

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/283308087>

# Pronóstico de lluvia en un punto desde diversos puntos geográficos de observación mediante Procesos Gaussianos

Conference Paper · June 2011

CITATIONS

2

READS

116

3 authors:



**Julian Pucheta**

National University of Cordoba, Argentina

59 PUBLICATIONS 228 CITATIONS

[SEE PROFILE](#)



**Josef Baumgartner**

National University of Cordoba, Argentina

12 PUBLICATIONS 54 CITATIONS

[SEE PROFILE](#)



**Cristian Rodriguez Rivero**

University of Amsterdam

50 PUBLICATIONS 139 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Análisis, modelado y diseño de sistemas para controlar procesos de dinámica lenta [View project](#)



Project of Intelligent and autonomous Navigation aerial Systems [View project](#)

# **PRONÓSTICO DE LLUVIA EN UN PUNTO DESDE DIVERSOS PUNTOS GEOGRÁFICOS DE OBSERVACIÓN MEDIANTE PROCESOS GAUSSIANOS**

**Josef BAUMGARTNER<sup>1</sup>, Cristian RODRIGUEZ RIVERO<sup>1</sup>, Julián PUCHETA<sup>1</sup>**

<sup>1</sup>Laboratorio de Investigación Matemática Aplicada a Control, Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba  
Vélez Sarsfield 1611, X5016GCA Córdoba, Argentina. Tel/Fax 0351-4334147 int. 107.  
josef.s.baumgartner@gmail.com

## **RESUMEN**

El objetivo de este trabajo es generar modelos de comportamiento espacial para el pronóstico de fenómenos meteorológicos, particularmente lluvia mensual acumulada.

Se propone un esquema para la predicción de series temporales que, además de emplear los valores de la serie del punto geográfico de interés, incorpora mediciones de lluvia de puntos geográficos distribuidos espacialmente en una determinada región. Se plantea el problema dividiéndolo en dos etapas. Primero, se busca identificar la dinámica espacial de lluvia con respecto a un punto definido. Segundo, se ajusta el algoritmo a los fines de pronosticar la lluvia para el punto de interés usando la dinámica espacial obtenida.

Para encontrar la relación entre los puntos de observación se usan series temporales de lluvia de cada punto y procesos gaussianos cuyas propiedades están adaptadas con un algoritmo genético; el non-dominated sorting genetic algorithm II (NSGA-II). Analizando el Pareto-Front obtenido por el NSGA-II se determina la influencia de cada punto de observación hacia el punto de interés. Luego se usa un proceso gaussiano con propiedades óptimas para hacer pronósticos puntuales de lluvia donde las series temporales de lluvia representan mediciones en varios puntos de Alemania.

El algoritmo propuesto se evalúa con distintos escenarios, cada uno con una cierta cantidad de puntos de observación y distintas relaciones entre las observaciones y el punto de interés. De esa manera se muestra la efectividad del algoritmo de detectar relaciones dinámicas entre los puntos de observación y un punto de interés. Además, se comparan los pronósticos obtenidos con resultados anteriores. En este paso se observa que el algoritmo propuesto da resultados prometedores salvo cuando la correlación entre los puntos de observación y el punto de interés sea muy débil.

**Palabras Claves:** Pronósticos de series temporales, procesos gaussianos, algoritmo genético, NSGA-II, aprendizaje automático.

## INTRODUCCION

Desde los principios de la agricultura la predicción de la lluvia es un tema importante para asegurar y mejorar el rendimiento de un campo (Wei et al., 2005). Como consecuencia se desarrollaron complejos modelos matemáticos en las últimas décadas para hacer pronósticos de lluvia. Por un lado, se usan modelos físicos que permiten pronósticos a corto plazo (Collins et al., 2004). Por el otro, existen modelos autorregresivos que se basan en datos históricos de lluvia (Chow et al., 1996; Pucheta et al., 2007). Estos modelos, como redes neuronales, son capaces de predecir la lluvia acumulada en un cierto tiempo a largo plazo (Pucheta et al., 2009). La desventaja de esos modelos es que generalmente se utilizan datos de lluvia obtenidos de un punto de observación. Ese esquema es conocido como predicción de series temporales (Santana, 2007).

En este trabajo se propone un algoritmo novedoso que predice la lluvia en un punto teniendo en cuenta los datos de lluvia del punto de interés además de varios puntos alrededor, utilizando así más información para predecir valores de una serie. Suponemos de esta manera que se logra mejorar los pronósticos de lluvia. Aunque ese esquema fue estudiado antes (Liu et al., 1999), se presenta un nuevo algoritmo que usa un modelo no-paramétrico y que es ensayado con problemas más amplios.

### **Importancia del problema**

Para poder alimentar la creciente población mundial es necesario aplicar métodos de control óptimo para aumentar el rendimiento de un campo y para evitar el ineficiente uso de agua (Cassman, 1999). En ese contexto es fundamental contar con pronósticos de lluvia de alta exactitud ya que ésta es uno de los factores más impactantes en el control de cultivos (Feliu-Batlle et al., 2009; Patiño et al., 2007).

### **Datos usados**

Para aplicar el algoritmo propuesto se usan series de lluvia acumulada mensual de varios puntos de observación en Alemania. Existen varias razones porque elegir esos datos. Por un lado hay series temporales de 46 estaciones distribuidas por toda Alemania, o sea existe la posibilidad de elegir entre varias estaciones y de tal manera genera un entorno útil. Por el otro, se proveen datos de lluvia de cada día y acumuladas diarios y mensualmente en los últimos 20 años. Todos esos datos de lluvia están disponibles para su uso<sup>1</sup>.

## FORMULACION DEL PROBLEMA

Este trabajo trata con una extensión del clásico problema de predecir futuros estados de una serie temporal. En el caso común se ajusta un modelo autoregresivo con los valores conocidos de una serie dada (Velásquez 2004). Luego se pueden hacer predicciones con el modelo ajustado realimentándolo en cada paso con valores anteriores.

El esquema de este trabajo consiste en varias series temporales que pueden ser usadas para ajustar un modelo autoregresivo con el fin de predecir futuros valores de una de las series dadas. En la figura 1 se muestra el esquema descripto. Se supone que los estados de cada serie están distribuidos uniformemente sobre el tiempo. Entonces, los estados de cada serie se

---

<sup>1</sup><http://www.dwd.de>

pueden expresar como  $x_i(t-T)$ ,  $x_i(t-2T)$ , ...,  $x_i(t-mT)$  donde  $i$  es el índice de la serie y  $T$  es el tiempo de muestreo.

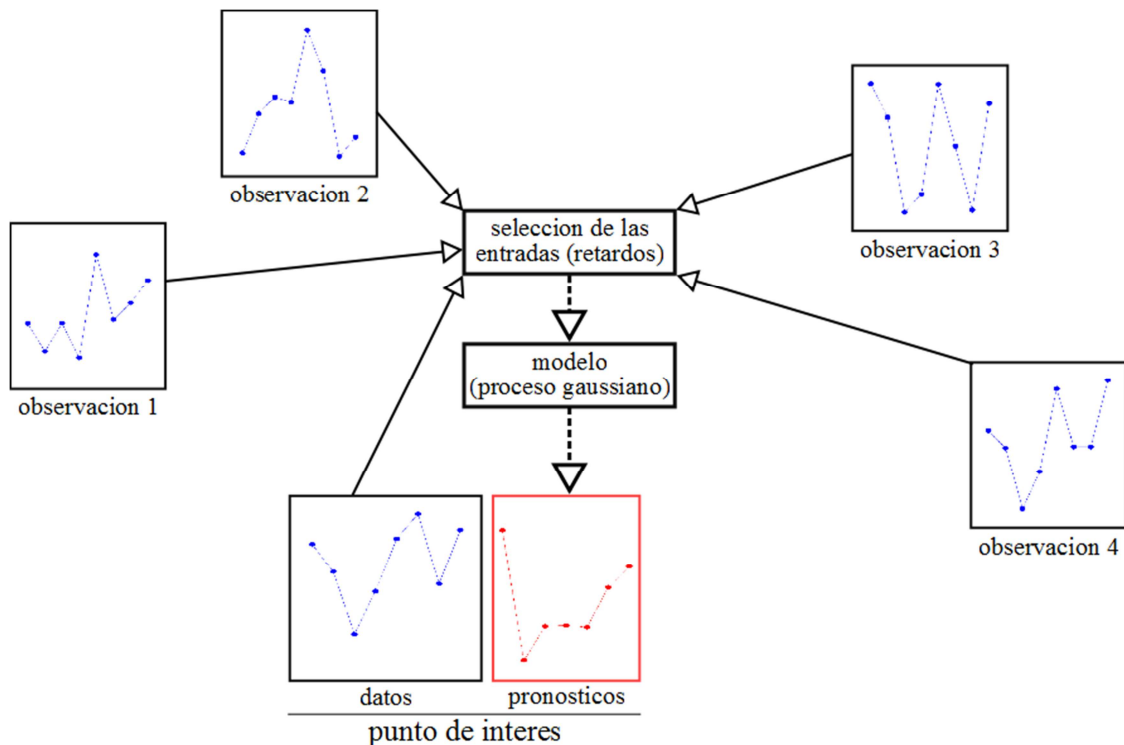


Figura 1. Esquema de predicciones con varios puntos de observación. Cada punto tiene una cierta influencia sobre el punto de interés que se expresa en los retardos aplicados. El conjunto de todos los retardos se puede interpretar como una dinámica espacial.

### **Objetivos principales**

Igual que en el caso común de solamente una serie temporal la pregunta principal para ajustar un modelo autoregresivo es: ¿Cómo se puede obtener la mayor cantidad de información útil de los datos dados? En el caso de varias series temporales se agrega la pregunta: ¿Qué información de cada serie sirve para predecir la serie de interés? Entonces se plantean los siguientes problemas que se afectan entre ellos: ¿Cómo elegir las entradas del modelo autoregresivo y que grado de libertad tiene que tener el modelo para que pueda representar la dinámica buscada?

### **Escenarios de investigación**

Para probar el algoritmo propuesto se estudian tres escenarios con distintas características. En cada escenario se usan mediciones realizadas en distintos puntos de Alemania cuyas ubicaciones se muestran en la figura 2. Dos de los puntos usados tienen una ubicación especial dentro de Alemania. El punto “Sylt” es el punto más norteño de Alemania mientras el punto “Zugspitze” está ubicado a 2962 metros sobre el nivel de mar que lo convierte en el punto más alto de Alemania. El punto de interés siempre es Stuttgart que se encuentra en el suroeste de Alemania.

Cada escenario incluye una cierta cantidad de series temporales donde los valores de una serie corresponden a la lluvia acumulada mensual en un punto de observación. Se conocen los

últimos 102 valores de cada serie – en otras palabras se usan datos históricos de los últimos ocho años y medio.



Figura 2. Mapa de Alemania con los puntos de observación.

El primer escenario consiste de ocho series temporales. Se usan los datos de las siguientes estaciones de medición: *Augsburg*, *Frankfurt*, *Konstanz*, *Nuremberg*, *Rheinstetten*, *Stuttgart*, *Sylt*, *Zugspitze*. En este escenario hay más datos históricos disponibles que en los otros escenarios. Luego se espera que en este caso las predicciones sean las más exactas.

En el segundo escenario se usan solamente cinco series temporales. En comparación con el primer escenario se sacan los puntos *Augsburg*, *Sylt* y *Zugspitze*. Mientras *Augsburg* es uno de los puntos de observación más cercanos de *Stuttgart* los otros dos puntos son especiales por su ubicación como se menciona arriba.

Para el tercer escenario se emplean los datos de *Berlin*, *Duesseldorf* y obviamente *Stuttgart*, el punto de interés. *Berlin* está ubicado a 510 km de *Stuttgart* hacia el noreste y *Duesseldorf* se encuentra a unos 340 km de *Stuttgart* en el noroeste de Alemania.

Para evaluar los tres escenarios y luego el algoritmo propuesto se comparan los resultados obtenidos con una serie de referencia. Esa serie se obtiene prediciendo futuros valores para *Stuttgart* usando una red neuronal adaptada con la ley de Levenberg-Marquardt (Levenberg, 1944).

## EL ALGORITMO PROPUESTO

El objetivo principal de este trabajo es la predicción de lluvia para un punto de interés con un modelo autoregresivo. Para lograr este objetivo el primer paso del algoritmo propuesto es

determinar las entradas del modelo. Luego se puede ajustar el modelo autoregresivo con los datos acomodados según las entradas encontradas.

En general el tipo y la estructura del modelo usado pueden influir en la selección de entradas. Entonces es importante encontrar no solamente un conjunto de entradas sino también un modelo con una estructura adecuada. Por ejemplo en el caso de redes neuronales un conjunto de entradas puede ser óptimo para dos capas ocultas y otro conjunto para una capa oculta con más patrones. Por lo tanto, se utilizan procesos gaussianos para simplificar el problema de elegir la flexibilidad, o sea, la cantidad de parámetros del modelo.

### **Procesos Gaussianos**

Muchos trabajos sobre la predicción de series temporales usan redes neuronales con una cierta ley de ajuste, sin embargo se aplican procesos gaussianos en este caso por las razones mencionadas precedentemente. A diferencia de las redes neuronales, los procesos gaussianos son modelos bayesianos, no-paramétricos basados en la siguiente teoría:

Se supone que los datos usados  $y(x)$  resultan de una función  $f(x)$  que se observa a través de ruido gaussiano  $\varepsilon \sim N(0, \sigma^2)$ :

$$y(x) = f(x) + \varepsilon \quad (1)$$

Entonces la verosimilitud se puede escribir como una distribución gaussiana sobre funciones:

$$p(y|f) \sim N(f, \sigma^2 I) \quad (2)$$

Luego se define una probabilidad a priori de  $f(x)$  dado las entradas  $X$ , donde  $K(X, X)$  es la matriz de covarianza, suponiendo que los datos tienen media cero:

$$p(f|X) \sim N(0, K(X, X)) \quad (3)$$

Si los datos no tienen media cero se puede aplicar una simple transformación restando la media, o sea el requerimiento de media cero no es ninguna restricción. La matriz de covarianza  $K(X, X)$  resulta de una función de covarianza como por ejemplo:

$$K(x_p, x_q) = cov(x_p, x_q) = c * \exp\left(-\frac{1}{2} \frac{(x_p - x_q)^2}{\ell}\right) \quad (4)$$

Donde  $c$  y  $\ell$  son hiperparámetros que definen la función de covarianza. En este trabajo se estudian varias funciones de covarianza que cumplen con dos características importantes: La resultante matriz  $K(X, X)$  es simétrica y definida positiva para cualquier conjunto de entradas  $X$ .

Con las distribuciones gaussianas de (2) y (3) se puede integrar sobre  $f$  obteniendo la marginal verosimilitud, que también es una distribución gaussiana:

$$p(y|X, \theta) \sim N(0, K(X, X) + \sigma^2 I) \quad (5)$$

Para ajustar la matriz de covarianza  $K(X, X)$  – o, más exacto los hiperparámetros – a los datos dados se busca maximizar la marginal verosimilitud de (5). En general se usa el logaritmo de (5) para la maximización que resulta en:

$$\log p(y|X, \theta) = -\frac{1}{2} \log |K(X, X) + \sigma^2 I| - \frac{1}{2} y' (K(X, X) + \sigma^2 I)^{-1} y - \frac{n}{2} \log(2\pi) \quad (6)$$

Para maximizar esta función se usa el método de gradiente que fue propuesto por Rasmussen et al., 2006.

Después de haber encontrado los hiperparámetros se busca hacer pronósticos para una nueva entrada  $X^*$ . La distribución de las observaciones y la nueva entrada es:

$$\begin{bmatrix} y \\ f^* \end{bmatrix} \sim N \left( 0, \begin{bmatrix} K(X, X) + \sigma^2 I & K(X, X^*) \\ K(X^*, X) & K(X^*, X^*) \end{bmatrix} \right) \quad (7)$$

Donde  $K(X^*, X)$  es la covarianza de la nueva entrada  $X^*$  con las entradas  $X$  y  $K(X^*, X^*)$  es la varianza de la nueva entrada. Entonces la probabilidad condicional de  $f^*$  dado los datos de observación y la nueva entrada  $X^*$  es:

$$p(f^*|X, y, X^*) \sim N(\mu, \sigma_*^2) \quad (8)$$

Con:

$$\mu = K(X^*, X) [K(X, X) + \sigma^2 I]^{-1} y \quad (9)$$

$$\sigma_*^2 = K(X^*, X^*) - K(X^*, X) [K(X, X) + \sigma^2 I]^{-1} K(X, X^*) \quad (10)$$

El mejor pronóstico para la nueva entrada  $X^*$  es obviamente la media  $\mu$  de (9). Además se puede calcular con (10) la varianza  $\sigma_*^2$  de la predicción que es una medida útil al momento de evaluar la incertidumbre de los pronósticos.

Recapitulando la teoría de procesos gaussianos se puede observar que, del punto de vista de un usuario, el único grado de libertad de un proceso gaussiano es la función de covarianza (Schölkopf et al., 2001). En este trabajo se aplican doce funciones distintas. Sin ir al detalle las funciones usadas son: una función lineal, una función exponencial cuadrada, una función racional cuadrada, una función arcoseno y una función Matérn con  $\nu=3/2$  y  $\nu=5/2$  (Matérn 1960) donde cada una de esas funciones esta usada con y sin una función de covarianza ruido. Una descripción de estas funciones de covarianza se encuentra en Rasmussen et al., 2006.

### **Selección de entradas y de la función de covarianza**

Después de haber determinado el tipo del modelo autoregresivo hay que encontrar un conjunto de entradas y una función de covarianza para poder ajustar el modelo. Poniendo restricciones como una máxima cantidad de 20 entradas y un retraso máximo de (t-30) aún quedan más de  $5,1 \cdot 10^{25}$  posibles combinaciones en el caso de cinco puntos de observación. Obviamente es imposible comprobar cada una de esas combinaciones si sirve para generar un modelo de alta exactitud. Por lo tanto se aplica un algoritmo genético que trata de encontrar una combinación optima con tan solo 120.000 evaluaciones.

El algoritmo que se implementa es el Non-dominated Sorting Genetic Algorithm II (NSGA-II) (Deb et al., 2002), uno de los algoritmos genéticos más usados (Correa Florez et al., 2008). El NSGA-II evalúa una generación de individuos y luego genera una nueva generación vía recombinación y mutación usando los individuos más exitosos. Una ventaja del NSGA-II es

que permite optimizar varios criterios. Aprovechando esa oportunidad se aplica el NSGA-II minimizando no solo el error de predicción sino también la cantidad de entradas. Esos dos criterios son elegidos para obtener un modelo que por más que tenga pocas entradas es capaz de hacer buenas predicciones.

Para calcular el error de predicción es necesario ajustar un proceso gaussiano con una parte de los datos conocidos y luego predecir la otra parte de los datos conocidos. Después se puede calcular el error de la raíz cuadrada media (ERCM) entre los datos originales y los datos predichos de la siguiente forma:

$$ERCM = \sqrt{\frac{\sum_{i=1}^n (y_{real(i)} - y_{pred(i)})^2}{n}} \quad (11)$$

La desventaja de esta evaluación es que en una computadora común tarda aproximadamente un cuarto segundo de ajustar y evaluar un proceso gaussiano. Por lo que 120.000 evaluaciones tardan más de diez horas.

Ya que el NSGA-II es un algoritmo de optimización de múltiples criterios no va a encontrar una sola solución sino varias soluciones que representan un compromiso entre todos los criterios usados. En ese contexto el conjunto de soluciones que superan las otras soluciones en por lo menos un criterio se llama Pareto Front.

Los parámetros importantes del NSGA-II son la probabilidad de mutación, el tamaño de cada generación y la cantidad de generaciones. En este trabajo una mutación ocurre con una probabilidad de 20% que, en comparación con otros trabajos, es un valor bastante alto. El motivo para elegir una probabilidad de mutación alta es que se trata de evitar que el algoritmo no pase por ciertas regiones teniendo en cuenta que el espacio de búsqueda es  $4,2 \cdot 10^{20}$  veces más grande que la cantidad de evaluaciones. En la figura 3 se puede observar que el NSGA-II converge asintóticamente lo que fortalece la selección de la tasa de mutación. El tamaño de cada generación es 300. Cada individuo evoluciona por 400 generaciones, luego son  $300 \cdot 400 = 120.000$  evaluaciones en total.

## RESULTADOS PRINCIPALES

El algoritmo propuesto consiste en dos partes. Primero, se aplica un algoritmo genético para determinar las entradas de un modelo autoregresivo implementado como un proceso gaussiano. Luego, se hacen pronósticos de lluvia para un punto de interés. La calidad de esos pronósticos depende fuertemente de las entradas encontradas por el algoritmo genético, ya que es necesario estudiar la confiabilidad de los resultados del NSGA-II antes de ajustar un proceso gaussiano. Teniendo en cuenta que cada entrada seleccionada representa un cierto retraso de un determinado punto de observación, el conjunto de todas las entradas se puede interpretar como una dinámica espacial de todos los puntos de observación hacia el punto de interés.

### Determinación de la dinámica espacial

Para estimar la confiabilidad de los resultados del algoritmo genético implementado la herramienta más importante es la evaluación de varias corridas. Esto quiere decir que el algoritmo empieza varias veces con condiciones iniciales aleatorias. Si en cada corrida se



encuentra una solución parecida es un indicio que las características del algoritmo, como la cantidad de generaciones etc., fueron elegidas adecuadamente al problema de optimización.

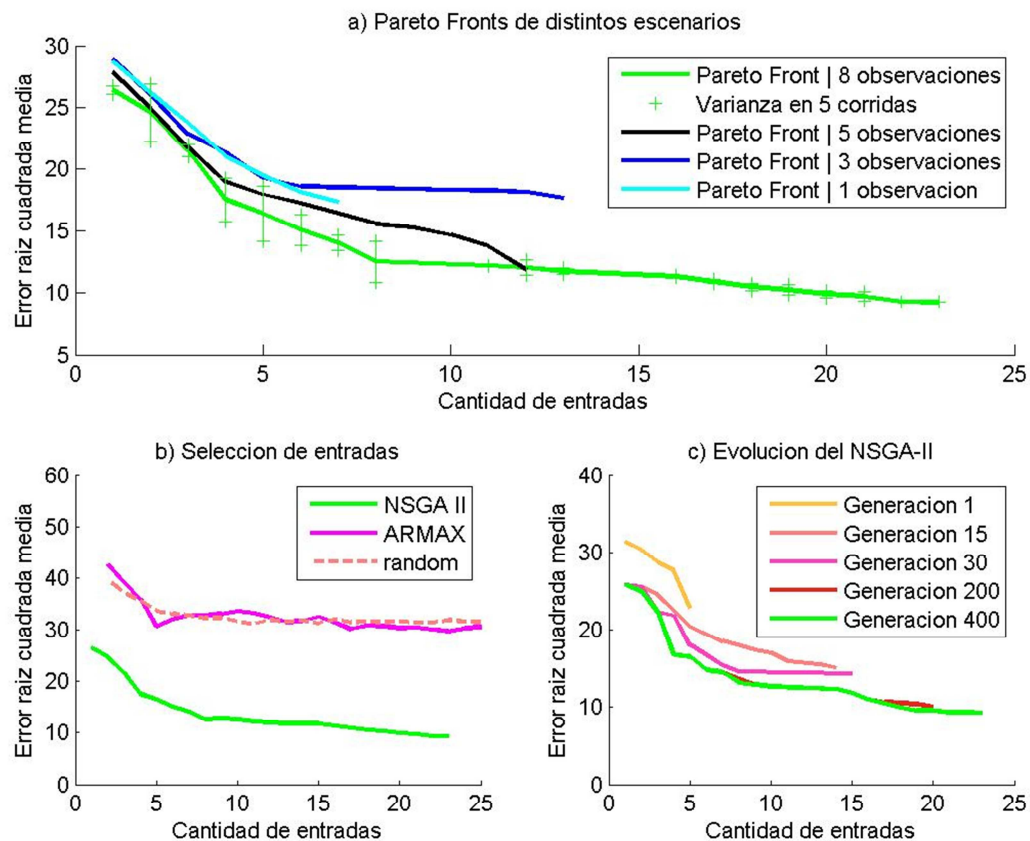


Figura 3. Evaluación del NSGA-II y de los Pareto Fronts encontrados. a) Se muestran los Pareto Fronts de distintos escenarios obtenidos en cinco corridas y la correspondiente varianza para el escenario de ocho observaciones. b) Se comparan distintos métodos para determinar las entradas del proceso gaussiano – con el NSGA-II, con un ARMAX cuyos coeficientes marcan las entradas más importantes o al azar. c) Se muestra cómo cambia el Pareto Front mientras avanza el NSGA-II. Observe que el Pareto Front converge asintóticamente y que cubre cada vez más el espacio de entradas. Después de 200 generaciones casi no hay cambios en el Pareto Front, lo que es un buen indicio de que la cantidad de entradas fue lo suficientemente grande elegida. El Pareto Front verde de c) no coincide exactamente con los Pareto Fronts verdes de a) y b) porque ahí se graficaron los mejores Pareto Fronts de cinco corridas mientras en c) se muestra solamente el Pareto Front de la primera corrida.

En este trabajo se realizaron cinco corridas para cada escenario. En la figura 3 se muestran los Pareto Fronts promedios de las cinco corridas de los tres escenarios. Además, se grafica la varianza para el caso de ocho observaciones. Se observa que la varianza es aceptable para pocas entradas y muy chica para muchas entradas. Las varianzas para los otros escenarios son notablemente más chicas debido a que, en estos casos, el espacio de búsqueda es más chico. En la misma figura se comparan los resultados del NSGA-II con una selección de las entradas

con un filtro lineal y una selección al azar. Finalmente se muestra cómo avanza el NSGA-II mientras crece la cantidad de generaciones terminadas. En resumen se puede concluir que el NSGA-II fue corrido con una suficiente cantidad de evaluaciones.

Luego se pueden evaluar las entradas de cada elemento de todas las Pareto Fronts. Contando las veces que aparece un retraso en una entrada se obtiene un mapa que indica la dinámica espacial del esquema estudiado. Hay que tener en cuenta que, el tamaño del espacio de búsqueda obstaculiza que el NSGA-II encuentre entradas más homogéneas. En la figura 4 se muestra el mapa de la dinámica espacial encontrada para el escenario de ocho puntos de interés.

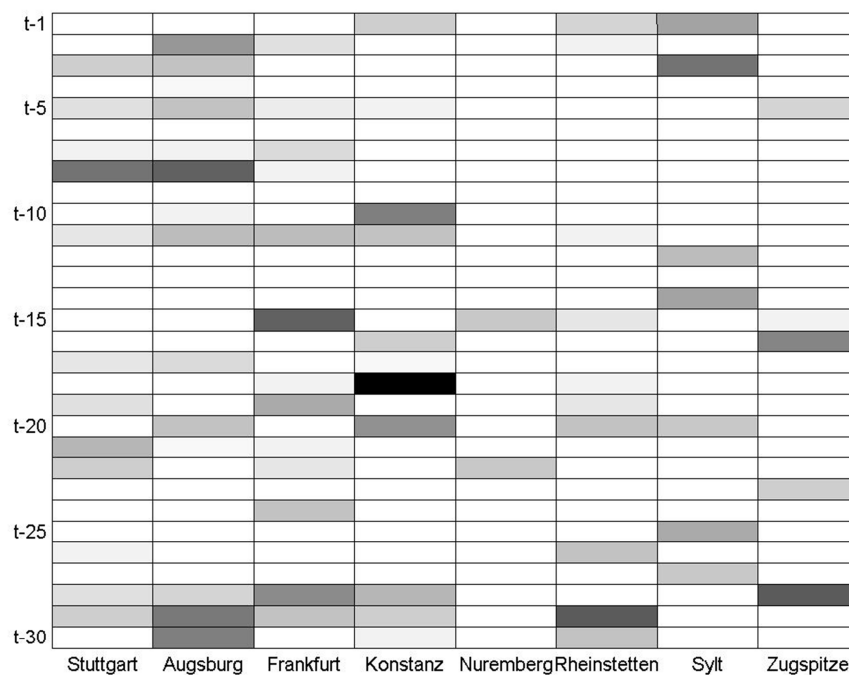


Figura 4. Mapa de la dinámica espacial para el escenario con ocho puntos de observación. Las entradas más oscuras son las más importantes para predecir valores de la serie *Stuttgart* según los resultados del NSGA-II. Mientras el punto Nuremberg no tiene mucho impacto sobre el punto de interés, las mediciones en Augsburg, Frankfurt y Konstanz por ejemplo son más importantes.

### **Pronósticos de lluvia**

Después de haber determinado la dinámica espacial se ajusta un modelo autoregresivo realizado como un proceso gaussiano. Antes de correr el algoritmo de ajuste es necesario definir las propiedades del proceso gaussiano como las entradas y la función de covarianza. Usando los resultados del NSGA-II se pueden determinar las propiedades de un individuo de la Pareto Front pero entonces, surge la pregunta: ¿Qué individuo elegir de la Pareto Front?

Es conveniente elegir a un individuo de la Pareto Front que representa un compromiso entre los dos criterios aplicados, dando más peso al ERCM que a la cantidad de entradas. Esto es así ya que la cantidad de entradas sirve sobre todo para evitar el efecto de sobreajuste en el

momento de encontrar las entradas del proceso gaussiano. Así es como se obtienen las propiedades mostradas en las tablas 1, 2 y 3 para los tres escenarios estudiados.

Tabla 1. Propiedades del proceso gaussiano para ocho puntos de observación.

Función de covarianza	arcoseno
Stuttgart	t-5, t-8, t-19, t-28
Augsburg	t-5, t-20, t-28
Frankfurt	t-5, t-11, t-15, t-18, t-29
Konstanz	t-18
Nuremberg	-
Rheinstetten	-
Sylt	-
Zugspitze	t-16, t-28

Tabla 2. Propiedades del proceso gaussiano para cinco puntos de observación.

Función de covarianza	arcoseno
Stuttgart	t-7, t-8, t-18, t-28
Frankfurt	t-1, t-5, t-7, t-19, t-30
Konstanz	t-17, t-20
Nuremberg	-
Rheinstetten	-

Tabla 3. Propiedades del proceso gaussiano para tres puntos de observación.

Función de covarianza	arcoseno
Stuttgart	t-1, t-3, t-7, t-10, t-17, t-18, t-19, t-20, t-22, t-28, t-29
Berlin	-
Duesseldorf	t-2, t-8, t-20, t-28

En la tabla 4 se muestran los errores de la raíz cuadrada media de los tres escenarios investigados para los pronósticos del valor 103 hasta el valor 120. Además se evalúa el mismo error para los dos escenarios de referencia.

Tabla 4. Error de la raíz cuadrada media para los escenarios estudiados y dos señales de referencia.

ERCM del escenario con 8 puntos de observación:	30.5657
ERCM del escenario con 5 puntos de observación:	25.3173
ERCM del escenario con 3 puntos de observación:	34.0023
ERCM del escenario de referencia con 1 punto de observación:	38.3490
ERCM de una red neuronal ajustada según (Rodríguez et al., 2010):	47.2301

Luego se muestran las predicciones para los escenarios con ocho y cinco puntos de observación en la figura 5. También se grafican los pronósticos de referencia que resultan de una red neuronal con una ley de ajuste (Rodríguez et al., 2010) utilizando solamente los datos de Stuttgart.

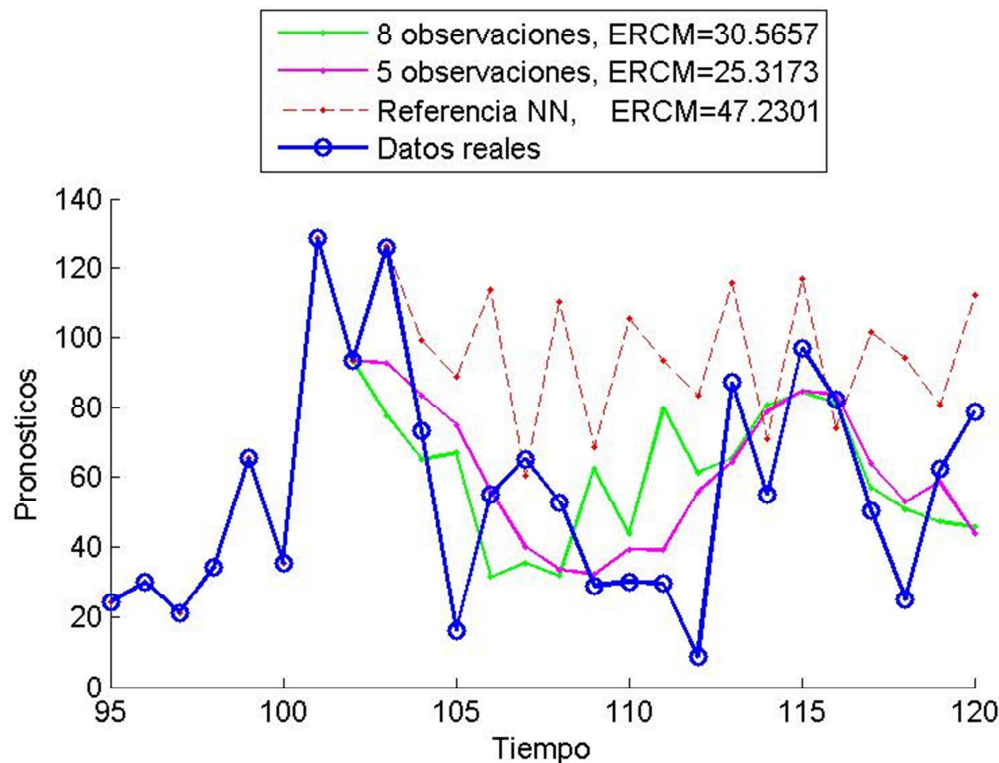


Figura 5. Pronósticos de lluvia para Stuttgart. Los datos predichos con el algoritmo propuesto tienen un ERCM menor al de la red neuronal. El ERCM del escenario con cinco observaciones es más chico que el de ocho observaciones para este muestreo.

## DISCUSION

Los resultados obtenidos por el NSGA-II son coherentes en cuanto la función de covarianza. En los tres escenarios el NSGA-II encontró una función de covarianza de arcoseno. Ese resultado es un fuerte indicio de que el NSGA-II realmente encontró la mejor función de covarianza para los datos usados.

Los retrasos encontrados por el NSGA-II sugieren que las mediciones en algunos puntos como Augsburg, Frankfurt, Konstanz y obviamente el punto de interés Stuttgart son más importantes para los pronósticos de Stuttgart que los otros puntos de observación. Luego los primeros dos escenarios tienen una ventaja sobre el tercer escenario y los dos escenarios de referencia. Esa ventaja se expresa en el ERCM de los pronósticos realizados, como se ve en la tabla 4. También se puede notar que el tercer escenario con pocos puntos de observación requiere más retrasos del punto de interés que los otros dos escenarios. En este contexto, se muestra la efectividad del algoritmo de detectar relaciones dinámicas entre los puntos de observación y un punto de interés.

Según los Pareto Fronts del NSGA-II el escenario con ocho puntos de observación tiene un ERCM menor que el escenario con cinco puntos de observación. Sin embargo en los pronósticos realizados del valor 102 hasta el valor 120 la relación del ERCM es al revés. Para investigar esa observación sería necesario aplicar el algoritmo a varios conjuntos de datos para excluir la posibilidad de que se trata de una desviación en el muestreo dado.

## CONCLUSIONES

Se estudiaron tres escenarios con distintas características. El primer escenario consiste en ocho puntos de observación que están ubicados alrededor del punto de interés más un punto en el norte de Alemania. El segundo y tercer escenario usan menos puntos de observación, además, los puntos de observación del tercer escenario están ubicados más lejos del punto de interés.

A través de la evaluación de los resultados del NSGA-II y de los pronósticos realizados se mostró que el algoritmo propuesto es capaz de mejorar los pronósticos hechos por un modelo autoregresivo que utiliza solamente datos de un punto de observación. Especialmente en la comparación de los modelos gaussianos con redes neuronales se puede observar una ventaja del algoritmo propuesto para predecir datos climáticos como lluvia acumulada mensual.

## Bibliografía

- Cassman, K.** (1999) "*Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture*". En: Proc. Natl. Acad. Sci. Vol.96, pp. 5952–5959. USA.
- Chow, T. W. S. y Leung, C.T.** (1996) "*Neural network based short-term load forecasting using weather compensation*". En: IEEE Transactions on Power Systems, Vol.11, Iss.4, pp. 1736-1742.
- Collins, W.; Rasch, P.; Boville, B.; Hack, J.; McCaa, J.; Williamson, D.; Kiehl, J. y Briegleb, B.** (2004) "*Description of the NCAR Community Atmosphere Model (CAM 3.0)*". En: University Corporation for Atmospheric Research. Retrieved 2011-01-03.
- Correa Florez, C. A.; Bolaños, R. A. y Cabrera A. M.** (2008) "*Algoritmo Multiobjetivo NSGA-II Aplicado Al Problema De La Mochila*". En: Scientia et Technica Año XIV. Vol.39, pp. 206-211. Universidad Tecnológica de Pereira.
- Deb K.; Pratap A.; Agarwal S. y Meyarivan T.** (2002) "*A Fast and Elitist Multiobjective Genetic Algorithm*". En: IEEE Transactions on evolutionary computation. Vol.35, No.2.
- Feliu-Batlle, V.; Rivas-Perez, R. y Castillo-Garcia F. J.** (2009) "*Fractional order controller robust to time delay variations for water distribution in an irrigation main canal pool*". En: Comput. Electron. Agric. DOI. 10.1016/j.compag.200908.005
- Levenberg, K.** (1944) "*A Method for the Solution of Certain Problems in Least Squares*". En: Quart. Appl. Math. 2, pp. 164-168.
- Liu, J. N. K. y Lee, R. S. T.** (1999) "*Rainfall forecasting from multiple point sources using neural networks*". En: Proc. of the International Conference on Systems, Man, and Cybernetics. Vol.3, pp. 429-434.
- Matérn, B.** (1960) "*Spatial Variation*". En: Meddelanden från Statens Skogsforskningsinstitut, 49, No.5. Almaanna Forlaget, Stockholm. Second edition, Springer-Verlag, Berlin, pp. 85, 87, 89.
- Patiño, H. D.; Pucheta, J.; Schugurensky, C.; Fullana, R. y Kuchen, B.** (2007) "*Approximate Optimal Control-Based Neurocontroller with a State Observation System for Seedlings Growth in Greenhouse*". En: IEEE International Symposium on Approximate Dynamic Programming and Reinforcement Learning. pp. 318-323.
- Pucheta, J.; Patiño, H. D. y Kuchen, B.** (2007) "*Neural Networks-Based Time Series Prediction Using Long and Short Term Dependence in the Learning Process*". En: Proc. of the 2007 International Symposium on Forecasting, 24th to 27th of June 2007 Marriott Marquis Times Square, New York.

- Pucheta, J.; Patiño, D. y Kuchen, B.** (2009) “*A Statistically Dependent Approach For The Monthly Rainfall Forecast from One Point Observations*”. En: IFIP International Federation for Information Processing Volume 294, Computer and Computing Technologies in Agriculture II, Vol.2, eds. D. Li, Z. Chunjiang, (Boston: Springer), pp. 787–798.
- Rasmussen, C. E. y Williams C. K. I.** (2006) *Gaussian Processes for Machine Learning*. The MIT Press.
- Rodríguez, C.; Pucheta, J.; Baumgartner, J.; Patiño y H., Kuchen, B.** (2010) “*An Approach for Time Series Forecasting by simulating Stochastic Processes through time lagged feed-forward neural network*”. En: Proceedings of the 2010 International Conference on Data Mining, Las Vegas, USA. Vol.6, pp. 287-293.
- Santana, J. C.** (2007) “*Predicción de series temporales con redes neuronales: una aplicación a la inflación colombiana*”. En: Revista Colombiana de Estadística. Vol. 29, No.1, pp. 77-92.
- Schölkopf, B. y Smola, A.** (2001) *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond*. The MIT Press.
- Velásquez, J. D.** (2004) “*Pronóstico de la serie de Mackey Glass usando modelos de regresión no-lineal*”. En: Dyna, Revista De la Facultad de Minas – Universidad Nacional de Colombia – Sede Medellín. Vol.71, No.142, pp. 85-95.
- Wei, H.; Li, J. y Liang, T.** (2005) “*Study on the estimation of precipitation resources for rainwater harvesting agriculture in semi-arid land of China*”. En: Agricultural Water Management, Elsevier. Vol.71, Issue 1, pp. 33-45. DOI: 10.1016/j.agwat.2004.07.002.