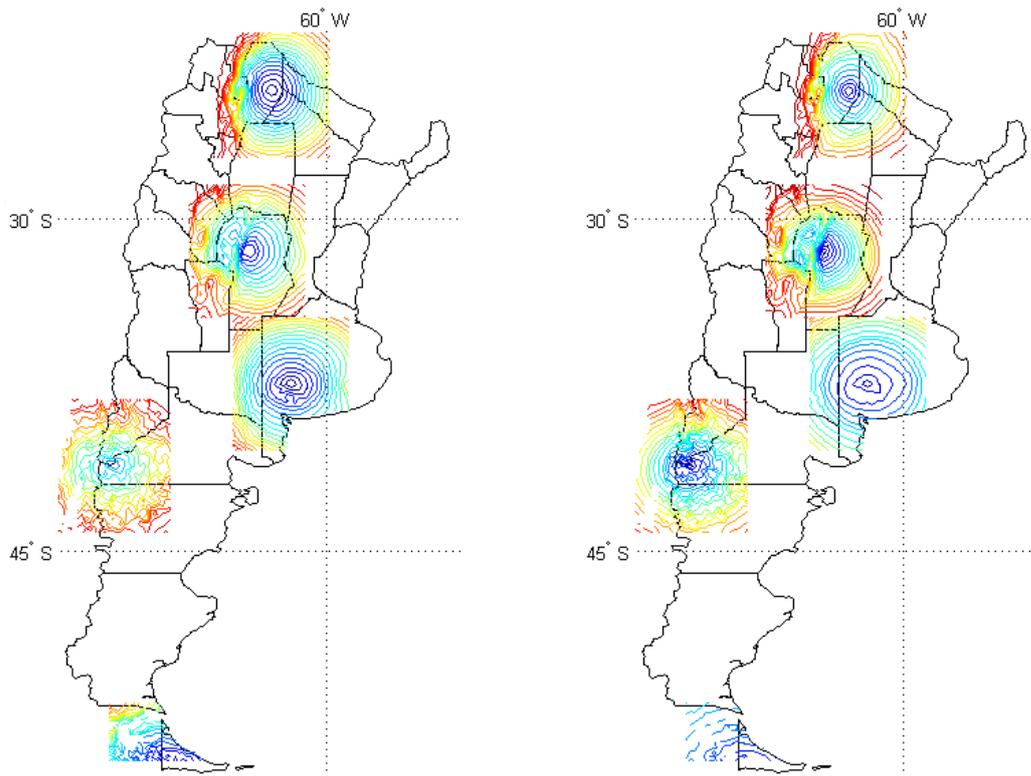


Figura 16 - Variograma para 5 puntos del territorio argentino, según modelos MLP #3 y #4

En el mapa correspondiente al 3er modelo, se puede observar cómo se suaviza la penalización de la diferencia de altura. Un buen ejemplo de esto es los puntos ubicados en Córdoba y Buenos Aires, donde los contornos vuelven a tomar forma circular. Otro cambio importante se ve en el punto de Río Negro, en zona de cordillera, que pasa a tener valores de variograma muy menores con su vecindad.

En el vector 4 no se observan cambios tan importantes, aunque sí se ven aumentos en el variograma del punto de Buenos Aires con su vecindad, y lo mismo para aquel de Tierra del Fuego y Neuquén.

La figura a continuación presenta para cada uno de los 4 modelos, una evaluación del error cometido por cada uno en función de 4 variables.

En los gráficos superiores en cada caso vemos el error cometido por el modelo en función de la distancia entre los puntos del par. En rojo el plot de muestras del conjunto de test y en negro el promedio de éstos (suavizado). En el 2do gráfico (en verde) se puede ver el error cometido en función del promedio de altura del par, también con el promedio en negro. En el tercer gráfico (azul) se ve el error cometido en función de la diferencia de altura del par, y en el cuarto se ve lo propio con respecto a la latitud promedio del par.

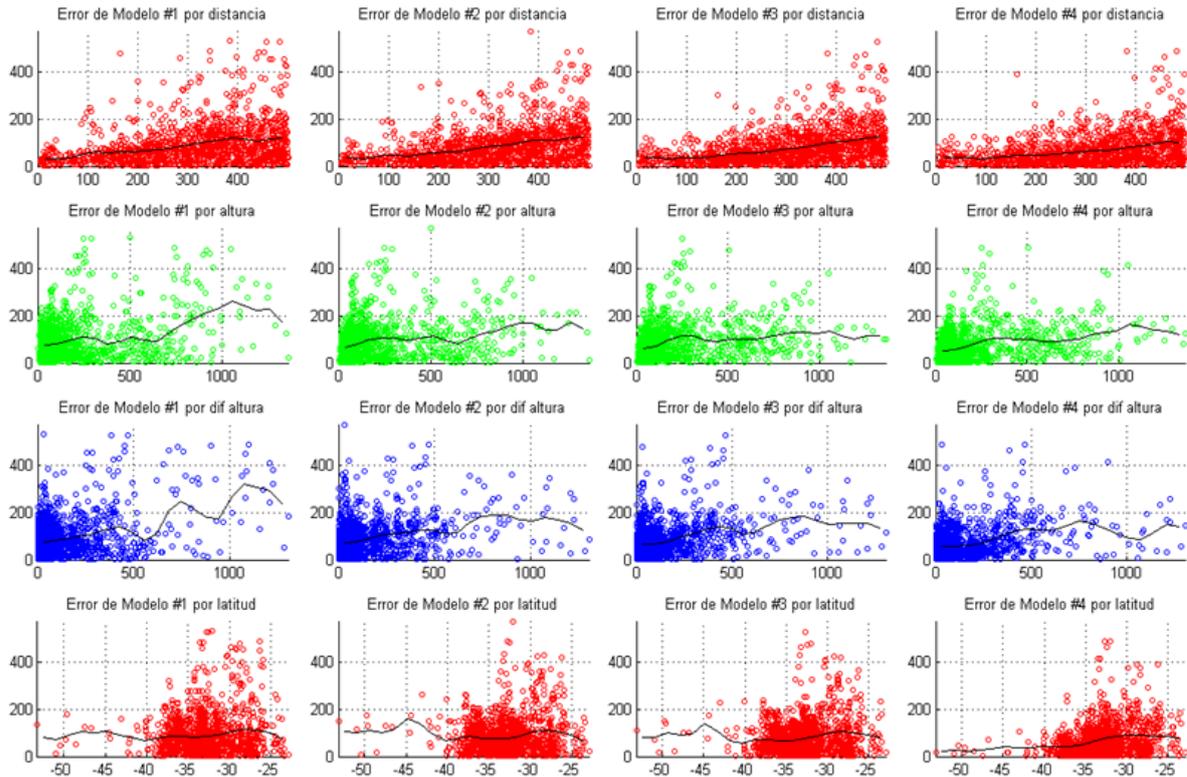


Figura 17 – Evaluación del error cometido en función de las variables

Algunas observaciones en cuanto a las diferencias de los modelos:

- El error cometido parece aumentar en los cuatro casos con la distancia. Esto se debe a que a mayor distancia mayor variabilidad en el variograma. De todas formas, el cuarto modelo presenta una baja notoria en el rango de 50 a 300 kilómetros.
- El modelo #2 mejora notablemente los resultados a mayor altura promedio y a mayor diferencia entre ambas. Estos dos cambios se acentúan en el modelo #3. El objetivo principal de los modelos #2 y #3 es justamente mejorar la predicción del variograma ante escenarios de alturas, y de diferencias de altura significativa.
- El modelo #2 modelo parece perder precisión en el sur (latitudes bajas) con respecto al primero, aunque esto se ve revertido totalmente en el 4to modelo.

2.7 Conclusiones. Hacia la interpolación...

Se presentó una aplicación de redes MLP al problema de modelar la función de variograma entre series de temperatura, para una zona geográfica concreta, y se compararon los modelos resultantes con aquellos utilizados frecuentemente en la literatura.

El principal problema de los modelos estadísticos de variograma es la pobreza de la información que utilizan.

La pregunta abierta que había antes de realizar este trabajo era: ¿Será positivo incorporar información adicional al modelo, en cuanto a la precisión de la estimación resultante? A partir de esta pregunta se desarrollaron 4 modelos sobre redes neuronales MLP, donde el primero se basaba simplemente en la distancia entre los puntos del par y los restantes 3 incorporaron progresivamente información adicional, incluyendo finalmente información de localización.

Pudimos corroborar que el 1er modelo MLP tenía resultados equivalentes al esférico y al exponencial, con lo que aceptamos que las redes neuronales pueden generar automáticamente funciones tan eficientes como las utilizadas en la literatura.

Finalmente, al incorporar más información al modelo, vimos que el error de estimación se redujo en un 25%, siendo especialmente notorias las diferencias en casos extremos de la distribución. Entre estos:

- Disminución promedio del error de estimación de 45% entre puntos con promedio de latitud menor a 42°S
- Disminución promedio del error de estimación de 40% entre dos puntos con diferencias de altura mayor a 900mts.
- Disminución promedio del error de estimación de 40% entre dos puntos con promedio de altura de entre 800 a 1200mts.

Estos casos particulares muestran dos cosas: La incorporación de información adicional al modelo y en particular a través del uso de redes neuronales hacen a un mejor modelo, uno más robusto para casos extremos, adaptable por naturaleza a escenarios muy distintos.

Con respecto a la complejidad computacional introducida por las redes neuronales, se puede ver que el tiempo de entrenamiento de estas aumenta 4500% al incorporar la información mencionada.

Sin embargo sabemos que en cualquier caso, el entrenamiento se realiza una vez para cada modelo, consumiendo hasta 45 minutos en los equipos con los cuales se realizó el trabajo. Luego de este entrenamiento inicial, cualquier red evalúa nuevos casos en pocos milisegundos. También sabemos que una red se representa en una computadora, en una cantidad de espacio acotada, relacionada a la complejidad de su arquitectura, pero siempre en el orden de pocos kilobytes. Esto hace que los modelos recién presentados o incluso algunos más complejos, puedan ser utilizados sin demandas especialmente grandes de poder de cómputo.

El trabajo recién presentado no pretende ser concluyente en cuanto a la información necesaria y suficiente para generar un modelo de variograma. En cambio, se llegó a la conclusión de que las redes neuronales de tipo MLP son una buena herramienta para generarlos, ya que permiten incorporar información adicional a un cálculo de forma transparente. También se vio que existe información fácilmente obtenible, relativa mayormente a la posición (latitud, longitud y altura), que permite mejorar de forma notoria los modelos actuales de variograma.

Este trabajo fue utilizado dentro del proyecto CLARIS para interpolación de temperatura de estaciones meteorológicas (Pausader, 2008), con el fin de generar grillas regulares de temperatura.

A continuación, como parte de este mismo proyecto, se estudia el uso de esta información generada como parte del trabajo en proyección de modelos de predicción de clima respecto a datos verídicos.

2.8 Bibliografía

- [1] Pausader, J. P. Boulanger. 2008. Etude du changement climatique en Argentine et projections pour le 21^o siècle. M.
- [2] Cressie, N., 1993, Statistics for spatial data, Wiley Interscience
- [3] Chiles, J.P., P. Delfiner, 1999, Geostatistics, Modelling Spatial Uncertainty, Wiley-Interscience
- [4] Haylock, P.D. et al, 2008. A European daily high-resolution gridded dataset of surface temperature and precipitation for 1950-2006

3 Proyección del los modelos de predicción del clima respecto a datos verídicos

3.1 Introducción

En esta sección se explicarán qué son los modelos de predicción del clima, de dónde provienen y se muestra cómo tomando en cuenta las predicciones a futuro que tienen, es posible clasificarlos y caracterizarlos en clases que representen a modelos con las mismas tendencias en distintos sectores.

Luego se hará un enfoque en Sudamérica y Argentina y se tomarán aquellos modelos que resulten relevantes para la región de acuerdo a como resulte la clasificación. Se comparará el comportamiento que predicen entre los años 1961-1990 contra los datos observados según el clima ocurrido en esas fechas.

Por último se muestra como la clasificación mantiene una coherencia a nivel mundial, habilitando el posible estudio de otras regiones en distintas partes del mundo.

3.1.1 Grupo Intergubernamental de Expertos sobre el Cambio Climático

Los modelos de comportamiento climático estudiados corresponden a desarrollos de distintos países basándose en los escenarios especificados por el IPCC.

El **Panel intergubernamental sobre el Cambio Climático (IPCC)** es un cuerpo científico encargado de evaluar el riesgo de cambio climático causado por la actividad humana. Este panel fue creado en 1988 por la Organización Meteorológica Mundial (OMM) y el Programa Ambiental de las Naciones Unidas.

3.1.2 Reporte Especial sobre Escenarios de Emisiones

Este reporte (conocido mejor por nombre en inglés **Special Report on Emissions Scenarios**), fue preparado por el IPCC para el Third Assessment Report (TAR) realizado en el 2001, sobre aquellos escenarios que serán usados para guiar a los modelos globales para que desarrollen escenarios de cambio climático. Estos escenarios reemplazaron a los IS92 utilizados para el Second Assessment Report del año 1995. También se utilizaron para el Fourth Assessment Report (AR4) del 2007.

Como proyectar el cambio climático depende fuertemente de la actividad humana a largo plazo, los modelos climáticos, de los cuales hablaremos más adelante, se ejecutan contra escenarios establecidos. Existen 40 escenarios distintos, donde cada uno realiza diferentes asunciones sobre la polución en el futuro, efectos de gas invernadero, uso de la tierra y otras fuerzas relevantes. Esto significa que cada escenario toma en cuenta diversos niveles de desarrollo tecnológico y desarrollo económico de la humanidad.

Los escenarios de emisiones se organizan en familias, las cuales contienen aquellos que poseen características similares entre sí. Los reportes de proyección del IPCC generalmente se hacen en el contexto de una familia de escenarios en particular.

3.1.3 Familias de Escenarios

Estas familias contienen escenarios con características comunes. Las familias mencionadas en el Third Assessment Report (TAR) y en el Fourth Assessment Report (AR4) del IPCC son A1FI, A1B, A1T, A2, B1, y B2. En el apéndice 6 (Familias de escenarios del IPCC) se comenta una breve descripción de las particularidades de cada una.

3.2 Modelos

Los modelos climáticos son las herramientas principales para proyectar el clima a futuro (Houghton et al). Han sido evaluados en base a sus habilidades de simular una gran variedad de situaciones y fenómenos. Este informe se centra en modelos que simulen el escenario A2, por ser el mismo el más pesimista de todos. Para la proyección, se utilizan los siguientes 17 modelos. En el apéndice se describen los orígenes de cada uno:

bccr_bcm2_0	ccma_cgcm3_1	cnrm_cm3
csiro_mk3_0	gfdl_cm2_0	gfdl_cm2_1
giss_model_e_r	inmcm3_0	ipsl_cm4
miroc3_2	miub_echo_g	mpi_echam_5
mri_cgcm2_3_2a	ncar_ccsm3	ncar_pcm1
ukmo_hadcm3	ukmo_hadgem1	

De cada modelo se utiliza una sola simulación (corrida), y además de simular al escenario A2, se utilizaron los resultados correspondientes al escenario 20c3m, del siglo 20, para luego compararlo contra los valores observados.

3.2.1 Datos utilizados (temperatura y precipitación)

Cada modelo provee muestras mensuales de los valores en los siglos 20 y 21, con distintas resoluciones espaciales. Para una buena comparación, se interpolaron los valores para concordar a una grilla espacial de 2.5 x 2.5 grados, de la misma manera que se hizo con los valores observados y tal como se menciona en (Boulangier et al.). Esto resulta en una matriz de 72x144 para cada punto del planeta: 72 latitudes y 144 longitudes.

En la figura 19 se muestra el mapa de Sudamérica, destacando la granularidad espacial de dicha grilla.

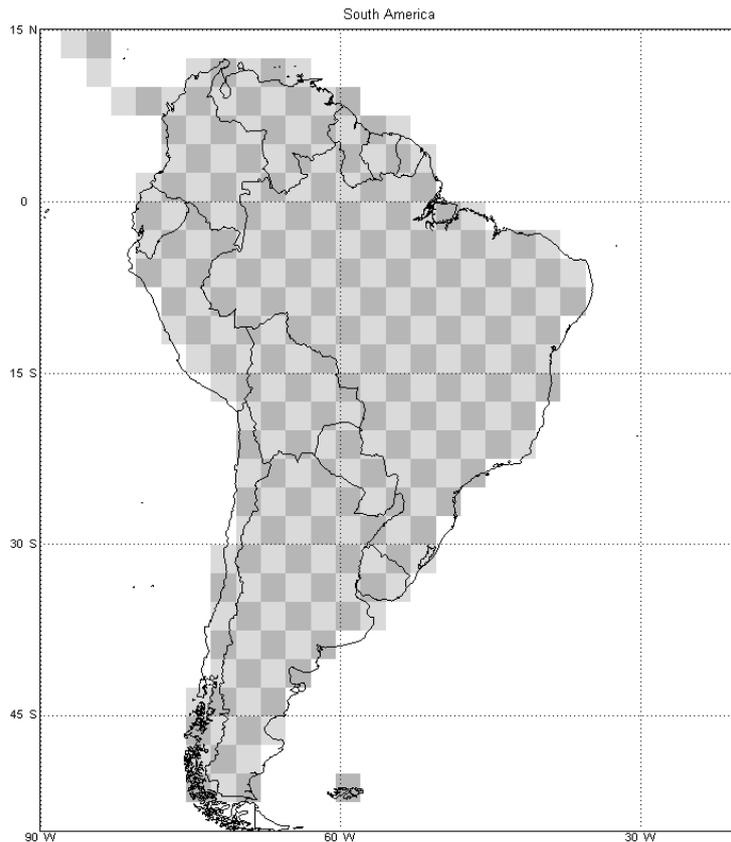


Figura 19 - Resolución espacial tomando una grilla de 2.5 grados

Cada modelo provee una variedad de datos en cada punto de grilla, de los cuales para este informe se utilizan la temperatura ambiente y la precipitación. Estos datos son promedios mensuales, para cada año del siglo 20 (20c3m) y del siglo 21 (escenario A2). La temperatura será utilizada para el proceso de clasificación de modelos. Luego, tanto ésta como la variable precipitación serán las medidas para verificar el rendimiento de los modelos en base a comparaciones. Los valores son promedios mensuales, para cada año (1901-2000) y (2001-2100).

3.3 Análisis de los datos de los modelos

3.3.1 Determinación temporal

Se establece como base el período 1961-1990, el cual se conoce como período de referencia del IPCC, y a partir de entonces se observa para cada modelo como va cambiando su temperatura a medida que pasa el tiempo. Para la obtención de un valor global promedio, se le aplicó un proceso de normalización a cada punto de grilla del globo.

$$\partial x = \cos \varphi \partial \theta \qquad \partial y = \partial \varphi$$

$$\iint \partial x \partial y = \int_{\theta=0}^{2\pi} \int_{\varphi=-\pi/2}^{\pi/2} \cos \varphi \partial \varphi \partial \theta = 4\pi$$

Donde θ sigue el camino de un paralelo (horizontalmente si se ve como un planisferio) y φ el de un meridiano. Computacionalmente hablando, la fórmula resulta en una matriz de pesos de 72x144 en la que todos los valores de una fila son iguales (correspondientes a un paralelo) y para cada columna, su i -ésimo valor (i toma entre 1 y 72) es:

$$\frac{\pi \sin \frac{i\pi}{72}}{144}$$

Siendo $\sin (i * \pi / 72)$ un cambio de coordenadas de los valores de la integral al rango 1..72; donde 1 y 72 son valores polares (resultando en un peso casi nulo) y 36 es donde la longitud se cruza con la latitud ecuatorial (resultando en un peso completo, o sea 1). Luego la sumatoria de 1 a 72 del seno da $1/\pi$, entonces se multiplica por π para que la sumatoria de 1 y se lo divide por 144 pues esa sumatoria se realiza para cada meridiano (y hay 144 meridianos en la grilla). Con esto se logra que todos los puntos de la grilla sumen 1 (pues es normalización).

Como se desean estudiar las tendencias anuales, sin resaltar a ningún mes en particular, se promediaron los 12 valores mensuales para obtener un valor único anual.

En la figura 20 se muestran los valores promedio para cada modelo. Se aplicó un suavizado de 25 años para evitar ruido. Se puede observar como todos los modelos predicen un marcado aumento de temperatura en el escenario A2.

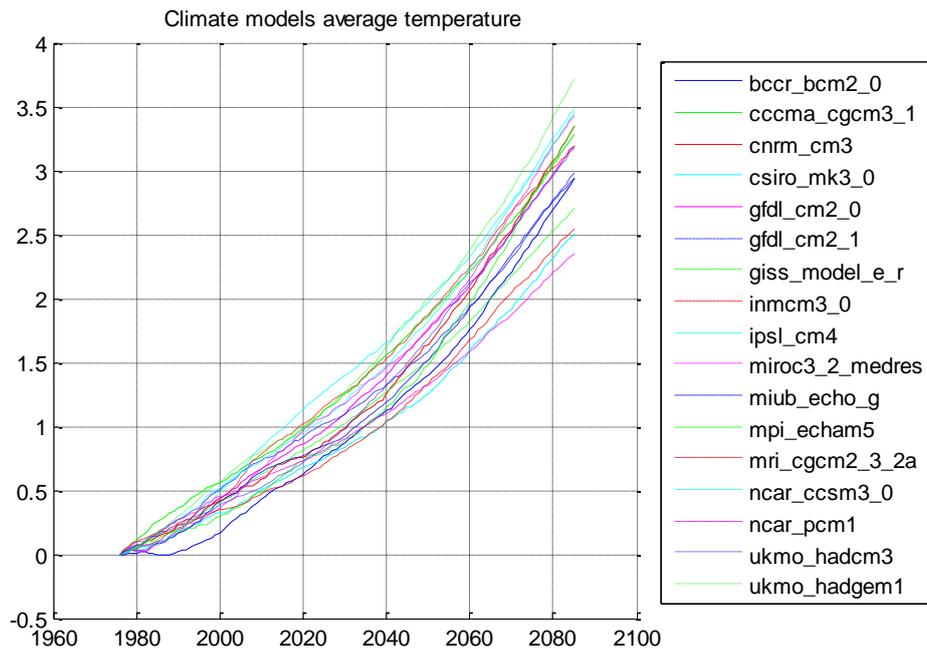


Figura 20 - Valores de temperatura anuales de cada modelo

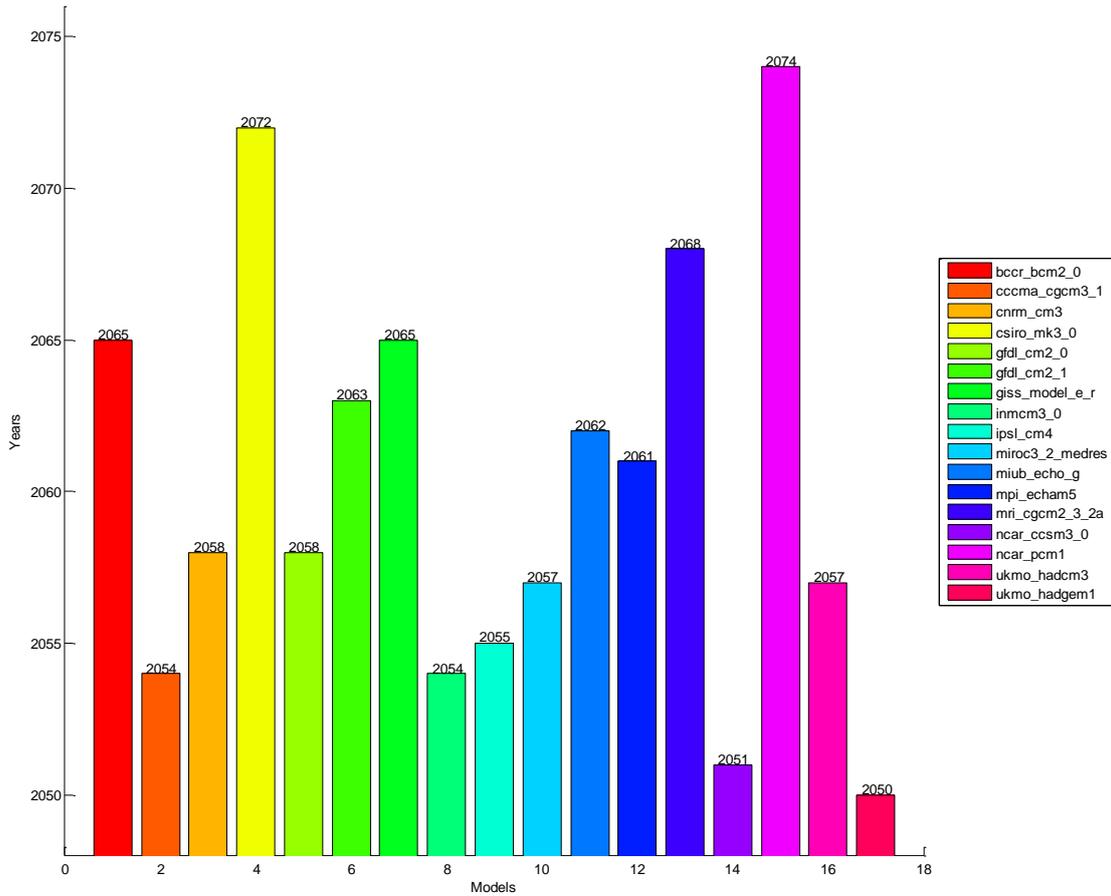


Figura 21 - Año en que cada modelo predice un aumento de 2 grados sobre el período 1961-1990

Este informe se enfoca en el momento en el que los modelos predicen un aumento promedio de 2 grados en la temperatura ambiental, pues este valor es el límite al cual el IPCC cree que no se debe sobrepasar. Este aumento se toma sobre el año 1976, año promedio del período 1961-1990.

Para obtener un año sobre el cual trabajar, se promediaron los años de aumento a 2 grados de cada modelo (figura 21), dando como resultado el año 2060. Sin embargo esa decisión trae aparejada una incertidumbre global entre los distintos modelos que dificulta el proceso de análisis. Como se desea estudiar el comportamiento local, se propuso tomar para cada modelo el momento en el que el mismo llega a sus 2 grados de aumento en promedio. En la figura 22 se toman los valores de aumento a 2 grados particulares y en la figura 23 los valores de cada modelo en el año 2060. En ambos gráficos se observa para cada caso cuál es el valor promedio de temperatura para cada latitud del globo.

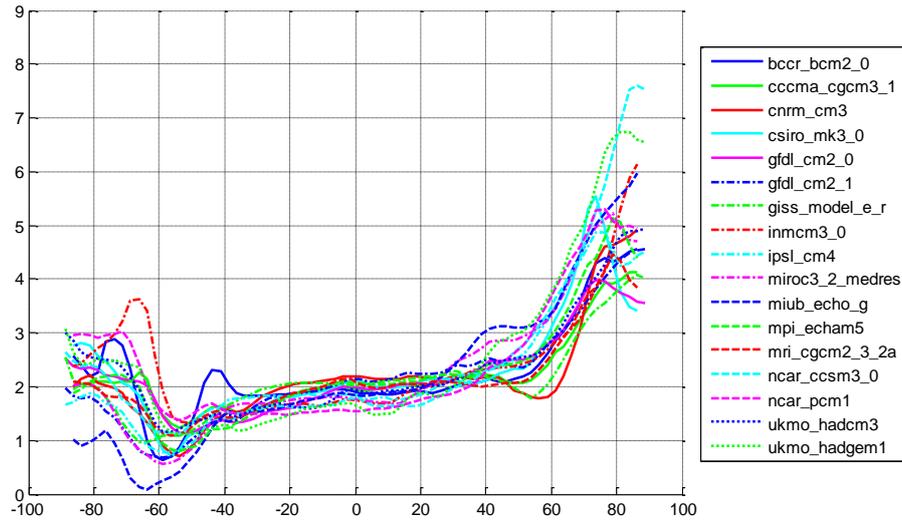


Figura 22 - Valores para cada latitud con datos de entrada de modelos en el año en el que cada uno predice un aumento a 2 grados

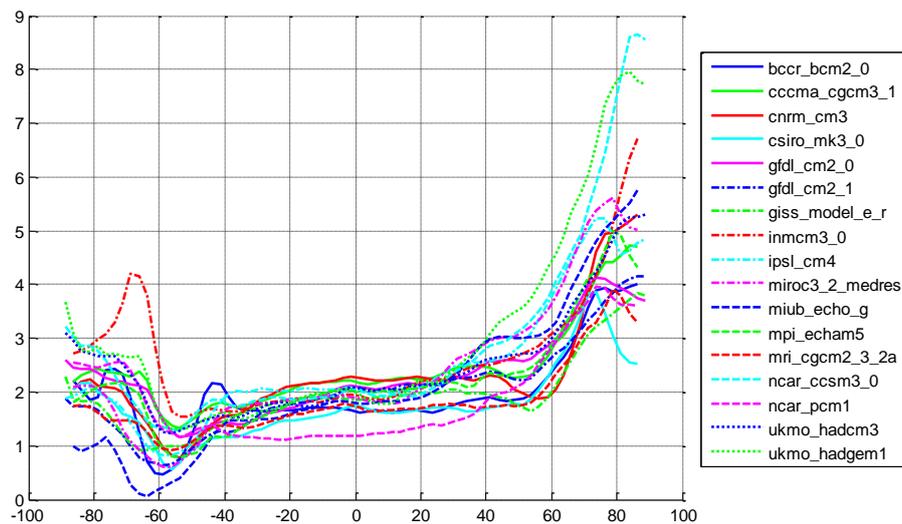


Figura 23 - Valores para cada latitud con datos de entrada de modelos en el año 2060

De estos gráficos se puede concluir que los valores representados en el primero muestran una menor incertidumbre respecto al aumento de 2 grados, con la consecuente mayor varianza en las zonas polares, debido al poco peso que tienen los datos allí. En el apéndice se ofrece un mayor detalle de la temperatura estimada en cada corte, para quien desee realizar un análisis en profundidad.

3.3.2 Pre-procesamiento de los resultados

Para cada punto de la grilla, se ordenaron los valores delta (diferencia con la temperatura promedio del período de referencia) de cada modelo y se tomó como dato su orden en lugar del tal valor, pues nuestro objetivo es de observar las zonas donde el orden de los modelos es similar.

3.4 Obtención de clases a partir de los modelos

A continuación, tomando como entrada los valores de temperatura de los modelos cuando éstos llegan a 2 grados, se realizará una clasificación para obtener un número de clusters que representarán a distintos climas de Sudamérica.

3.4.1 Aplicación de redes neuronales

Para el proceso de clasificación se empezará reduciendo el espacio con redes neuronales. En particular, a través de un Mapa Auto-organizado (SOM), el cual es un tipo de red neuronal artificial con aprendizaje no supervisado. Fue creado por Teuvo Kohonen por lo que a veces es llamado Mapa de Kohonen. Para más detalles del mismo, ver (Kohonen T.). Siguiendo el enfoque del trabajo, se toman en cuenta aquellos puntos de grilla que caen en Sudamérica. Estos conforman 264 posiciones, y siendo cada una un vector con 17 valores, uno para cada modelo, resulta que el espacio de entrada es de tamaño 264 y de dimensión 17.

Para el proyecto se utilizó la implementación de SOM conocida como SOM Toolbox, un paquete del programa Matlab. En (Vesanto J. et al., 2000) se explica cómo se encuentra hecho tal paquete y la manera de utilizarlo.

Al ejecutar el programa SOM, se dejó que el mismo realice un auto-afinamiento de los parámetros para que calcule el tamaño apropiado del mapa de grilla del SOM. La manera en la que calcula los valores por omisión es la siguiente:

Primero, se determina el número de unidades del mapa con la fórmula obtenida de la literatura (Vesanto J. et al., 2000):

$$um = 5 * tm^{0.54321}$$

donde um es el número de unidades del mapa y tm es el tamaño de la muestra.

Luego, los dos autovalores más grandes de los datos de entrada se calculan y se establece el radio entre el largo de los lados de la grilla. Entonces los valores de los lados son tales que su producto es tan cercano como sea posible al número deseado de unidades.

3.4.2 Clasificación de los resultados obtenidos

Como resultado del SOM, se obtuvo una grilla de 10x8 en la cual cada punto es un vector de 17 elementos. Es importante notar que el proceso del SOM no llevó más de 1 un segundo para la primera fase y 2 segundos para la posterior, lo cual es factible tomando en cuenta el tamaño de la matriz de entrada, las optimizaciones hechas por el SOM Toolbox y la velocidad de procesamiento existente hoy en día.

Error de cuantización (Quantization Error): Distancia promedio entre cada vector y su BMU (*Best Matching Unit, referirse a la bibliografía de SOM para más detalles*). Este error mide la resolución del mapa. Para la aplicación de SOM hecha en este proyecto, el QE resultante es 8.2912.

Error topográfico (Topographic Error): La proporción de los vectores de datos para los cuales las primera y segunda BMU no son unidades adyacentes. Mide la preservación topológica. En este proyecto el TE resultante es 0.0341.

La U-matrix (Ultsch A. et al.) es un método de visualización para poder obtener información de los resultados de las redes neuronales. Para este proyecto se observa en la figura 24.

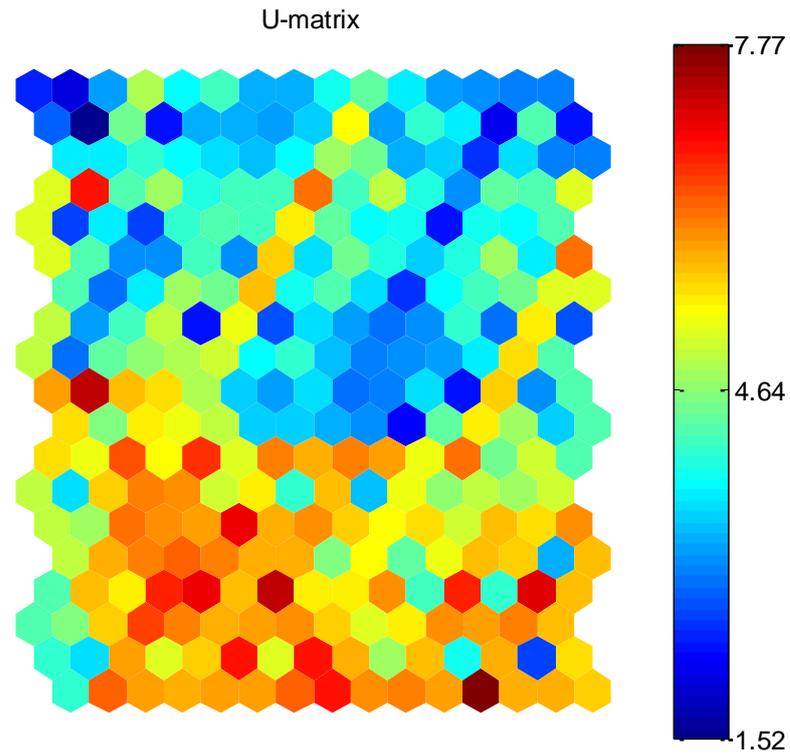


Figura 24 - U-matrix resultante de la aplicación del Mapa Auto-organizado a la matriz de entrada

Cada celda de la U-Matrix y de las component-planes mostradas a continuación, corresponde a un BMU. O sea es un vector de 17 elementos que mejor se ajusta a los valores de entrada. Cada valor de entrada representa a un punto de grilla del mapa de Sudamérica. Pero por simplicidad en el gráfico no se muestra para qué puntos del mapa se ajusta mejor cada celda.

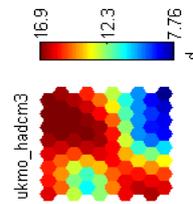
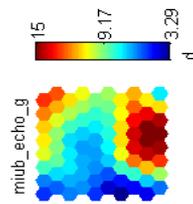
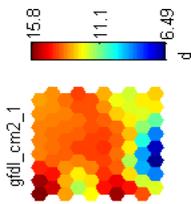
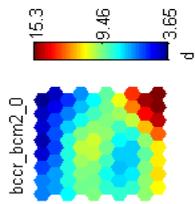
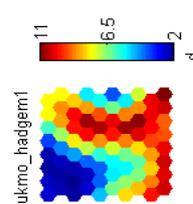
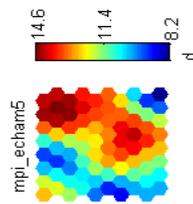
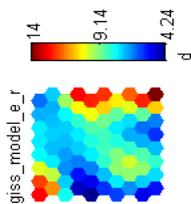
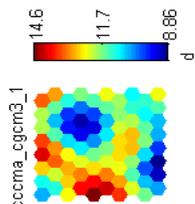
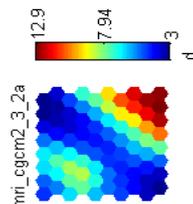
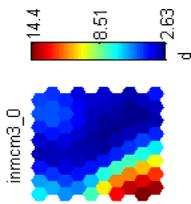
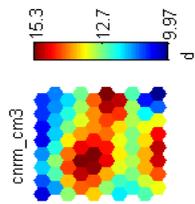
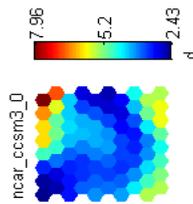
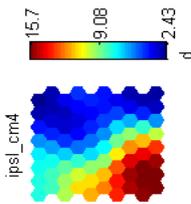
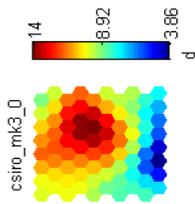
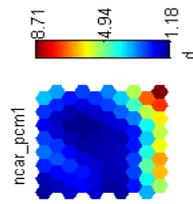
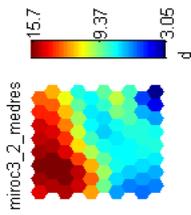
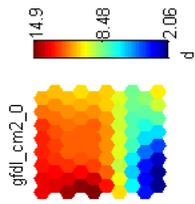


Figura 25 - Visualización de component-planes para cada modelo

Puntos de bajo valor en la U-matrix (tirando a azules) corresponden a grillas del mapa que forman un mismo cluster, y puntos de alto valor (naranjas y rojos) indican separaciones entre clusters. En la figura 25 se muestran los component-planes para cada variable (que corresponde a un modelo climático); éstos ayudan a descubrir la distribución espacial de cada modelo.

Como es difícil poder obtener clusters a partir de la U-matrix, se procede a realizar un Análisis Jerárquico de clústeres a los resultados del SOM, de manera de obtener un mejor detalle de los resultados arrojados por el mismo. Una explicación de esta técnica donde se combina SOM y clustering se puede leer en (Vesanto J. et al., IEEE).

3.4.2.1 Análisis Jerárquico de Clústeres

Este análisis es una generalización de los análisis de clústeres, en la cual el objetivo es agrupar conjuntamente objetos o registros que se encuentran "cerca" unos de otros. Una clave del análisis es el cálculo reiterado de distancias entre objetos, y entre clústeres una vez que los objetos comienzan a ser agrupados. La salida generalmente es representada con un *dendograma*.

Las dos categorías principales de métodos para el análisis son métodos divisivos y métodos aglomerantes. En cada paso, el par de clústeres con menor distancia entre sí se fusiona en uno solo. Los algoritmos más comunes para el clustering jerárquico son:

- Enlace simple
- Enlace completo
- Enlace promedio
- Enlace promedio grupal
- Enlace de Ward

3.4.2.2 Enlace de Ward (Ward's linkage)

Esta función de especificación de la distancia entre dos clústeres es computada como el incremento en la suma de cuadrados del error (ESS) luego de fusionar dos clústeres en uno solo. El método de Ward busca elegir pasos sucesivos de fusión tales que minimicen el incremento de ESS luego de cada paso.

El ESS de un conjunto X de N_x valores es la suma de los cuadrados de las desviaciones de la media. O sea:

$$ESS(X) = \sum_{i=1}^{N_x} \left| x_i - \frac{1}{N_x} \sum_{j=1}^{N_x} x_j \right|^2$$

Matemáticamente, la función de enlace (la distancia entre los clústeres X e Y) se describe con la expresión:

$$D(x,y) = ESS(xy) - [ESS(x) + ESS(y)]$$

donde

- XY es el clúster combinado resultante de la fusión entre X e Y
- ESS(...) es la función descrita anteriormente

3.4.2.3 Resultados de la clasificación

Se observa que la clasificación jerárquica, aplicada a la grilla de 80x17 puntos obtenida del SOM (siendo cada fila una celda y cada columna un modelo), arroja buenos resultados. Cada clase es distinguible y representa a fenómenos climáticos de la región.

Si bien el criterio de Ward nos muestra un óptimo en 8 clases (figura 27), se escoge la clasificación en 9 clases (figura 26) pues ella muestra una separación en la zona de la Cuenca del Plata la cual será relevante en la proyección de los modelos. En el apéndice se pueden apreciar distintos resultados de la clasificación y cómo se las regiones se van separando.

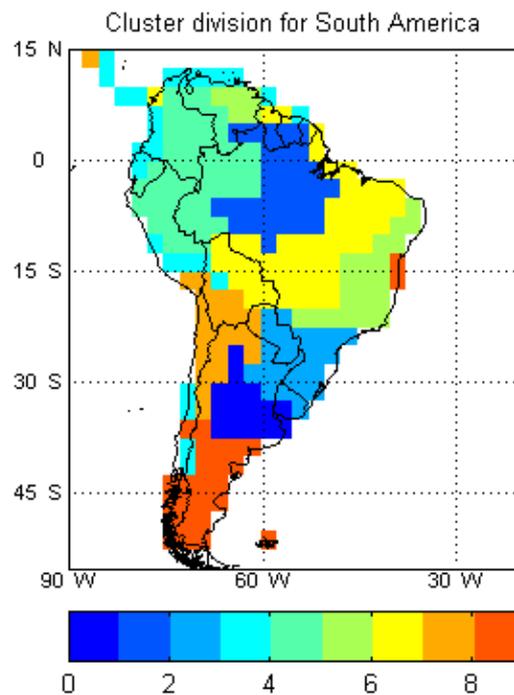


Figura 26 - Clasificación a 9 clusters

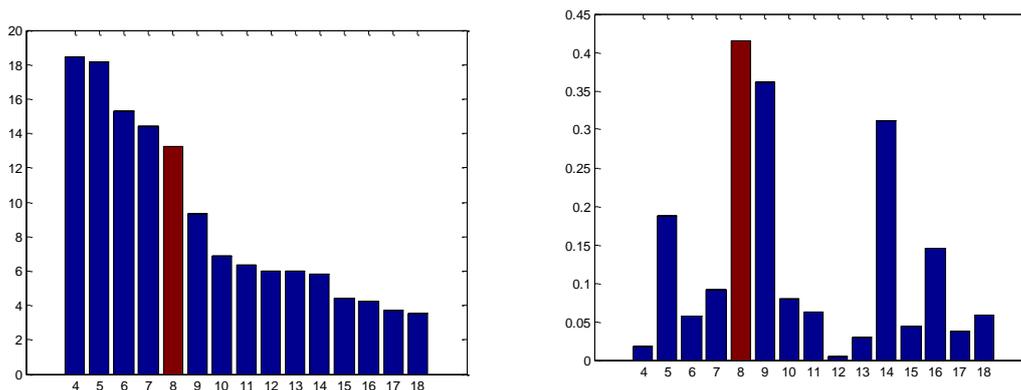


Figura 27 - función de Ward para elección de número de clusters; valores absolutos y relativos

En el mapa de 9 clases se puede apreciar lo siguiente:

- Clase 1: Sur de la Cuenca del Plata.
- Clase 2: Clima amazónico.
- Clase 3: Centro de la Cuenca del Plata. Clima cálido con variabilidad subtropical en invierno y más tropical en verano.
- Clase 4: Clima de la Cordillera. Fuerte influencia de las depresiones que vienen del Pacífico.
- Clase 5: Clima influenciado por el Pacífico Este Ecuatorial.
- Clase 6: Parte oriental de la zona del Monzón Sudamericano.
- Clase 7: Parte occidental del Monzón Sudamericano.
- Clase 8: Zona del altiplano boliviano.
- Clase 9: Zona sur Patagonia (Andes más bajos). Similitudes entre clima de Chile y Argentina.

3.5 Proyección de los modelos respecto a los valores observados

Como último objetivo se verá cómo se comportan los modelos al predecir valores de temperatura durante el promedio 1961-1990, y se compararán contra los valores observados obtenidos a través de interpolación (figura 28), y contra los valores de obtenidos cuando se produce un aumento global a 2 grados en el siglo 21 (figura 29).

En general se observa como en esta región los modelos se encuentran por encima de los 2 grados de aumento global, diferenciando un segmento de la Cuenca del Plata y obviamente a la región patagónica. Se destaca el modelo **ukmo_hadcm3** que evidencia un mayor aumento de temperatura en toda la región, seguido por el **gfrl_cm2_1**. En la parte inferior sobresale el **ncar_ccsm3_0**.

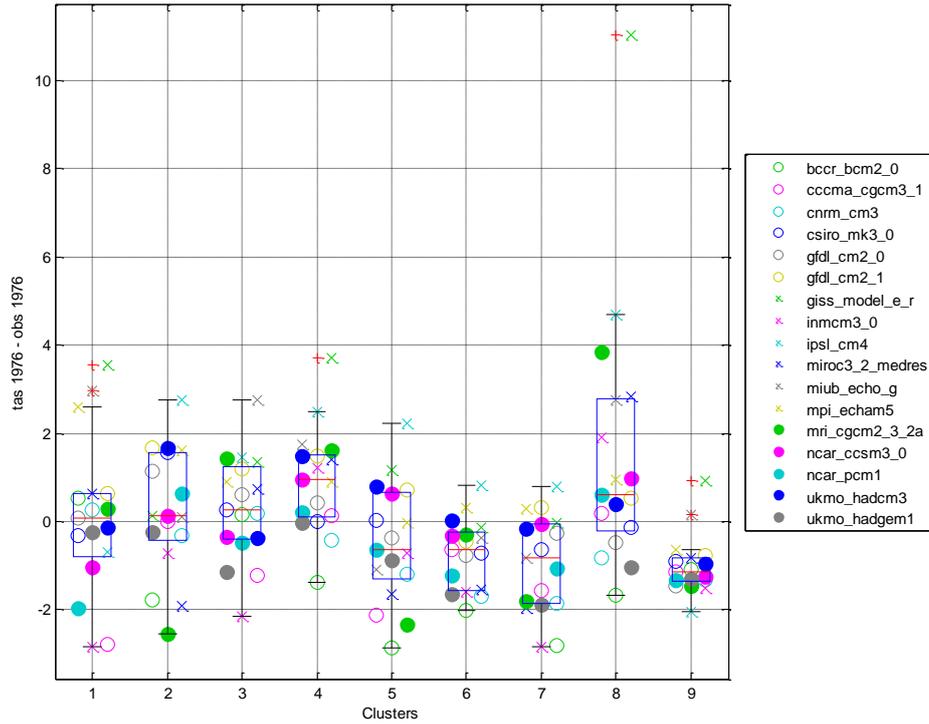


Figura 28 - Boxplot de dispersión en la exactitud de los modelos contra los valores observados en el año promedio 1976.

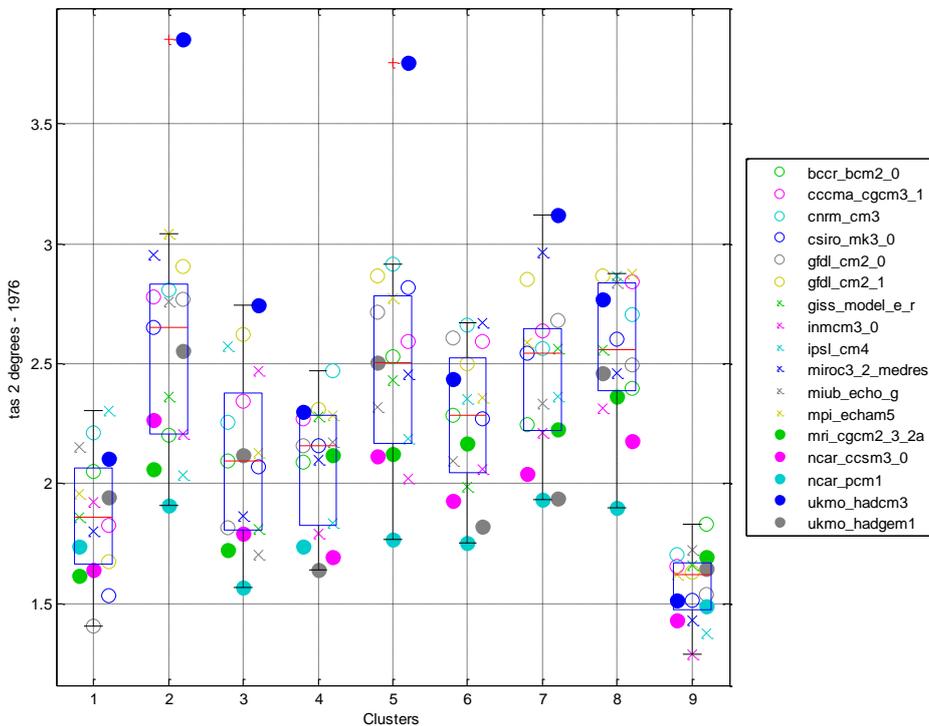


Figura 29 - Boxplot de dispersión del delta de temperatura (de nuevo entre 2060 y 1976) para todos los modelos para cada clase, adornado con los valores puntuales de cada uno.

Se observa que no hay coherencia entre clases respecto a si los modelos evalúan una mayor o una menor temperatura que la verdaderamente ocurrida. Aquí el modelo **giss_model_e_r** claramente predijo mayor temperatura de la que ocurrió en la época por toda Sudamérica. Lo contrario ocurre con el modelo **bccr_bcm2_0**.

Ahora se hará foco en las clases 1, 3, 8 y 9 que son las que tienen influencia en Argentina. Para cada una se mostrarán los siguientes plots de comparación (los datos observados corresponden a los resultados de la interpolación de la temperatura basada en los valores arrojados por las estaciones meteorológicas del SMN)

Ejes de los gráficos:

Aumento de temperatura Siglo 21 – Siglo 20: diferencia entre los valores registrados en la variable temperatura del modelo al momento de llegar al aumento global de 2 grados, contra lo que registra al modelar el año 1976 (promedio del período de referencia)

Cambios de precipitación Siglo 21 – Siglo 20: diferencia entre los valores registrados en la variable precipitación del modelo al momento de llegar al aumento global de 2 grados, contra lo que registra al modelar el año 1976.

Diferencia de temperatura entre Siglo 20 – valores observados: diferencia entre los valores registrados en la variable temperatura del modelo en el año 1976 contra los valores reales ocurridos en la región en dicho año.

Diferencia de precipitación entre Siglo 20 – valores observados: diferencia entre los valores registrados en la variable precipitación del modelo en el año 1976 contra los valores reales ocurridos en la región en dicho año.

Gráficos de comparación:

- Aumento de temperatura Siglo 21 - Siglo 20 vs diferencia de temperatura entre Siglo 20 - valores observados: *para tener una idea si los modelos tienden a sobreestimar la temperatura en general de la región.*
- Cambios de precipitación Siglo 21 - Siglo 20 vs diferencia de precipitación entre Siglo 20 - valores observados: *análisis del comportamiento de los modelos respecto a las precipitaciones.*
- Aumento de temperatura Siglo 21 - Siglo 20 vs diferencia de precipitación entre Siglo 20 - valores observados: *busca resaltar la relación entre la estimación de precipitación en el siglo 20 y la temperatura que el modelo predecirá en el siglo 21.*
- Aumento de temperatura Siglo 21 - Siglo 20 vs diferencia de precipitación Siglo 21 - Siglo 20: *ver si el aumento de temperatura refleja una disminución en la precipitación o también un aumento en la misma.*

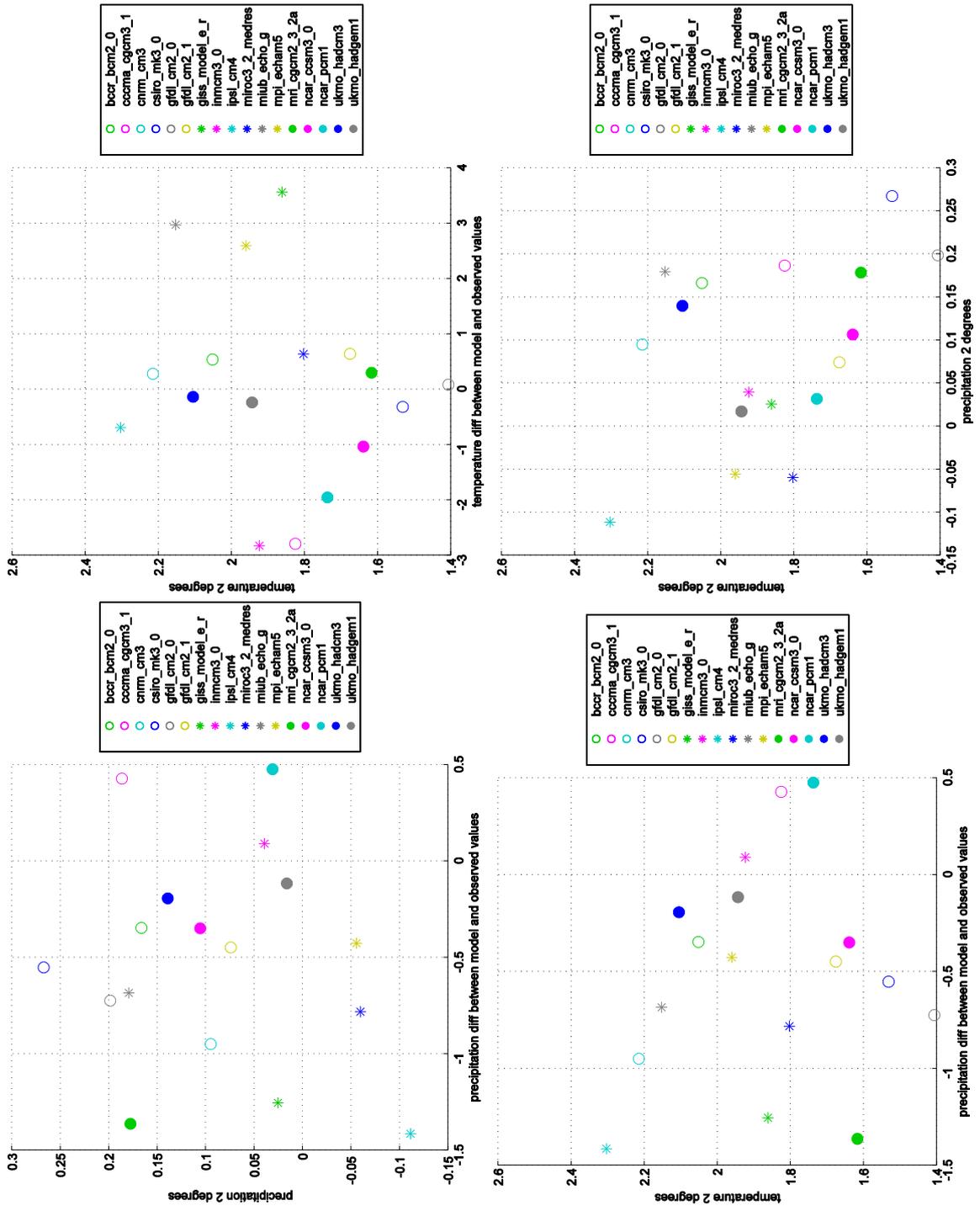


Figura 30 - Comparación de modelos en la clase 1

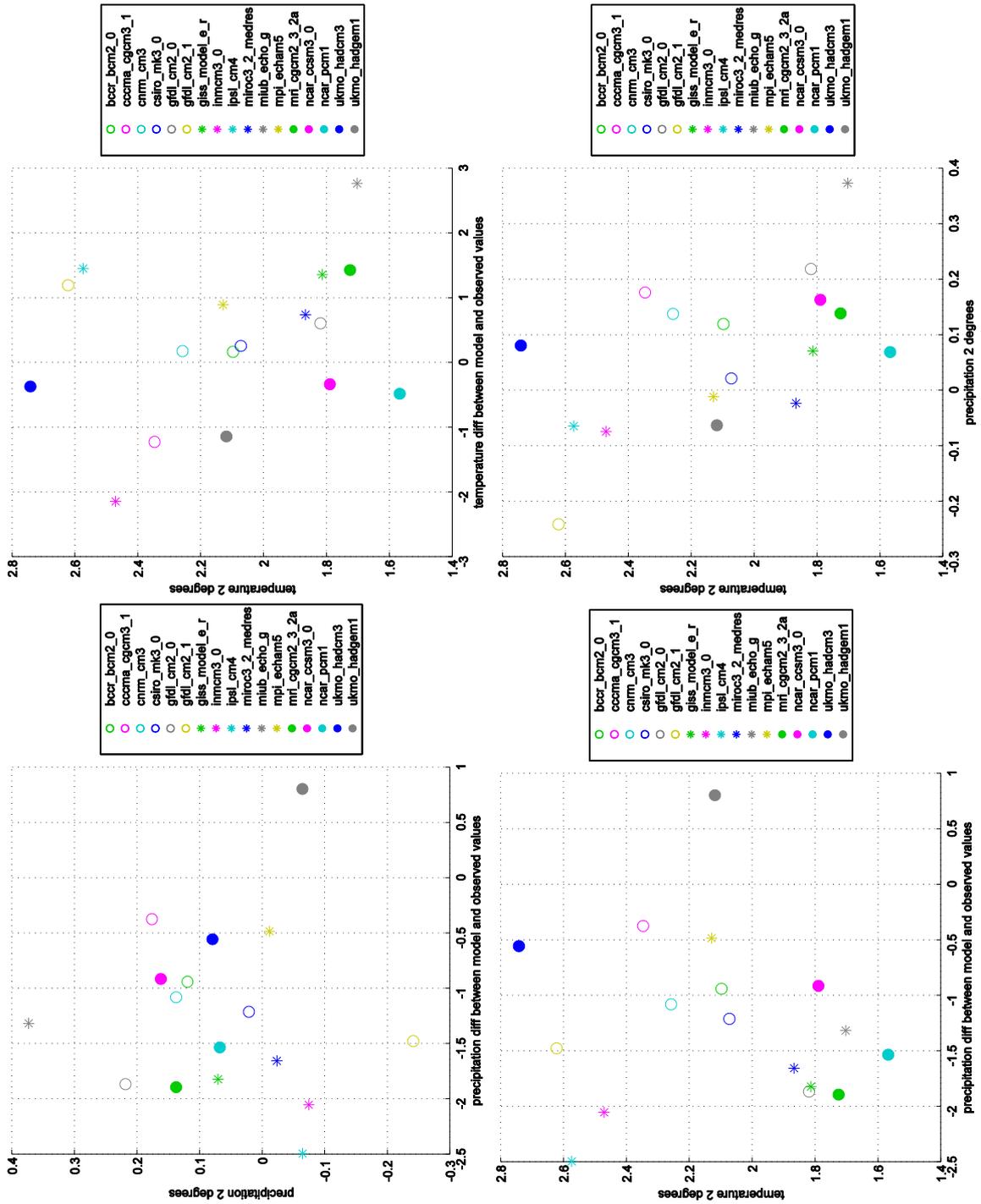


Figura 31 - Comparación de modelos en la clase 3

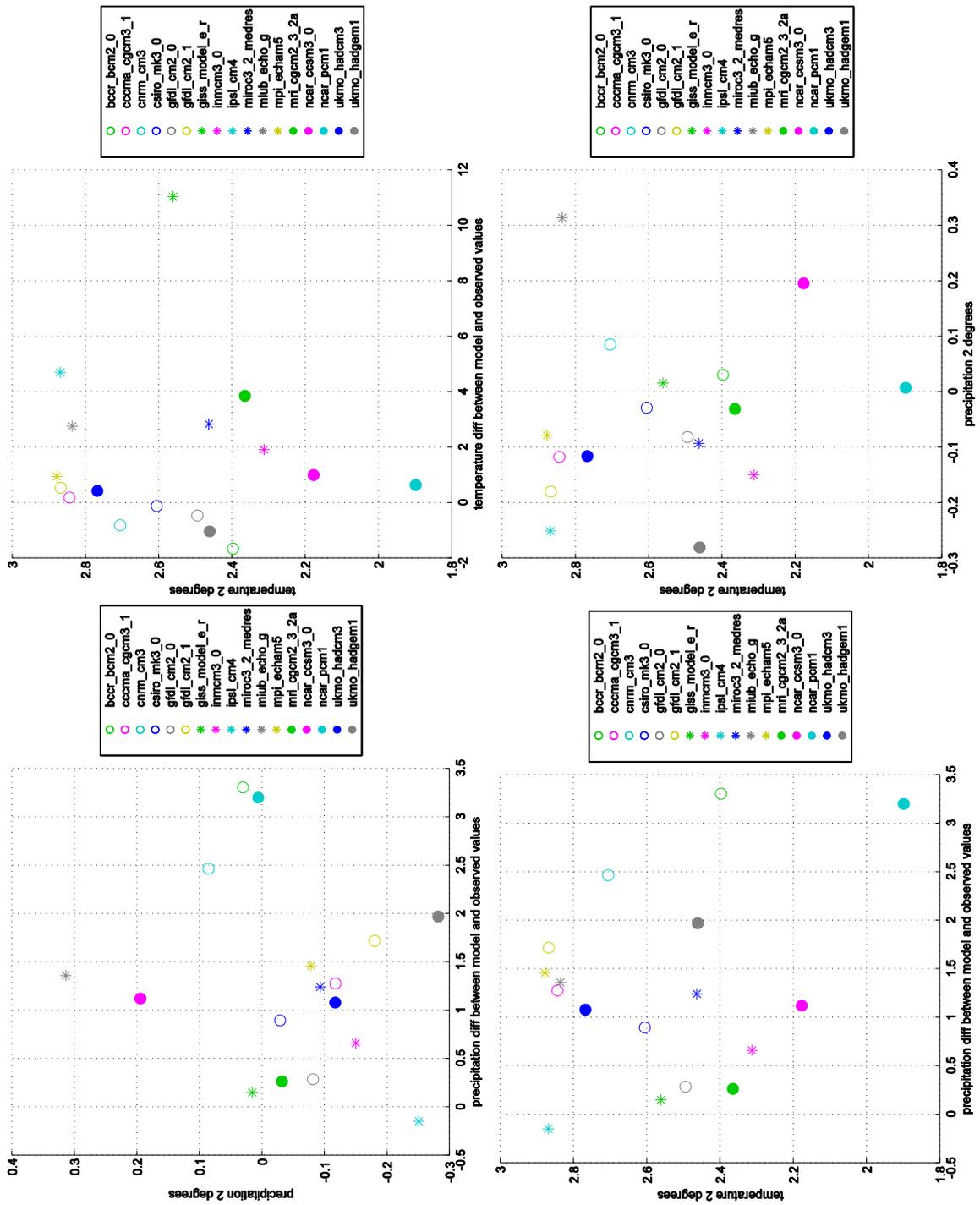


Figura 32 - Comparación de modelos en la clase 8

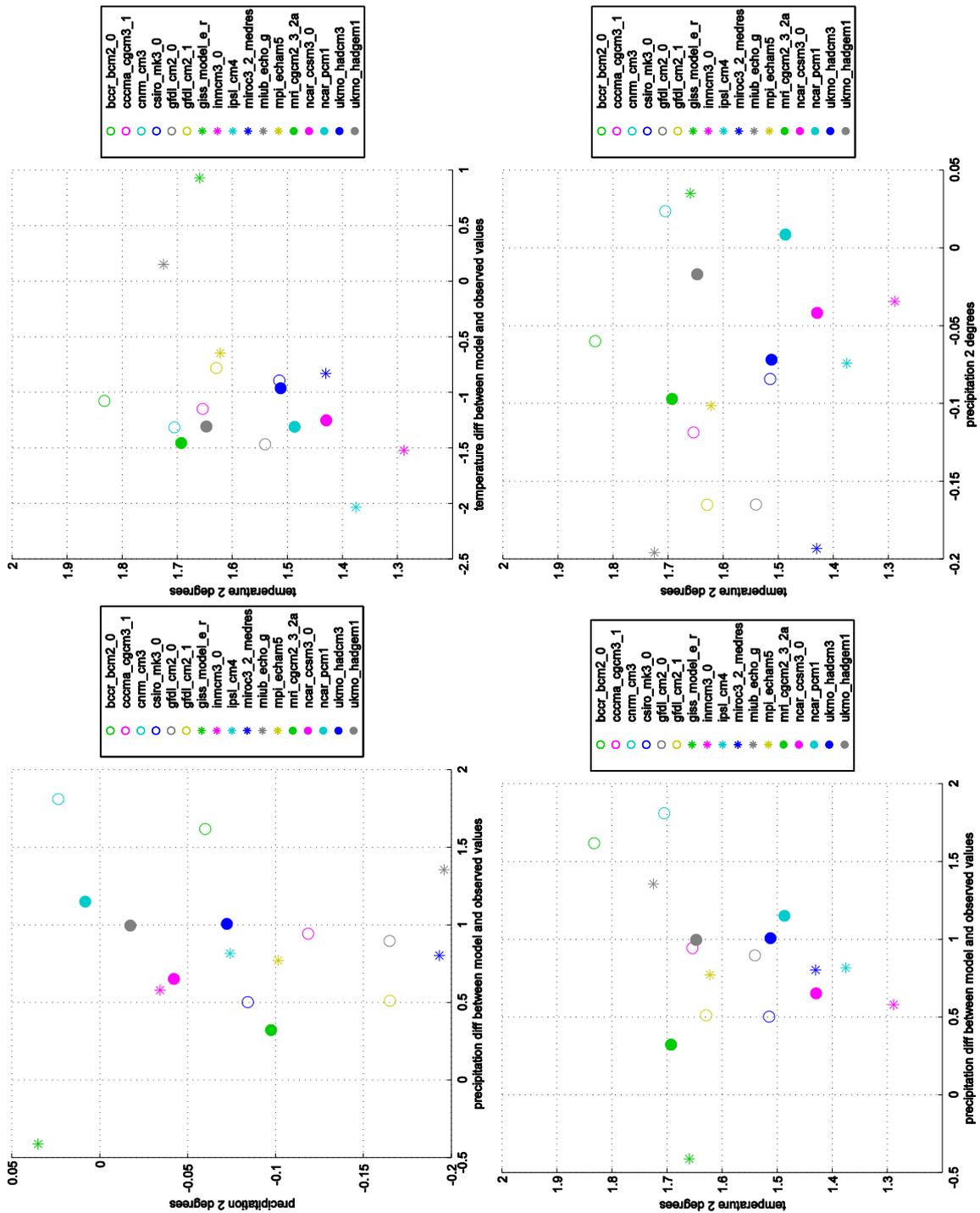


Figura 33 - Comparación de modelos en la clase 9

De los gráficos se puede concluir lo siguiente:

- Clase 1 (figura 30): Sur de la Cuenca del Plata

Se aprecia una proyección a 2 grados donde a mayor temperatura disminuye la precipitación promedio. Además, la mayoría de los modelos presentan un aumento de precipitación en el futuro sin importar cuán exacta sea su proyección en el siglo 20 contra los datos observados.

- Clase 3 (figura 31): *Centro de la cuenca del plata*

Casi todos los modelos subestiman la precipitación caída en el siglo 20. Sin embargo, se observa una tendencia en la que a menor temperatura a futuro, mayor es la subestimación de la precipitación en el presente. Viendo la precipitación a futuro se nota que gran cantidad de los modelos predice un aumento en la misma.

- Clase 8 (figura 32): *Noroeste argentino y altiplanos bolivianos*

Casi todos los modelos sobreestiman el monto de precipitación promedio en el siglo 20.

Para la proyección a futuro también se visualizan valores mayores a los 2 grados globales para los modelos. Existe una leve tendencia entre la precipitación estimada en el presente contra la temperatura proyectada a futuro, donde a mayor valor de precipitación, también es mayor el valor de la temperatura.

- Clase 9 (figura 33): *Zona de la Patagonia*

De aquí se puede concluir que a mayor aumento (desviación respecto a los valores observados) de la precipitación en los datos del siglo 20, ocurre un mayor aumento en la diferencia de temperatura entre siglo 21 y siglo 20.

Queda como trabajo futuro sacar mayor información a partir de los gráficos aquí mostrados.

3.6 Conclusiones

Los modelos de cambio climático estiman la temperatura que ocurrirá en el siglo 21 en base a complejos cálculos y simulaciones. Todos coinciden que en un momento habrá un aumento de temperatura a nivel global de 2 grados respecto al período de referencia del siglo 20 del IPCC. Para este trabajo se tomaron los valores de temperatura del año en que cada modelo llega a ese umbral, especializándose solamente en aquellos puntos que caen sobre la tierra de Sudamérica.

Para poder obtener información útil de tanta cantidad de datos, los mismos se ordenaron y organizaron para cada punto del mapa, para luego aplicarles un Mapa Auto-Organizado. Esta red tiene como salida un espacio más reducido de variables y en un primer intento se visualizó a través de una U-matrix con fines de obtener alguna clasificación notoria de los datos resultantes. Al no poder ser factible esta tarea se decidió aplicar un análisis jerárquico de clústeres para reducir el espectro a 9 clases. Se observa como aun habiendo dejado los parámetros de afinación del SOM en sus valores por omisión, el resultado final sobre el mapa de Sudamérica son 9 clases bien distribuidas en el espacio que reflejan a los distintos tipos de climas existentes en la región. La gran capacidad de aprendizaje sin supervisión de este tipo de redes se ve reflejada en los resultados semánticamente coherentes que ofrece.

En principio para las 9 clases se muestran las distribuciones de temperatura de los modelos tanto para lo que modelaron del siglo 20 como para lo que proyectan a 2 grados. Luego se hizo foco en aquellas clases que cubren el territorio Argentino (norte y sur de la Cuenca del Plata, zona

Patagónica y Noroeste Argentino) para realizar gráficos de comparación sobre cada una que muestren tendencias de los modelos. A partir de los resultados es posible conocer mejor la situación climatológica de la Argentina así como idear nuevas herramientas de análisis en base a las aquí presentadas.

3.7 Trabajo futuro

3.7.1 Clasificación a nivel global

En la figura 34 se muestra el resultado de aplicar el proceso de redes neuronales y posterior clusterización a los valores de temperatura en el momento de la proyección a 2 grados de cada modelo, a nivel global. Se ve como la coherencia espacial se sigue manteniendo, permitiendo un análisis de los distintos tipos de clima existentes en todo el planeta.

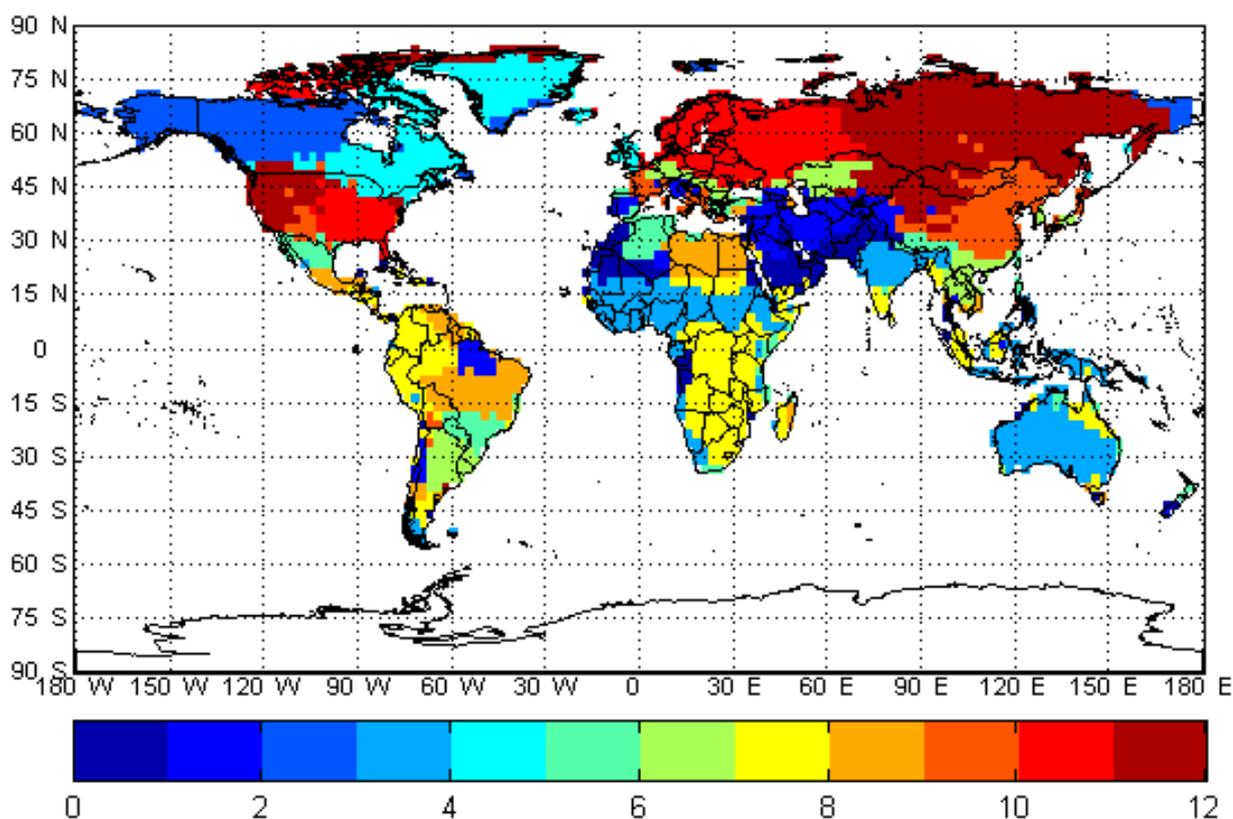


Figura 34 - Doce clases resultantes de aplicar el método de redes neuronales sobre todo el planeta (puntos terrestres)

Si bien un análisis a nivel mundial hace difícil la obtención de datos relevantes debido a la gran variedad de climas que se presentan (aunque sí se puede observar como las zonas tropicales o subtropicales se relacionan de la misma manera), aplicando este método es posible obtener primeras impresiones sobre hechos globales, y a medida que existan más modelos de simulación del clima y estos reproduzcan con mejor exactitud la situación cuando se los comparan con los valores observados, mayor información será posible obtener de los mismos.

3.7.2 Proyección de los modelos en otras regiones

Por otro lado, si no se desean realizar análisis globales sino observar los sucesos ocurridos localmente como se hace en este trabajo, es algo totalmente factible y pueden existir proyectos que realicen una metodología similar pero enfocándose en zonas como Australia, África o Europa.

3.8 Bibliografía

- [1] Houghton et al., Climate Change 2001: The Scientific Basis. Cambridge University, 2001
- [2] Boulanger et al., Projection of future climate change conditions using IPCC simulations, neural networks and Bayesian statistics, 2006
- [3] Kohonen T., Self-Organizing Maps Third Edition, 2001
- [4] Vesanto J. et al., SOM Toolbox for Matlab 5, 2000
- [5] Ultsch A. et al., Knowledge Extraction from Artificial Neural Networks and Applications, 1993
- [6] Vesanto J. et al., Clustering of the Self-Organizing Map, IEEE Transactions on Neural Networks Vol 11.

Conclusiones generales

Durante este informe presentamos tres herramientas desarrolladas en el marco de un proyecto común, con el fin último de estudiar proyección de indicadores de cambio climático para el siglo XXI.

En esta etapa de desarrollo, las herramientas fueron aplicadas al registro histórico del Servicio Meteorológico Nacional, con lo que finalmente se relevó la vigencia de ciertos modelos conocidos de cambio climático al territorio argentino. Estas herramientas están en curso a seguir siendo utilizadas en el marco del proyecto CLARIS, enfocado al estudio de cambio climático en territorio sudamericano.

Además de este fin natural que tiene el trabajo presentado, se presentaron algunos resultados interesantes.

En la primera parte del trabajo se presentó APACH, un procedimiento automatizado de control de calidad. Este trabajo se presenta en [1] J. P. Boulanger, coescrita por los autores de este trabajo y se encuentra, en versión adaptada, en este mismo informe. APACH fue aplicada al registro del SMN como parte de la consistencia del mismo, relevándose aproximadamente un 10% de registros diarios que presentan distintos tipos de situaciones anómalas. Dentro de este gran porcentaje se observan datos claramente erróneos y otros que seguramente se asumirán como válidos.

En general, la distribución de registros erróneos está totalmente relacionada al contexto histórico/geográfico en el que ocurren. Es decir que los resultados son coherentes con la realidad, con lo que APACH aportó información novedosa al Servicio Meteorológico Nacional, y al tratarse de una metodología no-paramétrica, sería de esperar que se pueda aplicar con la misma facilidad, en principio, a cualquier región del mundo. Esta herramienta fue presentada como aplicación para computadoras de escritorio, con lo que se probó su uso como solución general, reutilizable en cualquier contexto, por un usuario sin conocimientos de desarrollo de software.

En la segunda parte de este trabajo se presentó una metodología de estimación de variograma, basada en técnicas novedosas (redes neuronales). Esta metodología mostró grandes ventajas a la hora de generar un modelo más complejo de variograma, ofreciendo la misma capacidad inicial de predicción, mientras que posibilita la incorporación instantánea de información adicional al modelo, que es utilizada de forma transparente, mejorando la predicción en un 25%.

Para la última parte del trabajo se trabajó con modelos de predicción del clima, analizando los datos que modelan para el siglo XXI, en el momento que predicen un aumento a 2 grados, un valor del cual el organismo IPCC cree que no se debe sobrepasar. Utilizando ese corte a dos grados, se hizo un foco en Sudamérica para separarla en 9 regiones, donde cada una representa climas diferentes. Esta separación se realizó utilizando Mapas Auto-organizados, y se muestra como este sistema de redes neuronales aprende de manera automática las similitudes de cada modelo para lograr tal clasificación. Luego se tomaron las clases que ocupan el territorio argentino y sobre ellas se realizaron gráficos de comparación de los modelos, para analizar cuánto se desvían las proyecciones de éstos en el siglo 20 respecto a los datos verídicos, y cómo tal desviación influye en sus predicciones a futuro.

Apéndice

1. APACH: Prueba Espacial

Esta prueba se aplica de la siguiente manera:

1. Se preseleccionan estaciones vecinas, ubicadas a menos de 500km de la estación siendo analizada, con una correlación mayor a 0.8 (la correlación se computa solo con datos del mismo mes: enero, febrero, etc.) y con un nivel de significancia mayor a 99%. Los resultados no son significativamente diferentes si uno usa valores mayores o menores de correlación (en un rango 0.7-0.85). Se adoptó este valor para poder tener suficientes vecinos para computar las pruebas espaciales, y que estuvieran altamente correlacionados para evitar introducir ruido en la interpolación.
2. Para cada estación preseleccionada, se computó una regresión lineal entre los datos diarios de la estación analizada y la estación vecina (basados en todos los años comunes para el mes del dato a analizar), y mantener a los vecinos para los cuales el error residual de la regresión es a lo sumo 1.5 veces la variabilidad de la temperatura en la estación.

A pesar de la baja densidad en la red de estaciones meteorológicas, más del 80% (resp. 70%) de las observaciones máximas (resp. mínimas) diarias de temperatura tienen al menos dos estaciones vecinas que satisfacen a las condiciones de selección; alrededor del 10% de los datos de temperatura mínima y máxima sólo tienen una estación vecina, y alrededor del 10% (resp. 20%) de los datos de temperatura máxima (mínima) no tienen estaciones vecinas contra las cuales validar. Tales observaciones (y sus estaciones meteorológicas relacionadas) se encuentran ubicadas principalmente en la parte sureña de Argentina.

3. Entonces se computó el valor interpolado de los vecinos de la siguiente forma:

$$T_{\text{int}}(t) = \frac{\sum_{i \in [1, N]} w_i (\alpha_i^m T(t)_i + \beta_i^m)}{\sum_{i \in [1, N]} w_i}$$

donde α_i^m and β_i^m son los coeficientes de la regresión lineal entre la estación analizada y la vecina, N es el número total de estaciones vecinas, m es el mes del momento (t) del dato a ser analizado, y w_i es el peso de la estación vecina en la interpolación computada usando un método similar al usado por Thornton et al. (1997):

$$w_i = e^{-a \left(\frac{r_i}{R}\right)^2} - e^{-a} \text{ if } r_i \leq R$$

$$w_i = 0 \text{ if } r_i > R$$

R es la distancia máxima usada para elegir vecinos (500km), r_i es la distancia entre las estaciones vecinas y la analizada, y a es a un factor de escalado. Se probaron con diferentes valores y se escogió un valor de 3 de la misma manera que en Thornton et al. (1997). Cambios pequeños en este valor (de un rango 2-4) no afectaron significativamente al valor interpolado y por consecuencia a los resultados finales.

4. Luego se computa la diferencia entre el valor analizado y el interpolado. La diferencia es normalizada por la desviación estándar de la temperatura diaria del mes del día a comprobar. A mayor la distancia, mayor la confianza de que el dato es erróneo.
5. Finalmente, también se computó el ángulo alrededor de la estación analizar cubierto por todas las estaciones vecinas. Se definió para cada estación vecina el ángulo entre un paralelo al Ecuador y una línea derecha relacionando a la estación analizada (el punto de referencia) y la estación vecina. Se computó el ángulo total cubierto por todas las estaciones vecinas para poder cuantificar si se encontraban todas en una dirección similar o realmente cubrían gran parte del área alrededor de la estación.
6. En conclusión, para una distancia dada, a mayor el ángulo cubierto por las estaciones vecinas, más fuerte es la aserción de que el dato es útil (Useful) o erróneo (Suspect). Similarmente, para un ángulo dado, a mayor la distancia, más fuerte es la aserción de que el dato es útil o erróneo.

Distancia	Ángulos alrededor de la estación cubiertos por las estaciones vecinas elegidas			
	0-90°	90°-180°	180°-270°	270°-360°
>3	N2	N3	N4	N4
2-3	N2	N2	N3	N4
1.5-2	N1	N1	N2	N3
1-1.5	N1	N1	N2	N2
0-1	N0	N0	N0	N0

Tabla 1 – Matriz de decisiones para la prueba de confianza de vecinos. Si no hay vecinos, el código de la prueba es NN

2. APACH: Códigos de control de calidad de la temperatura y niveles de confianza

Prefijos (Pruebas locales)	DX	Valor de Dip muy largo, mayor a 5.5
	DL	Valor de Dip largo, entre 3 y 5.5
	DM	Valor de Dip medio, entre 2.5 y 3
	LE	El dato diario es un extremo de 3 días
	LN	El dato diario no es un extremo de 3 días
	SS	Series de 2 días con el mismo valor
	SL	Series de 3 días o más con exactamente el mismo valor de temperatura
	OX	La prueba de Dip no es computable y la prueba de periférico es mayor a 3.5
	OL	La prueba de Dip no es computable y la prueba de periférico está entre 2.5 y 3.5
	OM	La prueba de Dip no es computable y la prueba de periférico está entre 1.5 y 2.5.
	OO	La prueba de Dip no es computable y la prueba de periférico es menor a 1.5.
Sufijos (pruebas de vecindad)	NN	No hay vecinos contra los cuales comprobar
	Ni	Resultado de la prueba de vecinos (con i desde 1 a 4; ver Tabla 1).
	NEO	La proyección de temperatura de la estación vecina es un extreme de 3 días del signo opuesto
	NEI	La proyección de temperatura de la estación vecina es un extreme de 3 días del mismo signo
	NEZ	La proyección de temperatura de la estación vecina no es un extremo de 3 días
	NEN	La prueba de extremos de estaciones vecinas no puede ser computada

Tabla 2 – Códigos que describen el camino de cada dato diario por el árbol de decisión de la Figura 4

	NN	N0	N1	N2	N3	N4
DX	Doubftul	Useful	NeedCheck	Doubftul	Suspect	Suspect
DL	NeedCheck	Useful	NeedCheck	Doubftul	Suspect	Suspect
DM	NeedCheck	Useful	NeedCheck	NeedCheck	Suspect	Suspect
LENEO	NeedCheck	Useful	NeedCheck	Doubftul	Doubftul	Doubftul
LENEI	NeedCheck	Useful	NeedCheck	NeedCheck	Doubftul	Doubftul
LENEZ	NeedCheck	Useful	NeedCheck	NeedCheck	Doubftul	Doubftul
LENEN	NeedCheck	Useful	NeedCheck	NeedCheck	Doubftul	Doubftul

LN	NeedCheck	Useful	NeedCheck	NeedCheck	NeedCheck	Doubftul
SL	NeedCheck	Useful	NeedCheck	Doubftul	Suspect	Suspect
SS	Useful	Useful	NeedCheck	NeedCheck	Doubftul	Doubftul
OX	NeedCheck	Useful	NeedCheck	NeedCheck	Doubftul	Suspect
OL	NeedCheck	Useful	Useful	NeedCheck	NeedCheck	Doubftul
OM	NeedCheck	Useful	Useful	NeedCheck	NeedCheck	NeedCheck
OO	Useful	Useful	Useful	Useful	Needcheck	Useful

Tabla 3 – Tabla de confianza de temperatura

3. APACH: Detalles de las pruebas de vecindad en sequías

- a. **Primera prueba de vecindad.** Si la correlación es mayor que 0.8 y alcanza a un nivel de significancia del 99%, se “interpolan” los datos de precipitación de la estación vecina hacia la estación de referencia al computar los valores de precipitación con el mismo percentil que en la distribución de la estación vecina. Se computa el valor SPI (Standardized Precipitation Index, McKee et al., 1993) correspondiente. Para todas las estaciones vecinas, se escoge aquella con el SPI más seco y se computa la diferencia con el SPI de la sequía de la estación de referencia. El nivel de confianza es sugerido de acuerdo a una función de tal diferencia (Tabla 4a).
- b. **Segunda prueba de vecindad.** En la segunda prueba de vecindad, la correlación entre las series temporales de la estación de referencia y la estación vecina no son lo suficientemente grandes para aplicar la primera prueba. Por lo tanto, considerando que las características de la precipitación pueden ser similares (con la misma distribución, o sea, dos estaciones vecinas tienen el mismo clima), aunque la correlación no es tan alta como 0.8, se comparan las distribuciones de precipitación usando la prueba Kolmogorov-Smirnov. Si la prueba sugiere que las dos distribuciones son idénticas, con un nivel de significancia del 90%, se escoge la estación como una estación vecina relevante, se calcula su valor SPI (como en la primera prueba de vecindad) y la diferencia con el valor SPI de la sequía de la estación de referencia. El nivel de confianza es sugerido de acuerdo a una función de tal diferencia (Tabla 4b).
- c. **Tercera prueba de vecindad.** Si alguno de los registros de las estaciones vecinas es superior a 20 años, se computa su distribución de promedio acumulado. Entonces se computa el percentil de la precipitación en las estaciones vecinas registrados en el mismo período que la sequía de la estación de referencia. Se identifican tres códigos (Tabla 4d): MIN (si el valor es el valor mínimo de la distribución), P1 (menor al percentil 10^{mo}), P2 (menor al percentil 20^{mo}), y P3 (mayor al percentil 20^{mo}).

Primera prueba de Vecindad		Clasificación de la sequía en la estación		
		SPI_ED	SPI_SD	SPI_MD
Clasificación de la sequía en la estación vecina	SPI_ED	Useful	Useful	Useful
	SPI_SD	NeedCheck	Useful	Useful
	SPI_MD	Doubtful	NeedCheck	NeedCheck
	SPI_NN	Suspect	Doubtful	Doubtful
	SPI_MW	Suspect	Suspect	Suspect
	SPI_VW	Suspect	Suspect	Suspect
	SPI_EW	Suspect	Suspect	Suspect

Tabla 4a – Tabla de confianza cuando se aplica la primera prueba de vecindad

Segunda prueba de Vecindad		Clasificación de la sequía en la estación		
		SPI_ED	SPI_SD	SPI_MD
Clasificación de la sequía en la estación vecina	SPI_ED	Useful	Useful	Useful
	SPI_SD	NeedCheck	Useful	Useful
	SPI_MD	Doubtful	NeedCheck	NeedCheck
	SPI_NN	Doubtful	Doubtful	Doubtful
	SPI_MW	Doubtful	Doubtful	Doubtful
	SPI_VW	Doubtful	Doubtful	Doubtful
	SPI_EW	Doubtful	Doubtful	Doubtful

Tabla 4b – Tabla de confianza cuando se aplica la segunda prueba de vecindad

	Clasificación de la sequía en la estación			Estación con series temporales demasiado cortas
	SPI_ED	SPI_SD	SPI_MD	SS
NN	NeedCheck	NeedCheck	NeedCheck	NeedCheck
ND	Useful	Useful	Useful	Useful

Tabla 4c – Tabla de confianza cuando no hay datos vecinos contra los cuales comprobar (NN) o cuando alguna estación vecina presenta otra sequía en el mismo período (ND)

Tercera prueba de Vecindad	Estación con series temporales demasiado cortas
Código de estación vecina	SS
MIN	Useful
P1	NeedCheck
P2	NeedCheck
P3	Doubtful

Tabla 4d – Tabla de confianza cuando el valor SPI no puede ser computado en la estación (serie muy corta). Los códigos MIN, P1, P2 y P3 se refieren a la tercera prueba de Vecindad

4. Referencia para el usuario de APACH

Importación/Exportación de datos

Este menú engloba el sistema de importación, administración y exportación de datos en el sistema. La importación toma datos de uno o varios archivos de texto en formato CSV (separado por comas) los cuales cumplen con un formato específico, que requiere los siguientes campos:

- Estación: Identificador numérico, Nombre, Provincia
- Ubicación: Latitud, Longitud, Altura
- Fecha
- Valores registrados: Temperatura mínima y máxima (en decimas de grado) y precipitación (en milímetros caídos).

Una importación de datos comprende 3 tareas: Validar la información contra ciertas restricciones, persistirla en la base, y obtener información calculada.

Al finalizar una importación con éxito se garantiza que la información procesada es consistente.

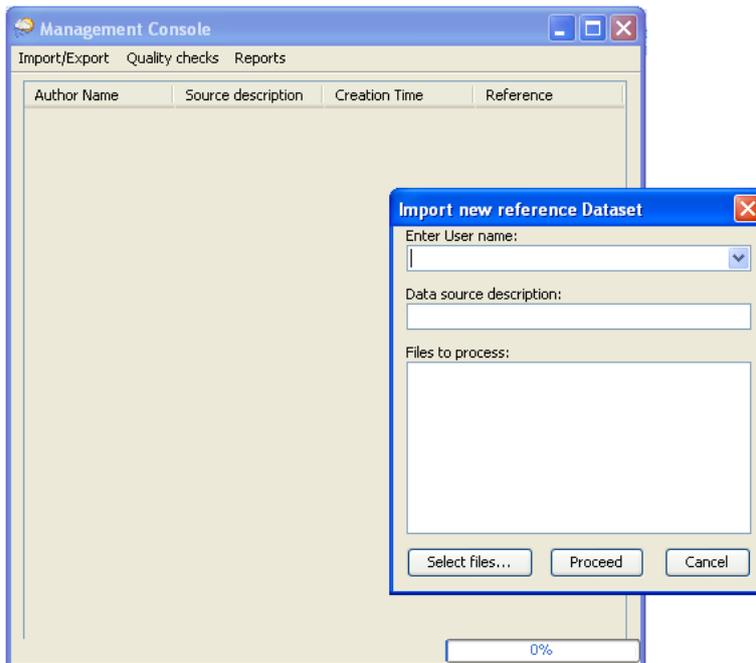


Figura 35 - Consola de Usuario, junto al menú de importación

Dentro de este mismo menú se ofrecen también las opciones de exportar la información de la base a un archivo de texto y de eliminar información de la base.

Control de Calidad

La funcionalidad central de la aplicación son los diversos controles de calidad que se pueden realizar sobre datos importados. Esto se presenta al usuario en el menú “Quality Checks”, comprendido por opciones para hacer control de calidad de Temperatura y de Precipitación (por separado).

Al invocar estas funcionalidades, se ejecutan los controles de calidad sobre los datos y se persisten los resultados en la base de datos. Esta información queda asociada a los mismos registros analizados y disponibles para futuros reportes y exportaciones.

Reportes

Dentro de este menú de la aplicación se encuentran opciones para visualizar gráficamente y exportar en formato de texto los resultados de los controles de calidad. Se pueden realizar los siguientes reportes:

- **Visor de sequías:** Este visor consta de dos pantallas. En la primera se muestra, para la estación seleccionada, un gráfico con la distribución formada por las sequías registradas, destacando a aquellas cuya longitud sea mayor al percentil 75% de la distribución. A estas sequías es posible analizarlas con mayor detalle, para lo cual se abre una segunda pantalla en la que se comparan los valores de precipitación caídos en el mismo período de la sequía pero en todos los años disponibles de la estación. También se revelan los valores registrados en estaciones vecinas en el mismo período, así como niveles de correlación entre estaciones, y valores de SPI. Es útil para revisar sequías dudosas en la región.

- Visor de precipitaciones mensuales: Muestra una ventana con las distribuciones de precipitación enfocadas en una estación y mes particulares. Para cada año se muestra la precipitación acumulada en la estación (solamente del mes elegido) y debajo se muestran los días en los que se registraron precipitaciones que superen al percentil 50% y con mayor resaltado a las que superan el 75%, con vistas de poder observar precipitaciones inusuales (y posiblemente erróneas).

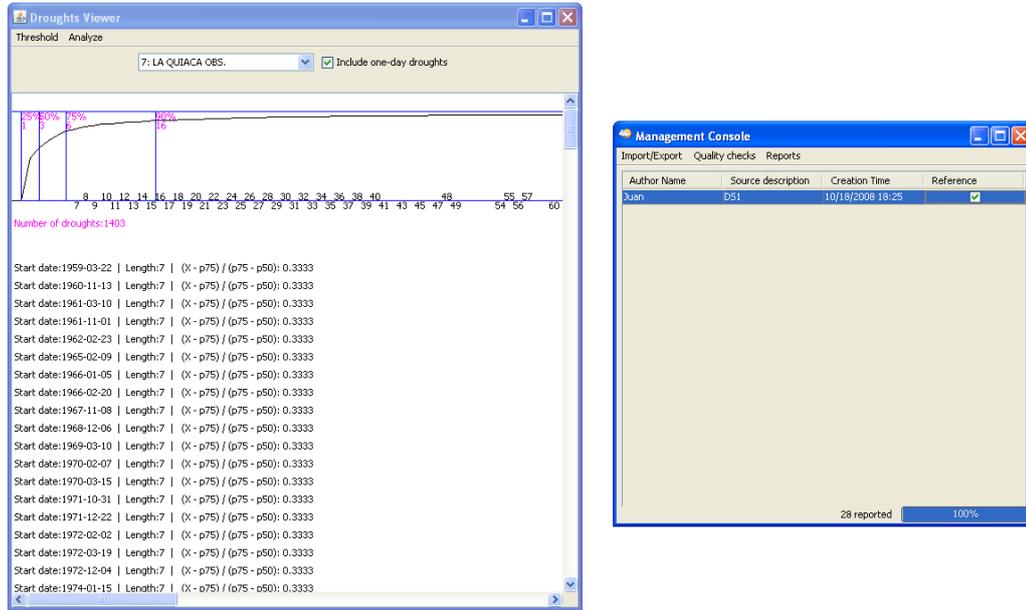
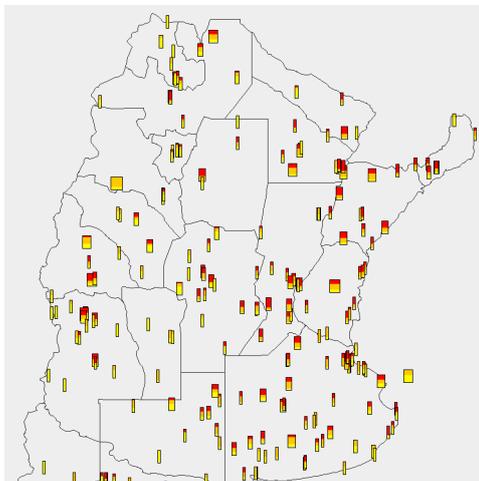


Figura 36 - Reporte de Visualización de sequias

- Reporte de precipitaciones extremas: Crea un archivo de texto en un formato de valores separados por comas (.csv) con los resultados de la clasificación de precipitación para las estaciones seleccionadas.
- Visor geográfico de problemas de temperatura: Se grafican en un mapa, un rectángulo por cada estación. Cada rectángulo representa con su ancho la cantidad de datos sospechosos, y con los colores internos la cantidad de datos sospechosos de cada tipo (NeedCheck, Doubtfull y Suspect). Este gráfico se realiza por separado para temperatura mínima y máxima.



- Reporte de problemas de temperatura: Este reporte presenta en un archivo de texto todos los registros que resultaron sospechosos, incluyendo datos sobre el registro en si junto a datos sobre la razón de la sospecha.

Concepto de Dataset y sus operaciones

Cuando se importa información desde archivos de texto a la aplicación esta se persiste agrupada en un Dataset. Un Dataset tiene además de la información meteorológica meta-información para poder identificarla.

La interfaz gráfica presenta justamente en su ventana principal el detalle de los Datasets cargados en el sistema. Se informa quién lo cargó, cuándo lo hizo y una breve reseña. La exportación, eliminación de información, controles de calidad y reportes se realizan luego sobre un dataset en particular. Esto permite que el mismo programa maneje varias bases en simultáneo.

5. Implementación de APACH

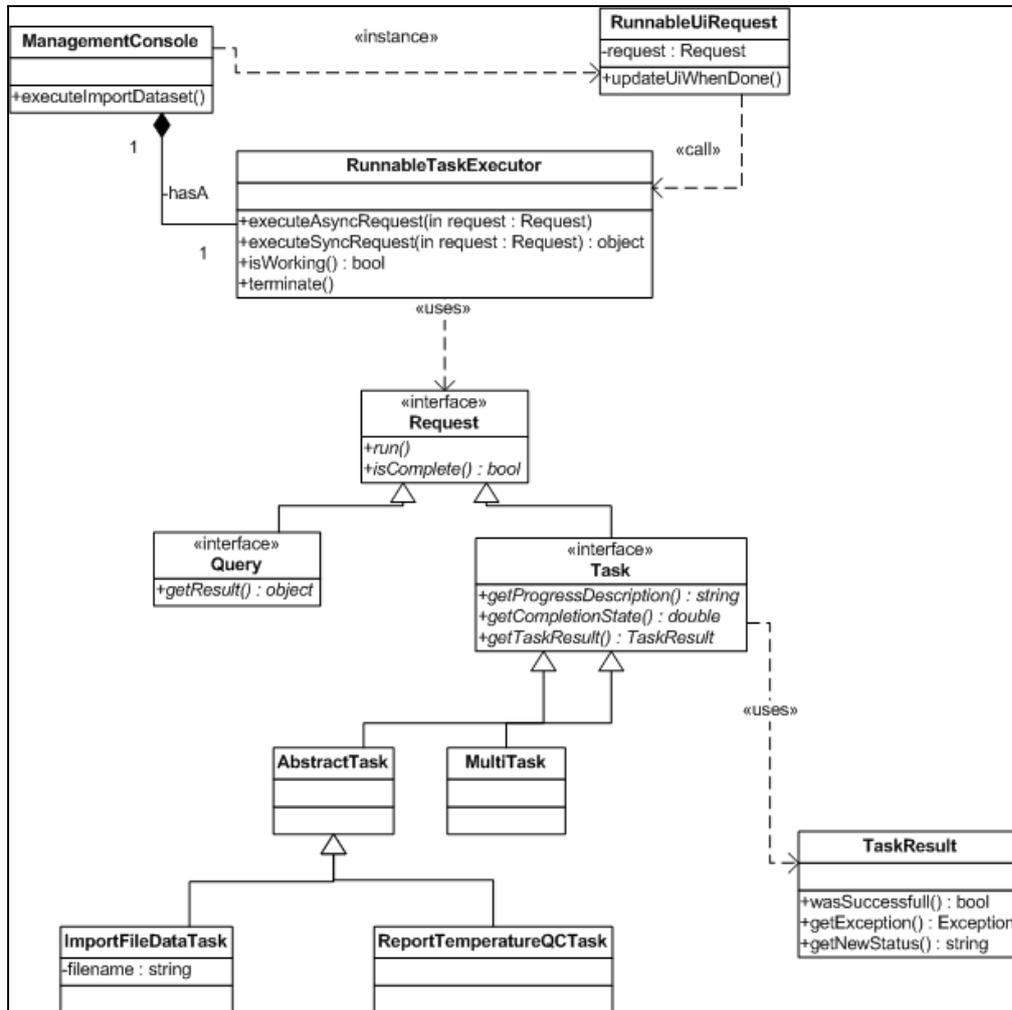
Interfaz gráfica

La interfaz gráfica desarrollada permite a un usuario interactuar con la aplicación para utilizar los procesos de control de calidad sobre un conjunto de series meteorológicas cualquiera. Las funcionalidades se encuentran divididas en 3 áreas disjuntas, cada una representada en un menú de la aplicación: Procesos de importación de datos y consistencia, Control de calidad, y Exportación de resultados y reportes.

Además de los menús, la interfaz presenta una barra de estado en el margen inferior. Aquí se informa si se encuentra ejecutándose alguna tarea (ej. importación de datos o comprobación de calidad). Dependiendo de cada tarea se muestra una estimación de avance, ya que algunas pueden llevar un tiempo considerable en completar su ejecución.

Diagrama de clases de la Capa “Interfaz Gráfica”

Éste es el diagrama de clases de la interfaz gráfica:



La clase [ManagementConsole](#) centraliza los pedidos de un usuario a la aplicación a través de la consola. Es decir, cada opción de menú tiene asociado un método `execute*`. Estos métodos compilan información sobre la tarea a ejecutar (a veces pidiendo detalles al usuario) e invocan a `executeRequest(Request)`.

Este método genera una instancia de [RunnableUiRequest](#) en cada invocación. Estos delegados supervisan la ejecución y se encargan de mantener al usuario al tanto del progreso. Para resolver la ejecución, los [RunnableUiRequest](#) invocan a [RunnableTaskExecutor](#), una variable Singleton que procesa los requests.

Cada pedido del usuario se modela como una instancia de [Request](#). Estas instancias se encargan de invocar al servicio necesario en la capa de servicios, y ofrecer información actualizada sobre el estado durante el transcurso del proceso.

[RunnableTaskExecutor](#) se encarga de invocar y supervisar los [Request](#), permitiendo ejecutarlos síncronica o asincrónicamente.

El primer caso es más adecuado para acciones de pocos segundos, y se modela en general con instancias de [Query](#), que extiende [Request](#) solo para exponer el resultado de su ejecución. Esta

interfaz suele utilizarse para consultas a la base de datos, con bajo tiempo de procesamiento y donde lo importante es el resultado de la búsqueda.

El segundo caso agrupa acciones más largas, con tiempos de ejecución de hasta varias horas, como el caso de control de calidad de temperatura para bases de datos grandes. En estos casos se utiliza `Task`, que extiende `Request` agregando información sobre el avance de la ejecución y el estado final de su resultado. Al ser procesos de cálculo, estos `Request` no suelen devolver a la interfaz más feedback que sobre la correctitud de la ejecución.

Multithreading

La aplicación involucra procesos que pueden llevar horas, durante los cuales se debe informar continuamente el estado al usuario y ofrecer la posibilidad de interrumpir el avance. Para lograr esto hizo falta manejar varios hilos de ejecución en paralelo. Vamos a detallar los hilos que maneja la aplicación, y cómo se relacionan estos con el diagrama de clases descripto más arriba.

Interfaz: La interfaz grafica tiene su propio hilo. Desde éste se dibujan las pantallas que ve el usuario (código genérico de Java), y se escuchan los eventos del usuario, como clics del mouse, ingreso de texto, etc. En este hilo corre la clase `ManagementConsole`. Cuando una tarea se corre secuencialmente, todos los participantes en ejecutar esa tarea corren en este hilo, mientras que en caso de tareas asincrónicas se usan dos hilos más.

Ejecución de tareas: Este es el hilo principal para los procesos de cálculo. En el mismo corre `RunnableTaskExecutor` junto a los `Task` que invoca. Naturalmente al utilizarse un solo hilo para la ejecución de tareas se impide que estas se realicen en paralelo. Este efecto es totalmente deseado, ya que se busca que cada proceso tenga exclusividad de acceso a datos.

Feedback: Este hilo hace de intermediario entre los otros dos. La interfaz `Task` pide a sus implementaciones que las variables acerca del estado de avance se actualicen con tanta frecuencia como sea posible sin entorpecer el proceso. Esta información esta sincronizada, por lo que el hilo de Feedback la busca periódicamente y la utiliza para mantener al tanto al usuario a través de la interfaz. De la misma forma, se pregunta periódicamente a la tarea si sigue en ejecución (a través de `Request.isComplete()`) y en caso de estar completada se utiliza su status final para actualizar por última vez la interfaz grafica. Para esta acción se combinan `Task.getTaskResult()` y `RunnableUiRequest.updateUiWhenDone()`.

Capa de procesos

En esta capa están implementados los controles de calidad de APACH, y en definitiva todos los procesos descriptos anteriormente.

Esta capa ofrece, a través de su API pública, la funcionalidad necesaria para resolver los pedidos de la interfaz gráfica. La API está compuesta por un subconjunto de las clases del modelo de dominio, interfaces con operaciones para actuar sobre estas y dentro de estas, por un conjunto de métodos, cada uno representando un servicio que se puede invocar desde la interfaz grafica. Los `Tasks` de esta capa deberían tener responsabilidades propias sencillas, básicamente para manejo de excepciones y compilación de información y delegar toda la funcionalidad concreta a través de la API de la capa de procesos.

Modelo de dominio

A continuación se presentan las clases del modelo de dominio, descritas visualmente en la figura 37.

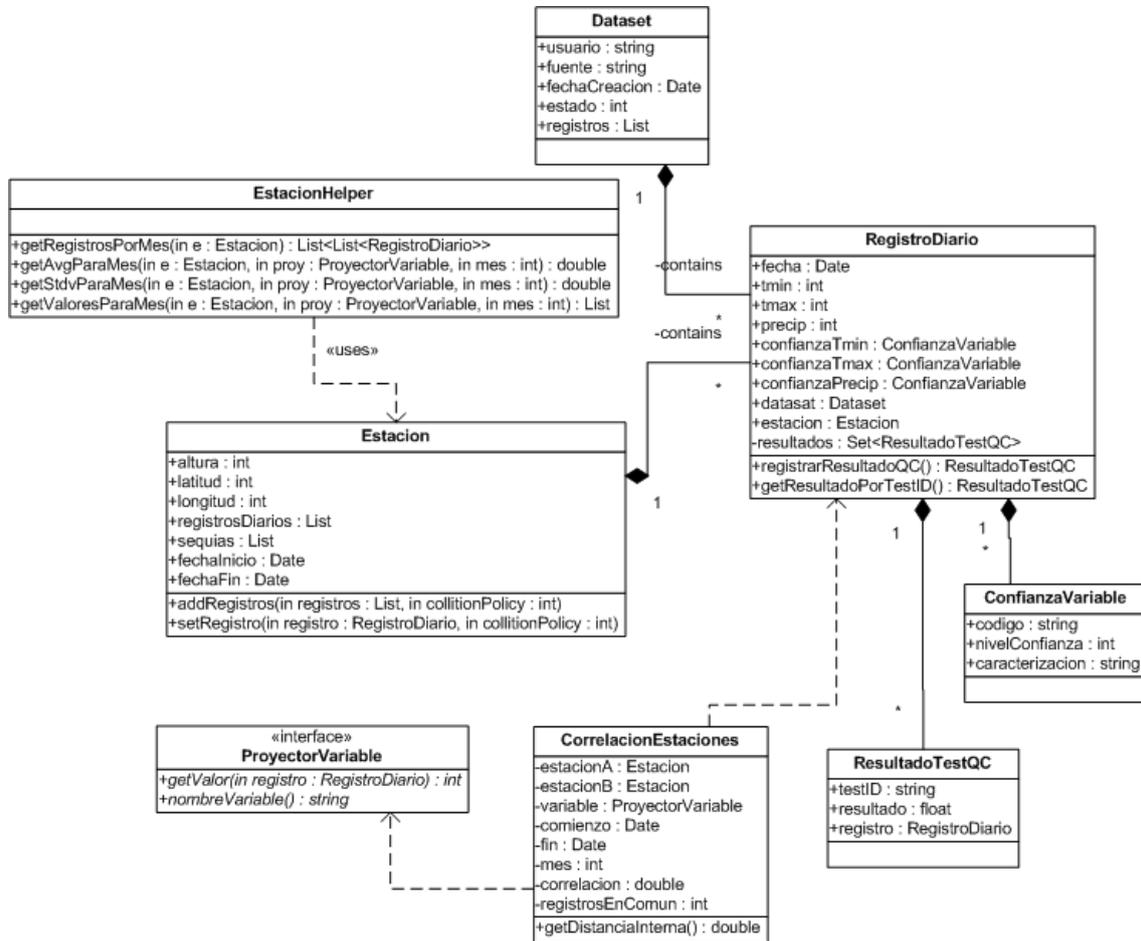


Figura 37 - Modelo de dominio de la aplicación

La entidad central en esta aplicación es el registro diario, por ser el objeto de todos los controles de calidad. Estos contienen información sobre su estación, fecha y variables medidas. Luego de pasar por el primer control de calidad (QC1) se le asocian un **ResultadoTestQC** por cada variable por cada test de la bacteria. Durante QC2 se utilizan estos resultados intermedios para asociar una categoría de confianza a cada una de las 3 variables. Los resultados de los tests de QC1 se persisten de todas formas para futura referencia.

Las categorías de confianza están modeladas en **ConfianzaVariable**. Cada categoría tiene una caracterización (i.e. porque un registros tendría asignada esa categoría), un nivel de confianza (OK, Need Check, Doubt o Suspect) y un código alfanumérico de 7 caracteres.

Dataset es un simple contenedor de datos, sin lógica propia. **Estacion** es también un acumulador de información y presenta lógica para agregar registros eficientemente: Los métodos **addRegistros()** y **setRegistro()** permiten agregar registros masivamente a la estación delegando en esta el manejo de colisiones. Esta funcionalidad fue clave para lograr un proceso de importación temporalmente eficiente.

ProyectoVariable es una interfaz muy utilizada a lo largo de la aplicación, principalmente para modelar las 3 variables de un registro, a través de instancias conocidas **TMIN**, **TMAX** y **PRECIP**.

CorrelacionEstaciones modela la correlación de Pearson entre dos estaciones para temperatura mínima o máxima. Estos se usan en distintos momentos de los controles y fuera de ellos, por lo que resulta más práctico persistirlos.

Acceso a la base de datos

Como se mencionó en el punto **¡Error! No se encuentra el origen de la referencia.**, esta capa maneja la base de datos en conjunto con la capa de persistencia. Para esto depende fuertemente de Hibernate, un framework ORM (object-relational mapping), útil para mapear un modelo de dominio basado en objetos a una base de datos relacional.

Utilizando las librerías de Hibernate se simplifica el acceso a base de datos, proveyendo sustitutos a las interacciones SQL, y permitiendo mantener la aplicación 100% en lenguaje Java. De esta forma se encapsulan (fuera de la aplicación) las complejidades propias de bajo nivel del uso de JDBC/SQL.

Hibernate funciona en esencia mapeando clases Java a tablas SQL, a través de archivos XML. Cuando se inicia la aplicación, Hibernate chequea la validez de estos mapeos contra la base, la actualiza de ser necesario, y permite generar objetos a partir del estado de la base (load), crear instancias nuevas, actualizarlas y borrarlas (insert, update y delete). Para poder mantener la coherencia de datos entre ambas, hibernate utiliza proxies de los objetos manejados, y un objeto **Session** con el que se interactúa explícitamente para ejecutar cambios en la base de datos. Los objetos proxy permiten entre otras cosas que la información de la base de datos se pueda recuperar solo en caso de ser necesario.

A continuación, las clases participantes en la persistencia de datos en la aplicación.

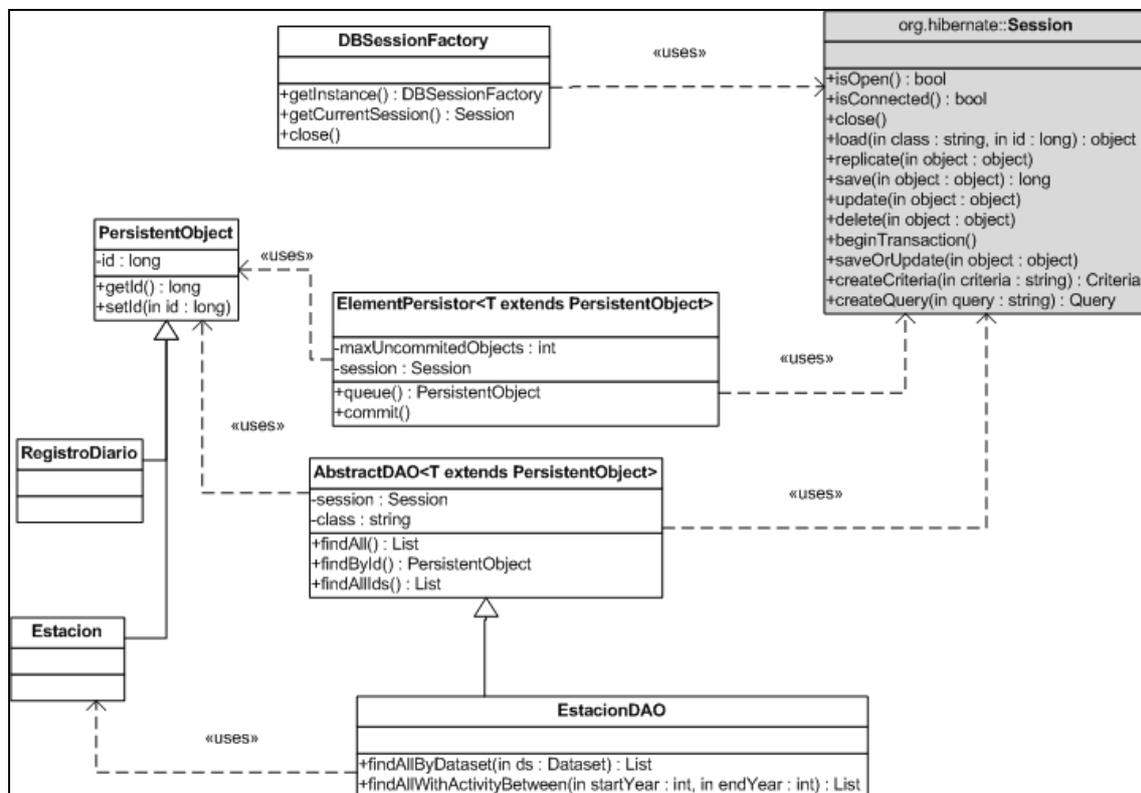


Figura 38 - Mecanismos de persistencia

La clase `Session` del framework de Hibernate es el punto más importante de acceso explícito a la base. Desde la session se pueden realizar las mismas operaciones que a través de SQL, ya sea utilizando los métodos `load`, `save()`, `update()`, `delete()`, o creando un `Criteria` (representación de un query dentro del paradigma de objetos) o un `Query HQL`.

Además del acceso explícito, ya mencionamos que Hibernate crea proxies para las instancias que maneja. Estos se utilizan entre otras cosas para recuperar la información de la base de datos solo cuando sea necesario. Es decir, al recuperar una estación `e` de la base, la variable `e.registros` será un proxy vacío. Pero cuando el usuario hace la primera llamada a `e.getRegistros()`, el proxy automáticamente se comunica con la sesión para recuperar la lista de registros de la base de datos. La aplicación depende fuertemente de esta modalidad para manejar una cantidad de información que de otra forma debería requerir decenas de millones de objetos en memoria.

Todas las operaciones realizadas a través de una `Session` deben ejecutarse en el contexto de una `Transaction`. Para esto, cuando un servicio realiza un acceso a la base tiene que encargarse explícitamente de manejar la transacción correspondiente.

Para obtener una `Session` (y por transitividad una conexión a la base) los servicios piden una instancia de `DBSessionFactory`. Esta clase instancia sesiones interactuando con los archivos de configuración de Hibernate, y si bien en este contexto mono-usuario solo maneja una instancia Singleton está preparada para manejar una instancia por thread en un contexto multi-usuario.

En el contexto de la aplicación, todas las entidades persistentes heredan de `PersistentObject`, y por lo tanto cuentan con un id de tipo Long. Estas entidades son exactamente las referenciadas en la Figura . Los servicios pueden manejar estas entidades directamente a través de la sesión o en algunos casos utilizando la infraestructura de DAO (por el patrón Data Access Object). Esta consiste en una clase abstracta `AbstractDao`, que ofrece queries genéricos para distintas entidades, y las subclases correspondientes a cada entidad, que ofrecen queries específicas, en general implementadas como queries HQL.

`ElementPersistor` se utiliza exclusivamente para la importación masiva de información. Estas entidades se encargan de validar persistir la información que les llega, de forma eficiente.

Base de datos

Durante el desarrollo de la aplicación, se utilizó MySQL como soporte SQL. La integración de este motor fue casi transparente, ya que solo es necesario crear una base y configurar Hibernate para que se conecte a la misma. Para esto se configuran los datos de conexión a la base y un Driver idóneo a la implementación, que en el caso de MySQL viene incorporado al mismo framework.

Una vez que se inicia la aplicación, Hibernate se encarga de sincronizar sus mapeos de datos con la estructura de la base de datos. Si hubiera diferencias significativas como falta de tablas o columnas, Hibernate las resuelve automáticamente.

Lo único que hizo falta manejar explícitamente desde la base de datos fueron los índices, imprescindibles dados los grandes cambios de eficiencia que implican. Por ejemplo durante varios casos de uso es necesario obtener todos los registros de una estación. Para una estación típica esto implica seleccionar 7000 registros de una tabla de 2,5 millones. El tiempo de resolver

esta consulta sin índice puede ser de un minuto o más, mientras que agregando un índice de tipo “Árbol B” simple este baja a 3 segundos en la misma instancia.

Más adelante en el desarrollo de la aplicación los requerimientos cambiaron. La eficiencia mejoró mucho y por líneas separadas tomo fuerza el requerimiento de empaquetar la aplicación para que fuera instalable y administrable por un usuario final. Estos dos sucesos nos llevaron a utilizar Apache Derby como soporte de persistencia de la aplicación.

Derby es un motor SQL escrito íntegramente en Java, muy eficiente dentro de las limitaciones que esto implica, y que aprovecha al máximo las posibilidades de embeberse con aplicaciones Java.

La migración de MySQL a Derby fue casi trivial, con gran ayuda de Hibernate. Cambio la ruta JDBC de conexión junto a algunos otros datos como el “dialecto” y “driver” utilizado para la generación de SQL. Fuera de esto la tarea central fue lograr que los recursos de Derby (librerías Java + carpetas de datos) se incluyan con la aplicación, y puedan moverse, duplicarse y eliminarse. Justamente el uso de Derby permitió que se puedan hacer back-ups de la base de datos de forma transparente al usuario, simplemente con una rutina de copia y compresión de archivos.

A continuación se muestra un Diagrama de Entidad-Relación del modelo de persistencia utilizado, con las entidades, sus atributos y las relaciones entre las mismas.

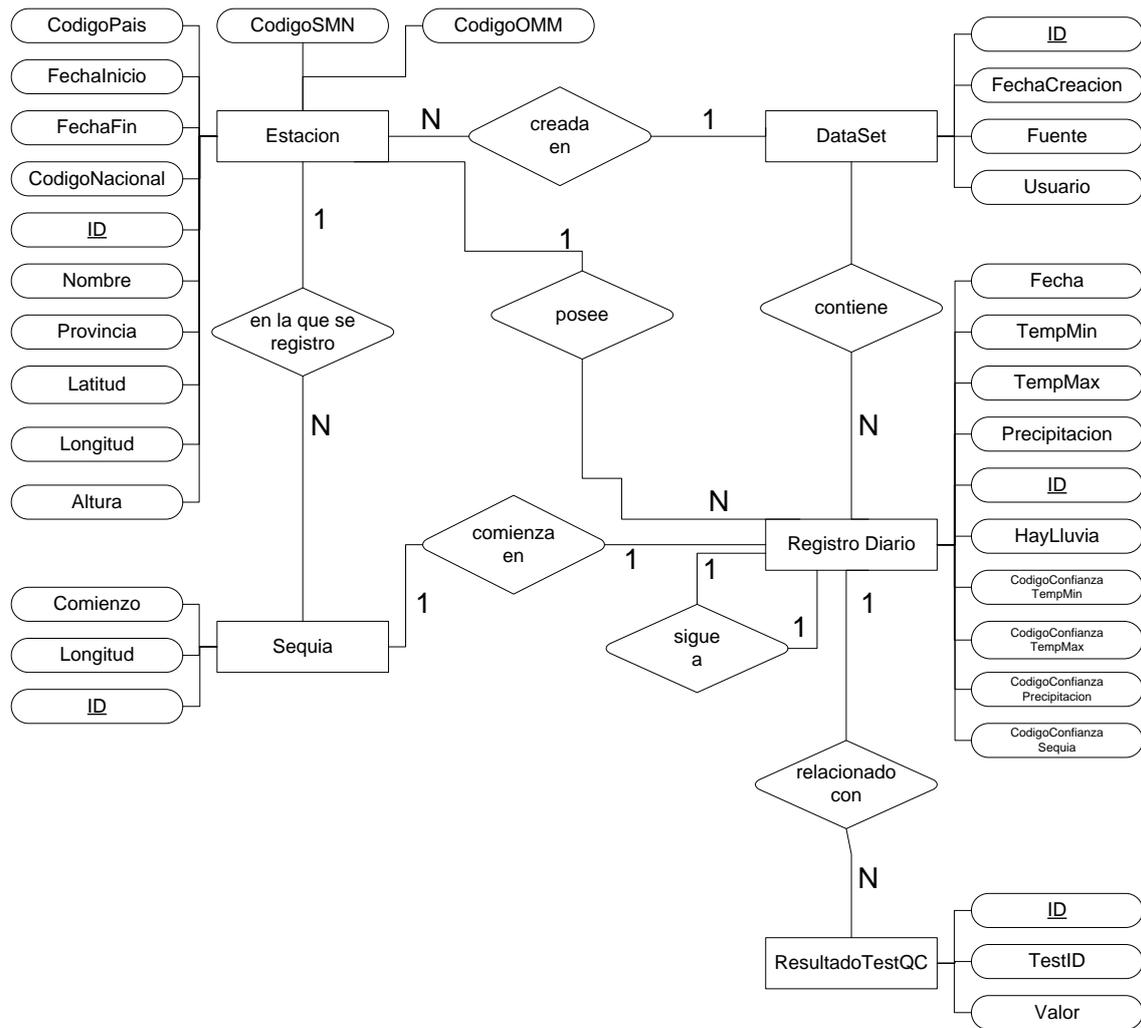


Figura 39 - DER de la aplicación

La entidad **DataSet** engloba al conjunto de datos que se agrega al sistema. Funciona como unidad mínima para importación, controles de calidad y exportación.

Entre los datos que engloba se encuentran los **Registros Diarios**, que contienen los sucesos ocurridos en una fecha en particular para una Estación dada. Se guardan valores de temperaturas mínimas y máximas y de precipitación. Luego de haberse corrido las pruebas de calidad, a cada registro se le asocia un código de confianza por cada variable, para poder generar reportes de calidad.

La entidad **Estacion** también engloba un conjunto de **Registros Diarios**, y representa a una estación meteorológica en particular. Entre sus atributos se incluye su ubicación, altura y códigos descriptivos (OMM y SMN para el caso Argentino). Al no relacionar una estación con un **DataSet** particular es posible utilizar la misma estación con registros de distintos datasets, y al mismo tiempo forzar la consistencia de la estación y sus registros. Esto es la base para poder realizar controles de calidad multi-dataset.

La entidad **Sequia** es calculada a partir de registros diarios consecutivos que no presenten precipitación. La misma se relaciona con el registro del primer día de la secuencia, y es una entidad muy útil durante las pruebas de calidad que de precipitación.

Por último, la entidad [ResultadoTestQC](#) relaciona a un registro diario con una prueba en particular, y contiene el valor obtenido en dicha prueba de calidad para ese registro.

6. Familias de escenarios del IPCC

Las descripciones están basadas en la versión AR4, que son las mismas que las de TAR.

A1

Los escenarios A1 son de un mundo unido. Esta familia se caracteriza por:

- Rápido crecimiento económico.
- Una población mundial que alcanza los 9 mil millones en el 2050 y luego decrece gradualmente.
- La rápida propagación de nuevas y eficientes tecnologías.
- Un mundo convergente – los ingresos y el estilo de vida se unifican entre regiones. Mundialmente extensas interacciones sociales y culturales.

Existen subclases de la familia A1 según el énfasis tecnológico:

- A1FI – Énfasis en combustibles fósiles.
- A1B – Balance en todas las fuentes de energía.
- A1T – Énfasis en Fuentes de energía no fósiles.

A2

Los escenarios A2 reflejan un mundo dividido. Sus características son:

- Un mundo de naciones independientes y autosuficientes.
- Una población creciente continua.
- Desarrollo económico orientado regionalmente.
- Lentos y fragmentados cambios tecnológicos y mejoras al ingreso per cápita.

B1

Estos escenarios muestran un mundo unido y consciente de la ecología. Sus características son:

- Crecimiento económico tan rápido como en A1, pero con cambios hacia una economía de información y servicios.
- Una población creciente hacia 9 mil millones en el 2050 y luego decreciendo, como en A1.
- Reducciones en intensidad material y la introducción de tecnologías limpias y eficaces.
- Énfasis en soluciones globales a la estabilidad económica, social y ambiental.

B2

Por último los escenarios B2 son de un mundo dividido, pero consciente de la ecología. Los mismos se caracterizan por:

- Una población continuamente creciente, pero de una manera más lenta que en A2.
- Con énfasis en soluciones locales a la estabilidad económica, social y ambiental (en vez de globales).
- Niveles intermedios de desarrollo económico.
- Cambios tecnológicos menos acelerados y más fragmentados que en B1 o en A1.

7. Descripción de los Modelos Climatológicos

bccr_bcm2_0 Bjerknes Centre for Climate Research, Norway - Bergen Climate Model Version 2.0

ccma_cgcm3_1 Canadian Centre for Climate Modelling and Analysis, Canada - Coupled Global Climate Model Version 3.1

cnrm_cm3 Météo-France / Centre National de Recherches Météorologiques, France - Climate Model Version 3.0

csiro_mk3_0 Commonwealth Scientific and Industrial Research Organisation, Australia - MK Climate Model Version 3.0

gfdl_cm2_0 Geophysical Fluid Dynamics Laboratory, USA - Climate Model Version 2.0

gfdl_cm2_1 Geophysical Fluid Dynamics Laboratory, USA - Climate Model Version 2.1

giss_model_e_r NASA / Goddard Institute for Space Studies, USA - ER Climate Model

inmcm3_0 Institute for Numerical Mathematics, Russia - Climate Model Version 3.0

ipsl_cm4 Institut Pierre Simon Laplace (IPSL), France - Climate Model Version 4.0

miroc3_2 Center for Climate System Research/National Institute for Environmental Studies/Frontier Research Center for Global Change, Japan - Climate Model Version 3.2

miub_echo_g Meteorological Institute of the University of Bonn, Meteorological Research Institute of KMA, and Model and Data group, Germany/Korea - Coupled climate model ECHO-G

mpi_echam_5 Max Planck Institute for Meteorology, Germany - ECMWF Hamburg Climate Model Version 5.0

mri_cgcm2_3_2a Meteorological Research Institute, Japan - Coupled General Circulation Model Version 2.3.2a

ncar_ccsm3 National Center for Atmospheric Research, USA - Community Climate System Model version 3.0

ncar_pcm1 National Center for Atmospheric Research, USA - Parallel Climate Model Version 1.0

ukmo_hadcm3 Met Office, United Kingdom - Hadley Centre Coupled Model, version 3.0

ukmo_hadgem1 Met Office, United Kingdom - Hadley Centre Global Environmental Model Version 1.0

8. Visualización de los cortes anuales elegidos

Para un mejor entendimiento de los datos, en la figuras 40 y 41 se muestra el aumento de temperatura, desde el período de referencia, predicho por cada modelo en nuestra región de interés (Sudamérica). Se toma como valor medio el aumento a 2 grados.

También se realizó una distribución tomando en cuenta los valores de los cortes de los modelos: para cada punto de grilla de Sudamérica, la figura 42 destaca el valor mínimo, 25%, promedio, 75% y máximo, así como la desviación estándar de los números. En la misma se observa que existe un mayor calentamiento en la zona tropical que en la polar. De hecho, la primera se calienta más que el promedio global, y la segunda se encuentra con valores por debajo de dicho promedio. Con respecto a la dispersión del ensamble, es del orden de los 0.4 grados; siendo el Amazonas la zona de mayor incertidumbre, tal vez porque no se tienen muchos datos de ese lugar.

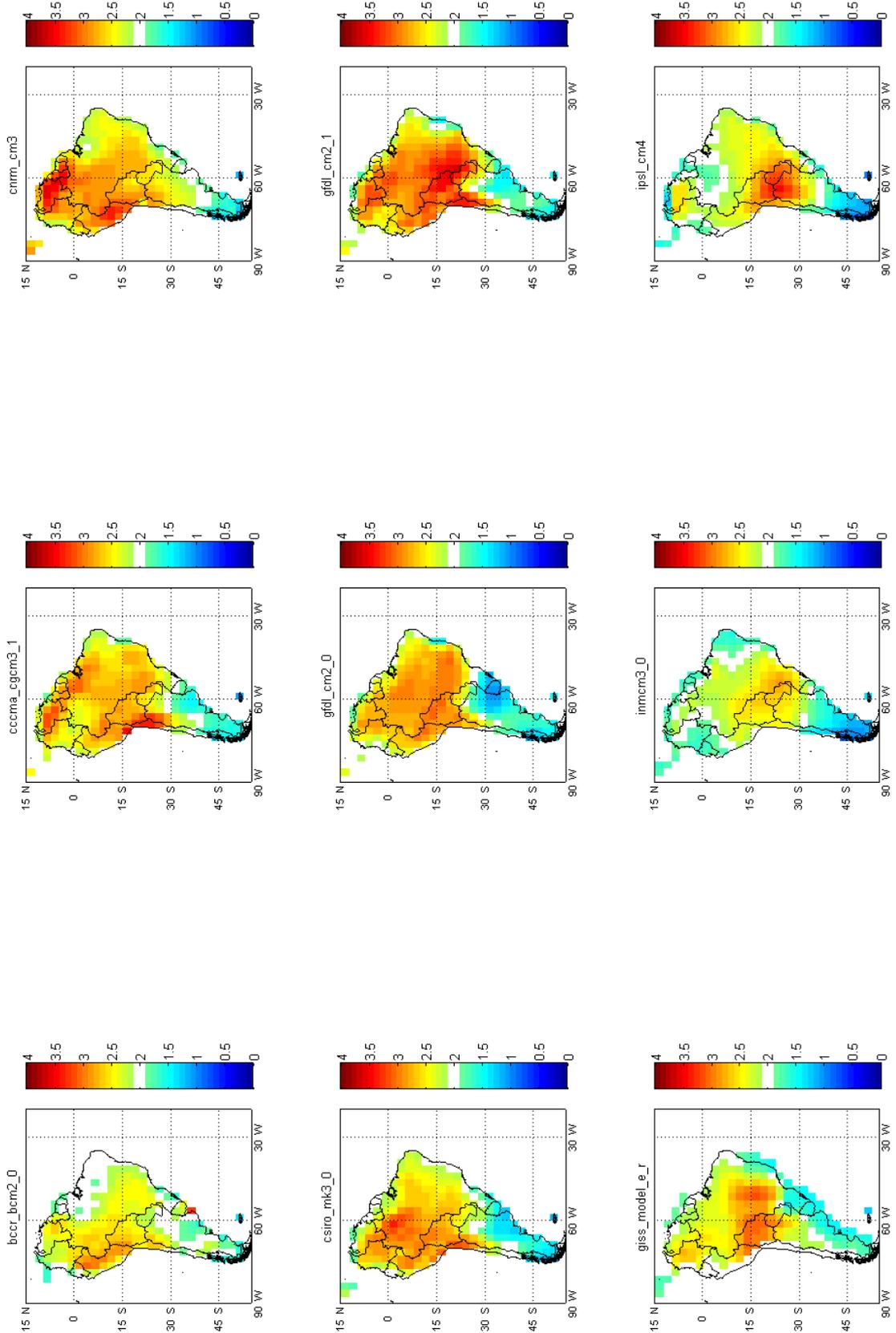


Figura 40 - Cortes de los modelos cuando predicen un aumento global a 2 grados

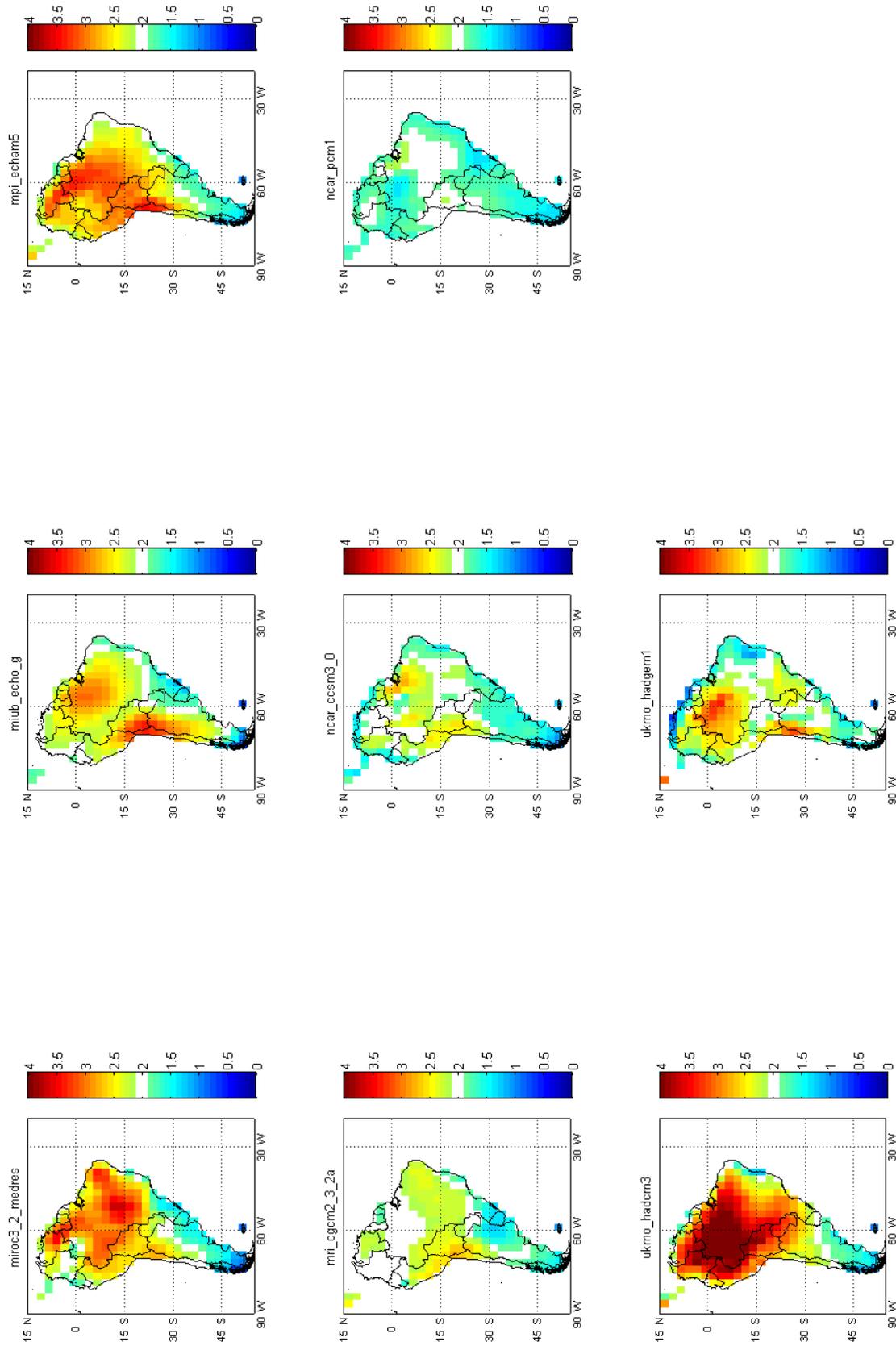


Figura 41 - Cortes de los modelos cuando predicen un aumento global a 2 grados (continuación)

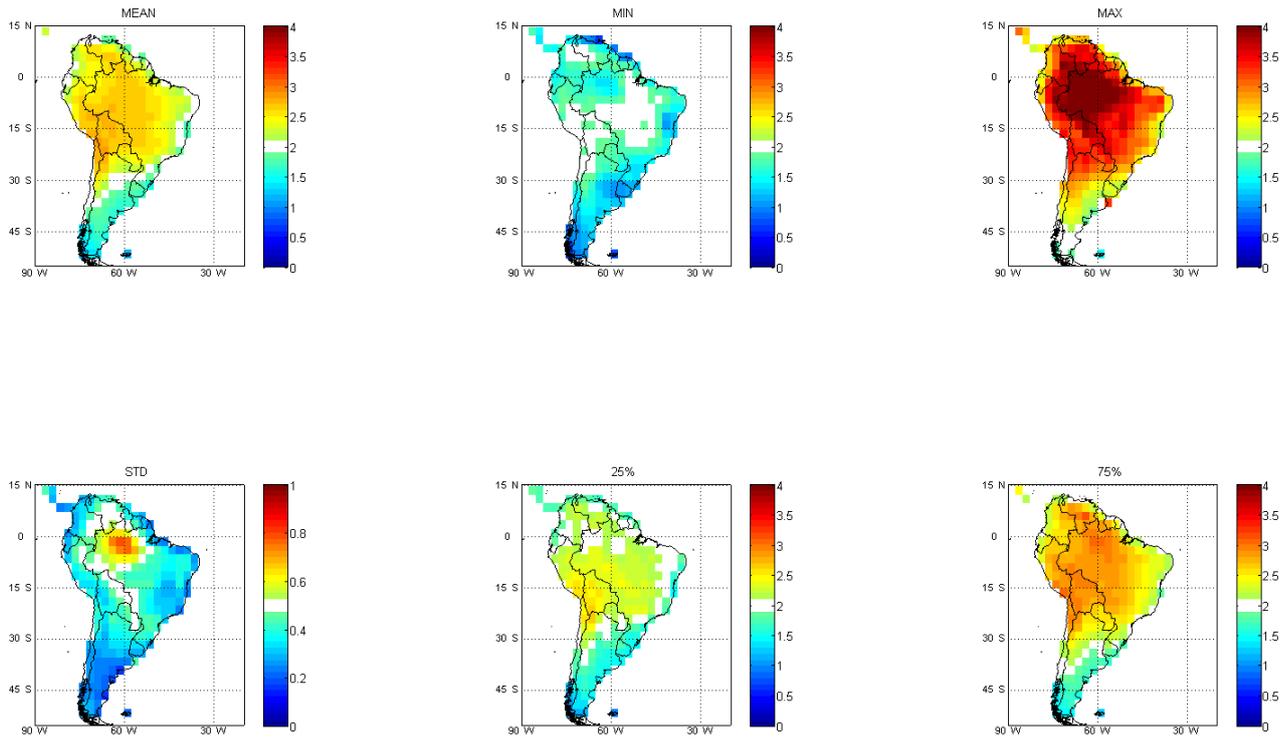


Figura 42 - puntos de interés en la distribución generada por los cortes de los modelos

9. Otros resultados de la clasificación

En las figuras 43 a 48 se observan los mapas para las clasificaciones con cortes de 6 a 11 clusters.

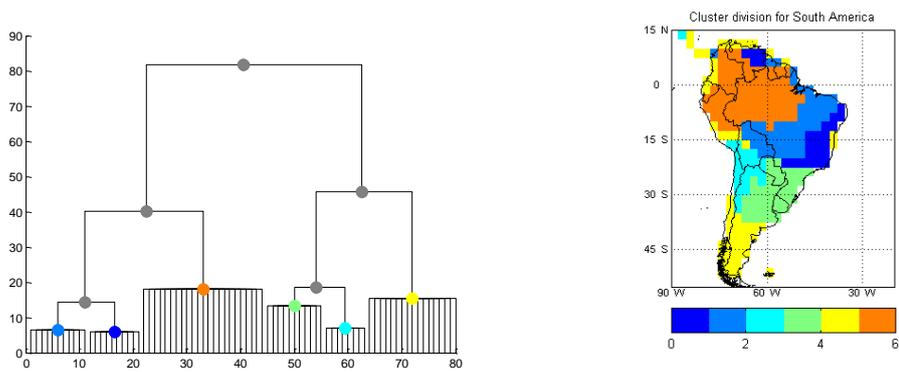


Figura 43 - Clasificación a 6 clusters

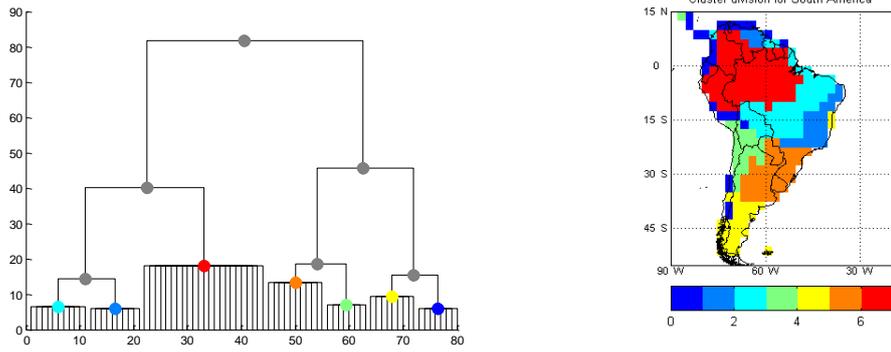


Figura 44 - Clasificación a 7 clusters

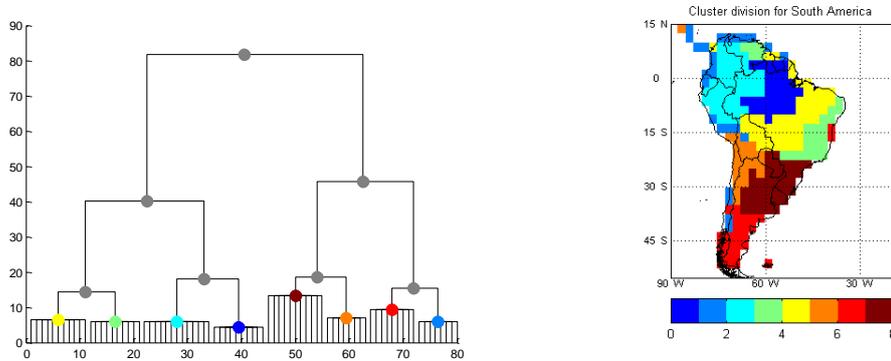


Figura 45 - Clasificación a 8 clusters

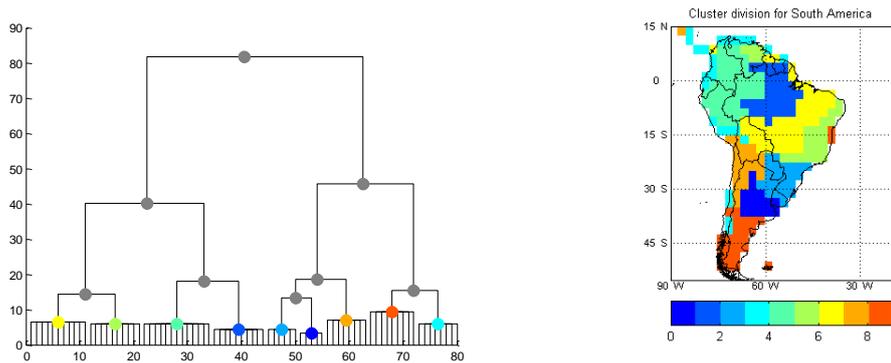


Figura 46 - Clasificación a 9 clusters

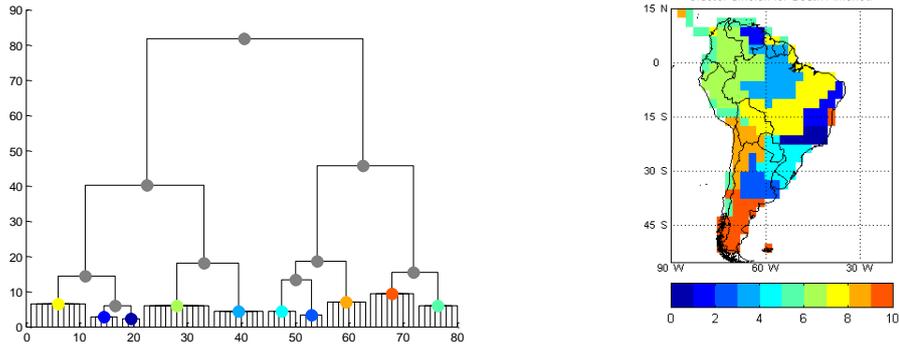


Figura 47 - Clasificación a 10 clusters

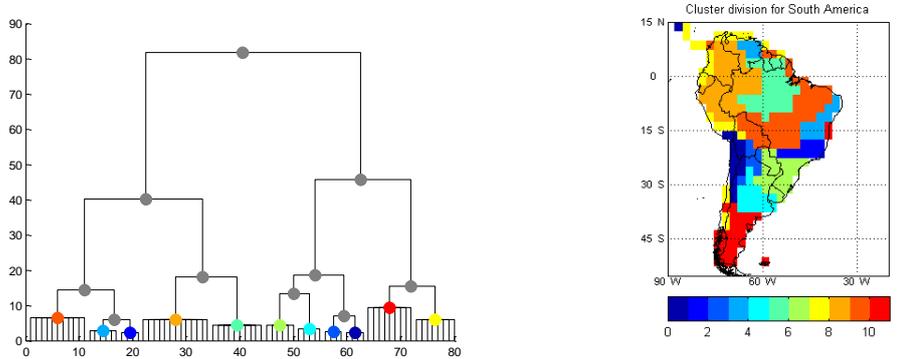


Figura 48 - Clasificación a 11 clusters