



Ingeniare. Revista Chilena de Ingeniería

ISSN: 0718-3291

facing@uta.cl

Universidad de Tarapacá

Chile

Salini Calderón, Giovanni; Pérez Jara, Patricio
ESTUDIO DE SERIES TEMPORALES DE CONTAMINACIÓN AMBIENTAL MEDIANTE TÉCNICAS
DE REDES NEURONALES ARTIFICIALES
Ingeniare. Revista Chilena de Ingeniería, vol. 14, núm. 3, diciembre, 2006, pp. 284-290
Universidad de Tarapacá
Arica, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=77225398011>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

ESTUDIO DE SERIES TEMPORALES DE CONTAMINACIÓN AMBIENTAL MEDIANTE TÉCNICAS DE REDES NEURONALES ARTIFICIALES

TIME SERIES ANALYSIS OF ATMOSPHERE POLLUTION DATA USING ARTIFICIAL NEURAL NETWORKS TECHNIQUES

Giovanni Salini Calderón¹

Patricio Pérez Jara²

Recibido 24 de marzo de 2006, aceptado 3 de julio de 2006

Received: March 24, 2006 Accepted: July 3, 2006

RESUMEN

Se diseñó una red neuronal artificial (RNA) para hacer predicciones de valores de concentraciones horarias de material particulado fino en la atmósfera. El estudio está basado en los datos de tres años de series de tiempo de pm2.5 (material particulado suspendido de 2,5 micrones de diámetro), obtenidos en una estación céntrica de la Red MACAM de la ciudad de Santiago de Chile, entre los años 1994 y 1996. Para obtener el espaciamiento óptimo de los datos, así como el número de datos hacia atrás necesarios para pronosticar el valor futuro, se aplicaron dos test estándar usados en estudio de sistemas dinámicos, como *Información Mutua promedio* (AMI) y *Falsos Vecinos más Cercanos* (FNN). De esta manera se encontró que lo más conveniente era considerar como entrada los datos de PM2.5 cada seis horas durante un día (cuatro datos), y en base a ellos predecir el dato siguiente. Una vez fijo el número de variables de entrada y elegida la variable a pronosticar, se diseñó un modelo predictivo basado en la técnica de RNA. El tipo de modelo de RNA usado fue uno de multicapas, alimentado hacia adelante y entrenado mediante la técnica de propagación hacia atrás. Se probaron redes sin capa oculta y con una y dos capas ocultas. El mejor modelo resultó ser con una capa oculta, a diferencia de lo obtenido en trabajo anterior que reportaba que la red sin capa oculta era más eficiente. Los resultados fueron más precisos que los obtenidos con un modelo de persistencia (el valor en seis horas más será el mismo que el actual).

Palabras clave: Contaminación ambiental, series temporales, redes neuronales artificiales, pronóstico, material particulado pm2.5, dinámica no lineal.

ABSTRACT

An artificial neural network for the forecasting of concentrations of fine particulate matter in the atmosphere was designed. The data set analyzed corresponds to three years of pm2.5 time series (particulate matter in suspension with aerodynamic diameter less than 2,5 microns), measured in a station that belongs to Santiago's monitoring network (Red MACAM) and is located near downtown. We consider measurements of concentrations between May and August for years between 1994 and 1996. In order to find the optimal time spacing between data and the number of values into the past necessary to forecast a future value, two standard tests were performed, Average Mutual Information (AMI) and False Nearest Neighbours (FNN). The results of these tests suggest that the most convenient choice for modelling was to use 4 data with 6 hour spacing on a given day as input in order to forecast the value at 6 AM on the following day.

Once the number and type of input and output variables are fixed, we implemented a forecasting model based on the neural network technique. We used a feedforward multilayer neural network and we trained it with the backpropagation algorithm. We tested networks with none, one and two hidden layers. The best model was one with one hidden layer, in contradiction with a previous study that found that minimum error was obtained with a net without hidden layer. Forecasts with the neural network are more accurate than those produced with a persistence model (the value six hours ahead is the same as the actual value).

Keywords: Air pollution, times series, artificial neural networks, forecasting, pm2.5, nonlinear dynamics.

¹ Facultad de Ingeniería, Universidad Católica de la Ssma. Concepción, Chile, gsalini@ucsc.cl; fono (56-41) 2735690.

² Departamento de Física, Facultad de Ciencia, USACH, Chile, pperez@usach.cl; fono (56-41) 7763322.

INTRODUCCIÓN

Un problema recurrente en Santiago de Chile (durante el invierno) es la alta concentración de material particulado en la atmósfera. Se han efectuado estudios que han determinado la influencia que la contaminación atmosférica tiene sobre el deterioro de la salud humana [1, 2]. Por ejemplo, la agudización de enfermedades respiratorias, cardiovasculares, dermatológicas, cáncer, entre otras. Además de daños a los animales y vegetales, en cuanto a la reducción de su desarrollo y productividad. Por ello, en el último tiempo Santiago ha sido catalogada como una de las ciudades de mayor contaminación en el mundo. Por otro lado, la causa principal de esta contaminación ambiental sería la alta tasa de emisión de origen antropogénico (orgánicos e inorgánicos).

Se sabe que es la situación geográfica (topográfica) la que impide que en Santiago haya una dispersión tanto horizontal como vertical de los contaminantes. Además, coadyuvan a ello la densidad de la población, el sistema de transporte público (aunque recientemente se está implementando un nuevo sistema, conocido con el nombre de TranSantiago, el cual debiera de mejorar esta situación), así como la situación meteorológica, entre otros factores.

La calidad del aire en Chile se define en términos del promedio móvil de 24 hr de PM10, es decir, 24MA de concentración de material particulado. Estas partículas son lo suficientemente pequeñas para penetrar el sistema respiratorio humano. Si se sobrepasa el límite de $240 \mu\text{g}/\text{m}^3$, se estaría en situación de peligro para la salud de las personas, de acuerdo a la OMS (1979). La CONAMA la ha denominado situación de “preemergencia”.

A nivel de salud de la población, existe una especial atención a las sustancias químicas que tienen propiedades carcinogénicas o mutagénicas (que son los contaminantes orgánicos atmosféricos). Estos compuestos pueden ser absorbidos directamente cuando están en fase gaseosa, o acumularse en el tracto respiratorio humano cuando están asociados a partículas menores a $10 \mu\text{m}$ (PM10: fracción respirable). Fuentes de PM10 son la combustión y los procesos industriales. Además, el tráfico vehicular genera polvo, siendo el humo negro generado por las fuentes vehiculares diesel. Aunque el estándar normal en muchos países ha sido puesto en partículas tipo PM10, estudios recientes de epidemiología han sugerido estudiar partículas de tamaño menores que $2.5 \mu\text{m}$ (PM2.5) [3]. En este tipo de partículas están basadas las tres series temporales que hemos trabajado.

A partir de 1988 funciona en Santiago y de manera permanente una red de monitoreo automático de contaminantes atmosféricos y parámetros meteorológicos (red MACAM), con la cual se han obtenido series temporales que contienen la información necesaria para hacer un análisis exhaustivo de los contaminantes indicados más arriba y que nos interesan analizar en este trabajo.

Mediante técnicas de modelación de sistemas dinámicos [4], se ha logrado caracterizar diferentes tipos de series temporales en naturaleza. Para ello, corrientemente se emplean como herramientas matemáticas la función AMI (Información Mutua promedio o Average Mutual Information) y la función FNN (Falsos vecinos más Cercanos o False Nearest Neighbours). Por otro lado, se han introducido como una técnica complementaria las Redes Neuronales Artificiales (RNA), aplicadas al análisis de series temporales.

La función AMI, $I(\tau)$ nos entrega el promedio de la información binaria que se transmite entre datos espaciados en un intervalo de tiempo τ . Esta función tiene su propio mínimo en cierto instante, simbolizado por τ , que corresponde al espaciamiento que implica una correlación apropiada entre datos sucesivos para permitir un mejor modelamiento. Este primer mínimo hallado por el método AMI se considera un estimador del tiempo como para realizar la reconstrucción del atractor en el espacio de fases [5]. Por otro lado, la función FNN permite hallar la dimensión de encaje $d_e(\tau)$, correspondiendo a la mínima cantidad necesaria de variables dinámicas para modelar al sistema. En nuestro caso, corresponde a la cantidad de datos pasados consecutivos necesarios para poder hacer una predicción, un cierto intervalo de horas hacia el futuro. Para esta función, si su curva decrece rápidamente, pero nunca se va a cero, entonces se tienen indicios de que se estaría en presencia de ruido.

La técnica de RNA consiste de un sistema compuesto de una gran cantidad de procesadores simples (neuronas o PE's), funcionando colectivamente para resolver tareas complejas. Pero también se pueden entender como un algoritmo con varios parámetros ajustables, los cuales son fijados después de la etapa de entrenamiento. El grado de interacción entre los PE's está dado por los pesos de conexión entre ellos. Así, los pesos son determinados mediante el aprendizaje por ejemplos.

Los tipos de RNA más usados son el perceptrón y la red de multicapas (alimentada hacia delante), con función de transferencia no lineal (tipo sigmoide o tangente hiperbólica). Esta función, denominada también “función de activación”, usualmente tiene la siguiente forma

$$f(X) = \frac{1}{1 + e^{-X}} \quad (1)$$

Los inputs son los valores de las variables considerados relevantes para generar las salidas (outputs). El método de optimización para fijar los pesos es el denominado *backpropagation*. Así, la información codificada adicionalmente fluye sólo desde una capa a la siguiente.

El perceptrón con función de transferencia lineal se puede decir que es un método de regresión lineal, en la cual todas las neuronas de la capa de entrada se conectan con una única (o más) neurona en la capa de salida.

En cambio, la red de multicapas incorpora una función no lineal, cuyo argumento es una combinación lineal de valores de la capa previa, generando un valor final en las neuronas de la capa de salida. El número de neuronas en la capa intermedia se obtiene, en general, por ensayo y error, o por experiencia en problemas similares.

METODOLOGÍA EN EL TRATAMIENTO DE LA BASE DE DATOS

Se determinó la función AMI para cada una de las series de tiempo, las cuales consistieron de tres series temporales entre el 1º de mayo al 30 de septiembre (correspondiente a los meses de invierno) y entre los años de 1994, 1995 y 1996, y así calcular lo que se denomina el tiempo de rezago (o lag) τ . Para ello se evaluó $I(\tau)$ y a partir de ésta se calculó el tiempo de rezago buscado. La idea detrás de esta función es que una buena elección para el tiempo de rezago τ es uno tal que, dado el estado del sistema $s(n)$, proporcione la máxima información nueva medida en $s(n + \tau)$. Un gráfico de $I(\tau)$ comienza en un valor alto. A medida que τ crece, $I(\tau)$ decrece, entonces usualmente se incrementa nuevamente. Se ha sugerido que el valor del retardo temporal (o tiempo de rezago) donde $I(\tau)$ alcanza su primer mínimo, sea usado para la reconstrucción del espacio de estado. Es en este punto donde $s(n + \tau)$ agrega máxima información al conocimiento que se tiene de $s(n)$, es decir, la redundancia es menor.

Es así que se halló la separación temporal óptima de datos necesarios para capturar la dinámica interna, a través de los rasgos más relevantes de tres series temporales semejantes (en cuanto a condiciones de borde), relativos a la concentración de material particulado PM2.5 en la atmósfera, durante los meses de invierno y en un período de tres años consecutivos, en Santiago [6, 7, 8].

A continuación se calculó la función FNN, la cual permite determinar la dimensión encajada, d_e . Este es un método para escoger la mínima dimensión encajada de una serie de tiempo. La idea de este método es hallar los vecinos más cercanos de cada punto en una dimensión dada, luego chequear para ver si estos puntos son aún vecinos próximos en una dimensión superior. El porcentaje de FNN debería caer a cero cuando la dimensión encajada ha sido alcanzada.

Más explícitamente, la dimensión encajada de un sistema dinámico es una dimensión global entera que da el número de coordenadas necesarias para desplegar (unfold) su trayectoria, en el espacio de fases. Para el cálculo de esta dimensión se necesita tener el rezago temporal, τ , para cada serie de tiempo. Una vez que se tiene este valor, se escoge una dimensión d para algún punto $s(k)$. Con ello será suficiente conocer el vector:

$$y(k) = (s(k), s(k + \tau), \dots, s(k + (d - 1)\tau)), \quad (2)$$

para poder predecir la evolución del sistema dinámico. Entonces, en el espacio d -dimensional, el vector $y(k)$ tiene un vecino más cercano $y^{NN}(k)$, el cual es el vector con la mínima distancia euclídea a él. Si d es menor que d_e , es muy probable que este vecino más cercano sea falso; vale decir, que corresponde a una proyección desde el espacio dimensional d_e real. Este método requiere colocar dos valores umbrales a algún valor más arbitrario, los cuales se usan para determinar los falsos vecinos. Cabe señalar que este método es sensible a la tasa de muestreo de la serie de tiempo. Por último, se puede decir que este método aunque es intuitivo y fácil de implementar, no es directo su uso e interpretación.

La función AMI se define como [3]:

$$I(\tau) = \sum_{s(n), s(n+\tau)} P(s(n), s(n+\tau)) \log_2 \left[\frac{(P(s(n), s(n+\tau)))}{(P(s(n))P(s(n+\tau)))} \right] \quad (3)$$

siendo $P(s(n))$ la probabilidad de obtener el valor s en el instante n , y $P(s(n), s(n + \tau))$ es la probabilidad conjunta de conseguir un valor dado de la variable s en el instante $n + \tau$, dado el valor en el instante n .

Por su parte, la función FNN se define como sigue: el número de falsos vecinos próximos para un d dado es el número de veces que la ecuación (4) supera un umbral dado [3].

$$\sqrt{\frac{R_{d+1}^2(k) - R_d^2(k)}{R_d^2(k)}} = \frac{|s(k + d\tau) - s^{NN}(k + d\tau)|}{R_d(k)} \quad (4)$$

en que R^2_d es el cuadrado de la distancia euclíadiana entre los puntos vecinos más cercanos, vistos desde un espacio en d dimensión; y lo mismo para R^2_{d+1} .

La base de datos se procesó usando los valores calculados de los parámetros τ y d_e . Específicamente, se seleccionaron cuatro entradas (las mismas que la dimensión de encaje), una única salida que corresponde al pronóstico de un paso en adelante que se quiere efectuar. Si bien el gráfico de la función AMI sugiere un $\tau = 7$, se usó un $\tau = 6$, dado que siendo este múltiplo de 24, permite utilizar de mejor manera los datos de concentraciones horarias de PM2.5.

Los datos disponibles se agruparon en tres grupos, los cuales se generaron desplazando una ventana de tiempo cada seis horas, sobre la serie completa. Así, de los 3.672 datos correspondientes a una determinada serie temporal, quedaron finalmente 122 datos por columna, de un total de cinco columnas, siendo esta última la 06 a.m. que se quiere pronosticar (las cuatro primeras columnas corresponden a las 06 a.m., 12 p.m. 18 p.m. y las 00 a.m., respectivamente). Este procedimiento se efectuó con todas las bases de datos: pm2594, pm2595 y pm2596. Con todo esto, la nueva base de datos que surgió de la pm2594 se utilizó para entrenamiento; la nueva base de datos surgida de la antigua pm2595 se usó para test de validación durante el aprendizaje, y finalmente la nueva base de datos que emergió de la antigua pm2596 se usó para test, luego de alcanzar la red neuronal el aprendizaje óptimo.

Se diseñaron tres tipos de redes neuronales artificiales: un perceptrón, una red de multicapas con una sola capa escondida y otra red de multicapas, pero con dos capas escondidas. La capa de entrada para los tres tipos de redes tenía cuatro neuronas. Mientras que la capa de salida tenía sólo una neurona. Los cuatro PE's de entrada se debe a que es la cantidad óptima que se obtiene a través de la función FNN, lo cual concuerda con [9]. Además, la neurona de salida se debe a que se requiere hacer sólo un pronóstico en adelante.

Es necesario señalar acá que el perceptrón simple es un modelo dinámico lineal, al implementarse con una función de transferencia lineal. Por lo que conjuntamente con el método de persistencia sirve de comparación respecto de la red de multicapas.

La actualización de los pesos se realiza de acuerdo a la siguiente ecuación:

$$\Delta\omega_{ji}^{[s]}(t+1) = lcoef e_{ji}^{[s]} x_{i}^{[s-1]} + mom \Delta\omega_{ji}^{[s]}(t) \quad (5)$$

donde $x_{i}^{[s]}$ es la entrada de la neurona i en la capa previa ($s - 1$); $e_{ji}^{[s]}$ es una medida del error local en la j -ésima neurona en el nivel s ; $lcoef$ es un coeficiente de aprendizaje; mom es un término, llamado momentum, que evita un lento aprendizaje, vale decir, sirve para aumentar la velocidad de convergencia; $\Delta\omega_{ji}^{[s]}$ es la regla del gradiente descendente para los pesos que unen a la i -ésima neurona en la capa ($s - 1$) con la j -ésima neurona en la capa s . Esta ecuación es una variante del algoritmo estándar conocido. Este método usado que ha demostrado ser un buen método, aunque un poco lento, es suficiente para nuestra base de datos. Requiere más tiempo de computación solamente.

El método de pronóstico, por otro lado, de una serie de tiempo a través de las *Persistencias*, consiste en suponer que un cierto valor en el pasado, de la serie de tiempo, se mantendrá (o persistirá) en el futuro (en un cierto intervalo de tiempo). Así, en nuestro caso, si tomamos el tiempo de rezago como 6, Persistencia implica suponer que el dato de las 6 a.m. es igual al de las 12 de la noche anterior. Podremos comparar el pronóstico de este método de predicción con el de las redes neuronales artificiales.

ANÁLISIS Y RESULTADOS

En la figura 1 se muestra función AMI para la base de datos pm2594, en la cual se observa que el valor del tiempo de rezago es 7. Lo mismo se obtuvo para pm2596. Para pm2595 se obtuvo un tiempo de rezago de 8. En la figura 2 se muestra la función FNN para la base de datos pm2596, obteniéndose idéntica dimensión de encaje para las otras dos bases de datos.

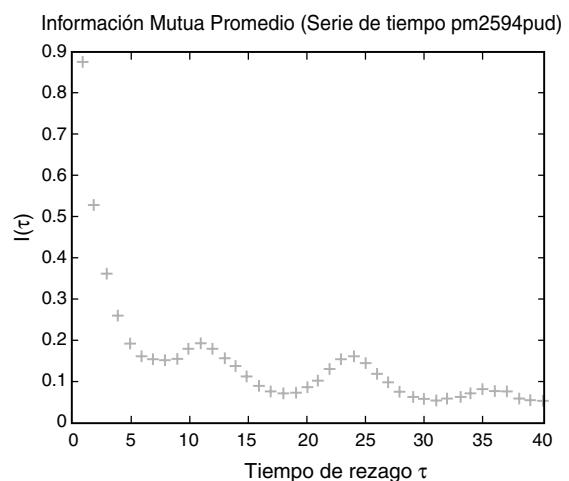


Figura 1. Información Mutua Promedio para la serie pm2594.

A pesar de que la función AMI arrojó valores de 7 y 8 como tiempo de rezago, como se observa en la tabla 1, se prefirió trabajar con un tiempo de 6. Se sabe que es el valor $\tau = 7$ que debemos emplear para calcular la función FNN. Sin embargo, se ha demostrado [3] que dicha función no va exactamente a cero. En cambio, si se elige $\tau = 6$ se consigue que FNN vaya exactamente a cero y con ello logramos capturar la dinámica correcta de nuestras series temporales. Además, demostramos que en nuestro caso se logran predicciones significativas con la técnica de redes neuronales artificiales, con hasta 6 horas en adelante y no con 7, como se había reportado en [3]. Por ello se ha escogido $\tau = 6$ como tiempo de rezago.

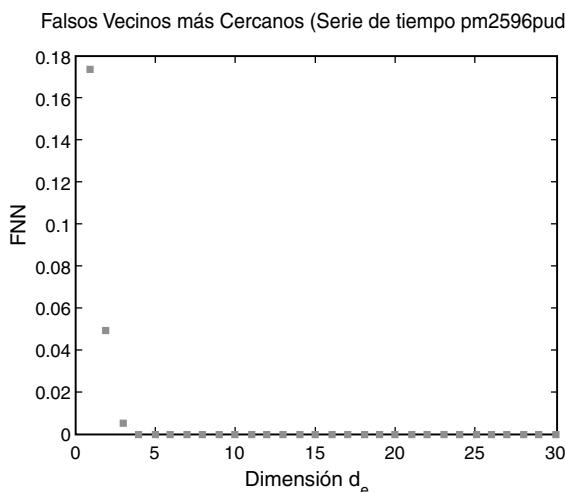


Figura 2. Falsos vecinos más cercanos para pm2596.

Como criterio de comparación de los modelos se usó el error RMSE – por root normal mean square error-, normalizado (RNMSE), definido como:

$$\text{RNMSE} = \frac{\sqrt{\frac{1}{N} \sum_i (O_i - P_i)^2}}{\sum_i O_i} \quad (6)$$

en que O corresponde al valor real y P al predicho. El error en la predicción se calculó comparando el valor predicho por la red en la etapa de test, con el valor que efectivamente se presentó.

En la tabla 1 se tabulan los parámetros de tiempo de rezago, τ y la dimensión encajada, d_e , para las tres bases de datos.

Se puede observar que a pesar de que la función AMI para la segunda base de datos da un valor de rezago

de 8 (y las restantes de 7), por las razones expuestas anteriormente se ha escogido el valor para AMI de 6, en las tres bases de datos.

Como se mencionó más arriba, el valor de la dimensión de encaje, para las tres series temporales, resultó ser de 4. Este dato sirve para determinar el número de PE's en la capa de entrada, para el diseño de nuestra red neuronal.

Tabla 1. Funciones AMI y FNN para las tres bases de datos estudiadas.

	AMI	FNN
pm2594	7	4
pm2595	8	4
pm2596	7	4

Tabla 2. Valores RNMSE para distintos tipos de modelos RNA.

NNCO	Mom	R Apren	TF Trans	R Lcoef	RNMSE
Perceptrón simple					
0	0	N-C	L	0,5	0,5505
0	0	N-C	S	0,5	0,5963
0	0	N-C	S	0,5	0,6087
0	1	D-R	S	0,5	0,5372
0	1	N-C	L	0,5	0,5385
0	0	N-C	L	0,9	0,549
Red de Multicapas (MLP)					
1	0	N-C	S	0,5	0,8375
1	0	D-R	T	0,5	0,5698
1	0	D-R	T	0,9	0,5499
1	0	N-C	S	0,9	0,5479
5	1	D-R	S	0,5	0,5590
7	1	D-R	S	0,5	0,5156
10	0	D-R	S	0,5	0,5648
10	1	D-R	S	0,5	0,5259

L : Lineal; S: Sigmoide; T: Tanh.

NNCO : número de neuronas en capa oculta.

Mom : momentum.

R Apren : regla de aprendizaje.

TF Trans : tipo de función de transferencia.

R LCoef : razón LCoef.

RNMSE : error de la raíz cuadrática promedio normalizado.

N-C : Normalized cumulative learning rule.

D-R : Generalized delta rule.

En la tabla 2 se presentan los parámetros usados para las diferentes redes neuronales implementadas. Se aprecia que la “mejor” red neuronal fue aquella con una capa escondida y un número de 7 neuronas en dicha capa. Además, se muestra el caso del perceptrón simple, con función de transferencia lineal, así como función no lineal, siendo el mejor perceptrón aquel con un valor RNMSE de 0,5372 (con función no lineal sigmoide). Para algunos modelos de red neuronal se han cambiado ciertos parámetros, que aparecen en la ecuación (5). Además, se han incluido dos tipos de reglas de aprendizaje: la regla Delta generalizada (D-R) y una variante denominada Normalized cumulative learning rule (N-C), la cual para aliviar el problema de presentación estructurada del conjunto de datos en la entrada acumula los cambios de los pesos sobre varias presentaciones de entrenamiento y, posteriormente, hace la actualización de todos a la vez, a diferencia de la regla delta usual, en la cual los pesos se van actualizando secuencialmente. Los cambios en los pesos para la regla delta, en el primer caso, son acumulados en el término de momentum estándar ecuación (5).

En resumen, podemos observar de la tabla 2 que perceptrones simples con función de transferencia lineal son los que peor aprendieron. Al cambiárseles la función de transferencia de lineal a no lineal (sigmoide), se aprecia que mejoraron su aprendizaje. Por otro lado, las redes multicapas con una capa escondida y entre 7 y 10 neuronas en esa capa, obtuvieron mejor aprendizaje, a diferencia de trabajo [3] en el cual un perceptrón superó a red de multicapa en cuanto a predecir ciertas horas en adelante, en todas las situaciones.

Por otra parte, mediante el método de persistencia se obtuvo un valor RNMSE de 0,7649.

En la figura 3 se muestra la arquitectura de la mejor red neuronal que obtuvo el mejor valor RNMSE, de acuerdo a la tabla 2. Se usó una regla de aprendizaje delta generalizada, con función de transferencia no lineal tipo sigmoide.

Finalmente, podemos decir que ambos tipos de redes neuronales lograron superar al método de persistencias, siendo para el caso de red de multicapas, entre un 21,4% y un 32,4% mejor. En el caso del perceptrón simple, fue entre un 16,5% y un 28,2% mejor al método de persistencias.

En el futuro se piensa investigar el efecto de incorporar como entradas, valores de variables meteorológicas, velocidad del viento, humedad, etc.

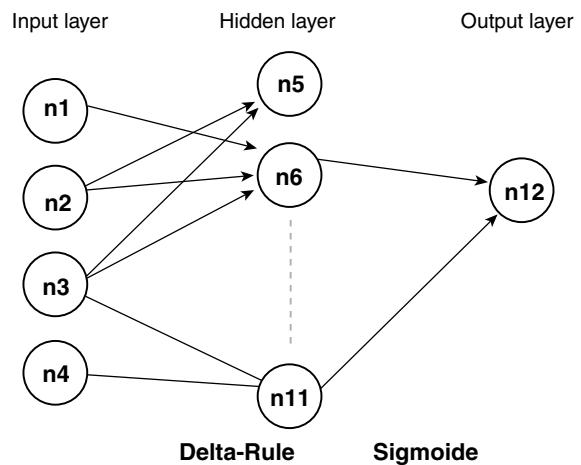


Figura 3. Arquitectura de la mejor red con valor RNMSE 0.5156. Se muestran algunas conexiones entre los nodos de entrada y la capa oculta y entre esta última y el PE de salida.

CONCLUSIONES

Se puede apreciar que con la metodología de las redes de multicapas diseñadas se obtiene una automatización satisfactoria del proceso de predicción de un paso en adelante, de nuestra serie de tiempo, en comparación con el método de predicción de persistencias.

Además, el perceptrón simple con función lineal fue superior en efectuar predicciones de un paso en adelante que el método de persistencia.

Se observó, además, que el perceptrón simple mejora su capacidad de predecir cuando se usa una función de transferencia no lineal. Sin embargo, no logra hacer mejores predicciones en relación a nuestro modelo de multicapas.

En consecuencia, y a diferencia de lo mencionado en el artículo de Pérez-Reyes [3], apreciamos que un perceptrón simple no superó a una red de multicapas, en su capacidad de predecir. Por lo que una función de transferencia no lineal no mejora la predicción.

Se puede afirmar que cuando efectos no lineales no son demasiado importantes en la modelación, las redes de multicapas no son significativamente mejores que el perceptrón. Sin embargo, como ocurrió en este caso, cuando estos efectos no lineales pasan a ser importantes, las redes de multicapas son mejores en cuanto a su capacidad de predicción, respecto de los modelos lineales.

Finalmente podemos afirmar que la predicción de la concentración de contaminantes en la atmósfera en la ciudad de Santiago puede llevarse a cabo empleando técnicas utilizadas para modelar sistemas dinámicos, así como la técnica de las RNA.

AGRADECIMIENTOS

Este trabajo fue en parte financiado por el proyecto FONDECYT N° 1050015 y en parte por la Dirección de Investigación y Postgrado de la Universidad Católica de la Ssma. Concepción.

REFERENCIAS

- [1] M. Ilabaca, et al. "Association between levels of fine particulate and emergency visits for pneumonia and other respiratory illnesses among children in Santiago, Chile". *Journal of the Air and Waste Management Association*. Vol. 49, pp. 154-163. 1999.
- [2] J. Schwartz, F. Laden and A. Zanobetti. "The concentration-response relation between PM_{2.5} and daily deaths". *Environmental health perspectives*. Vol. 110 N° 10, pp. 1025-1029. Octubre, 2002.
- [3] P. Pérez and J. Reyes. "Prediction of particulate air pollution using neural techniques". *Neural Computation & Application*. Vol. 10, pp. 165-171. 2001.
- [4] H. Abarbanel. "Analysis of Observed Data". Institute for Nonlinear Science, Springer-Verlag. 1996.
- [5] A.M. Fraser and H.L. Swinney. "Independent coordinates for strange attractors from mutual information". *Physical Review A*. Vol. 33, pp. 1134-1140. 1986.
- [6] J. Reyes y P. Pérez. "Predictibilidad del material particulado PM2.5 en Santiago de Chile utilizando técnicas de modelación de sistemas dinámicos y redes neuronales". *Contribuciones Científicas y Tecnológicas*. ACB. N° 121, pp. 107-108. Diciembre 1998.
- [7] P. Pérez, A. Trier and J. Reyes. "Prediction of PM_{2.5} concentrations several hours in advance, using neural networks in Santiago, Chile". *Atmospheric Environments*. Vol. 34, pp. 1189-1196. 2000.
- [8] J.B. Ordieres, E.P. Vergara, R.S. Capuz and R.E. Salazar. "Neural network prediction model for fine particulate matter (PM2.5) on the US-Mexico border in El Paso (Texas) and Ciudad Juarez (Chihuahua)". *Environmental Modelling & Software* 20, pp. 547-559. 2005.
- [9] P. Pérez and J. Reyes. "Prediction of maximum of 24-h average of PM10 concentrations 30 h in advance in Santiago, Chile". *Atmospheric Environment*. Vol. 36, pp. 4555-4561. 2002.