



# Un estudio sobre la influencia de la función objetivo en evolución gramatical para regresión simbólica

J. Manuel Colmenar

Universidad Rey Juan Carlos

C/. Tulipán s/n, 28933, Móstoles, Spain

Email: josemanuel.colmenar@urjc.es

J. Ignacio Hidalgo

Universidad Complutense de Madrid

C/. Prof. José García Santesmases 9, 28040, Madrid, Spain

Email: hidalgo@dacya.ucm.es

**Resumen**—La evolución gramatical es una variante de la programación genética, que permite, entre otras características, introducir conocimiento del problema en el proceso de búsqueda. Una de las aplicaciones más importantes de esta técnica, al igual que de la programación genética tradicional, es la regresión simbólica. La regresión simbólica consiste en obtener expresiones matemáticas que modelen un conjunto de datos. Como todo algoritmo evolutivo, la búsqueda tiene que estar guiada por una función objetivo. Esta función objetivo, en el caso de la regresión simbólica, debe medir la proximidad de la función matemática reportada por el algoritmo a los datos objeto del problema. Por lo tanto, lo más habitual es utilizar el error cuadrático medio, el valor de  $R^2$ , o alguna otra variante de la diferencia entre los puntos. En este artículo investigamos la influencia de la utilización del error cuadrático medio y de  $R^2$  en el proceso de búsqueda y en la calidad de las soluciones obtenidas con evolución gramatical en problemas de regresión simbólica.

## I. INTRODUCCIÓN

Uno de los paradigmas más importantes en cuanto a las aplicaciones de la inteligencia artificial es la computación evolutiva y dentro de ella es de fundamental importancia la programación genética (PG, o GP, de sus siglas en inglés: *Genetic Programming*) [6]. La PG es una metodología que permite obtener de una forma automática programas que representen soluciones a problemas o que realicen una determinada función. Utiliza una representación de las soluciones en forma de árboles, lo que hace que en ocasiones el tamaño de memoria necesario para representar las soluciones de la población sea excesivo. Por ello se han desarrollado alternativas que funcionen bajo los mismos principios pero que permitan reducir tanto el espacio ocupado en memoria, como el tiempo efectivo de búsqueda de soluciones. La evolución gramatical, también conocida como gramáticas evolutivas (de una mala traducción del inglés, *Grammatical Evolution*) [7], es una variante de la programación genética en la que se utilizan una serie de reglas recogidas en una

gramática para construir los árboles que representan una solución. Esto permite introducir conocimiento del problema en el proceso de búsqueda y controlar el tamaño de los árboles que representan a las soluciones.

Una de las aplicaciones más relevantes de esta técnica, al igual que de la programación genética tradicional, es la regresión simbólica. La regresión simbólica consiste en obtener expresiones matemáticas que representen un conjunto de datos. Sus aplicaciones en la actualidad son innumerables ya que es una técnica fundamental en el análisis de datos y modelado de sistemas a partir de registros de información [3, 4]. Como todo algoritmo evolutivo, la búsqueda tiene que estar guiada por una función objetivo. Esta función objetivo, en el caso de la regresión simbólica, debe medir la proximidad de la función matemática reportada por el algoritmo a los datos objeto del problema. Por lo tanto, lo más habitual es utilizar el error cuadrático medio, el valor de  $R^2$ , o alguna otra variante de la diferencia entre los puntos. Sin embargo, aunque el objetivo buscado sea el mismo, al utilización de una u otra función puede tener consecuencias importantes en el resultado final del proceso de búsqueda.

En este artículo se analiza la influencia de la utilización del error cuadrático medio y de  $R^2$  tanto en el proceso de búsqueda como en la calidad de las soluciones obtenidas con evolución gramatical en problemas de regresión simbólica. Como se verá en la discusión del trabajo, a partir de las expresiones de las funciones objetivo, intuitivamente se puede pensar que cuando se utiliza el valor de  $R^2$  los resultados de la búsqueda deben reportar expresiones que tengan una alta correlación con los datos, pese a que no estemos obteniendo una función que represente exactamente todos los datos. Sin embargo, al utilizar el error cuadrático medio (*RMSE*, del inglés *Root Mean Square Error*) se esperaría que las buenas soluciones representen los datos de manera más fiel. Para realizar el estudio se han utilizado varios problemas clásicos de

regresión simbólica, y se han comparado los resultados utilizando gramáticas iguales con funciones objetivo diferentes. Los resultados muestran la diferencia entre los modelos y soluciones obtenidas así como la influencia de la función objetivo en estos resultados.

El resto del artículo está organizado como sigue. En la sección II se describe brevemente la técnica de la evolución gramatical. En la sección III se revisa el diferente enfoque de las funciones objetivo bajo estudio. La exposición de los resultados y su análisis se realiza en la sección IV y la sección V resume las principales conclusiones de este trabajo.

## II. EVOLUCIÓN GRAMATICAL

La evolución gramatical es un paradigma evolutivo que deriva de la programación genética. En resumen, la aportación de la evolución gramatical consiste en modificar la representación de los individuos utilizando cromosomas en lugar de árboles. Este cambio permite que se puedan utilizar operadores clásicos de algoritmos genéticos como el cruce, mutación o selección, cuya implementación es más sencilla y eficiente al ejecutar sobre cromosomas que sobre árboles. Además, para realizar la decodificación del cromosoma se utiliza una gramática, que incluye reglas de producción que permiten generar las expresiones que corresponden al fenotipo del individuo.

Pese a ser una técnica reciente, durante los últimos años se ha utilizado de forma extensa en problemas de modelado. Para mayor información sobre evolución gramatical se puede recurrir a los trabajos mencionados en la introducción y también acudir a [2] ó [5] para ver el detalle del proceso de decodificación.

Como ejemplo de gramática, la Figura 1 muestra la que se ha utilizado en los experimentos de este artículo. En particular, esta gramática corresponde al conjunto de datos de referencia Vladislavleva-F8 porque considera dos variables de entrada, como se indica en la última regla. Para el resto de las pruebas de referencia, esta regla será la única modificada, adaptada al número de variables de entrada.

## III. FUNCIONES OBJETIVO EN PROBLEMAS DE REGRESIÓN SIMBÓLICA

A continuación, se planteará la formulación de las funciones objetivo bajo estudio en este artículo. Siendo  $O$  el vector de valores observados (aquellos provenientes de la optimización) y  $E$  el vector de valores esperados (formado por la variable objetivo), y  $\bar{O}$  y  $\bar{E}$  los valores promedio de cada uno de ellos, las expresiones para las funciones objetivo son las siguientes:

```
# Rule to create the main structure of the
# produced expression
<expr> ::= (<expr> + <expr>) / (<expr> + <
  <expr>)
  | <expr> + <expr> | <expr> - <expr>
  | <expr> * <expr> | <expr> / <expr>
  | <c> | <c>.<c> | <c><c>.<c><c> | <
    <var> | <var>
  | <expr> <op> <expr> | <c>^<var>
  | exp(abs(<c> <op> <var>)) | log(abs
    (<c> <op> <var>))

# Digits
<c> ::= 0|1|2|3|4|5|6|7|8|9

# Arithmetic operands
<op> ::= +|-|*|/

# Input variables
<var> ::= X1 | X2
```

Figura 1: Gramática diseñada para abordar los experimentos. El ejemplo corresponde a Vladislavleva-F8, que trata 2 variables de entrada.

$$R^2 = \left( \frac{\sum_i ((E_i - \bar{E})(O_i - \bar{O}))}{\sqrt{\sum_i (E_i - \bar{E})^2 \sum_i (O_i - \bar{O})^2}} \right)^2$$

$$RMSE = \sqrt{\frac{1}{N} \sum_i (O_i - E_i)^2}$$

$R^2$  no es sólo una medida de la bondad de ajuste, sino también una medida de hasta qué punto el modelo (el conjunto de variables independientes que se seleccionó) explica el comportamiento (o la varianza) de su variable dependiente. Por lo tanto, si un modelo tiene una  $R^2$  de 0,6, esto explica el 60 % del comportamiento de su variable dependiente. Ahora, si se se usa el  $R^2$  Ajustado que esencialmente penaliza el  $R^2$  por el número de variables que usa el modelo, se puede obtener una idea bastante buena de cuándo hay que dejar de agregar variables al modelo (y eventualmente simplemente obtener un modelo que se ajuste). Si el  $R^2$  Ajustado de los modelos obtenidos con un conjunto de variables es 0,8, y al añadir una variable sólo aumenta en 0,01, lo más indicado es que no sea necesario añadir esa variable. Realmente, si el algoritmo de evolución gramatical está bien diseñado, esto no debe suceder ya que de inicio se habrán seleccionado las variables que intervienen y el propio algoritmo proporcionará un modelo que implícitamente indica qué variables son las más importantes, ya que son las que aparecen en el modelo.

En el caso de RMSE, éste devuelve una estimación de la desviación estándar de los residuales<sup>1</sup> [1]. Por

<sup>1</sup>La diferencia entre los valores reales y los estimados es el residual.



tanto, a mayor RMSE, mayor desviación estimada en los errores del modelo.

En principio, por tanto,  $R^2$  mide el ajuste de un modelo sobre un conjunto concreto de datos, mientras que RMSE mide la dispersión de los errores. A priori, podría parecer que estas medidas indican ambas la calidad de un modelo, por lo que podrían ser intercambiables. Sin embargo, como se verá en la sección de resultados, valores altos de una de las métricas no implican valores altos de la otra.

#### IV. RESULTADOS EXPERIMENTALES

Los experimentos de este artículo se han realizado sobre un conjunto de tres problemas de regresión estándar, que se denominarán *benchmarks* a partir de este punto, haciendo uso de la expresión comunmente utilizada en inglés. Estos problemas se han seleccionado como parte de los problemas recomendados para PG en [10], y son los siguientes: *Tower*, *Vladislavleva-F8* y *Spatial co-evolution*. A continuación se detallan algunas de las características de estos problemas:

- En el conjunto de datos *Spatial co-evolution* [8], la variable objetivo se define de la siguiente manera:

$$F(x, y) = \frac{1}{1 + x^{-4}} + \frac{1}{1 + y^{-4}}$$

En el conjunto de entrenamiento original existen 676 registros en los cuales los valores para las variables  $x$  e  $y$  se muestrean desde -5 a +5 en intervalos de 0.4. En el conjunto de test hay 1000 registros en los que las variables  $x$  e  $y$  se muestrean en el intervalo  $[-5, \dots, +5]$  aleatoriamente. En conjunto, forman un total de 1676 datos, y se han utilizado como serie de regresión al completo.

- El conjunto de datos *Tower* [9] proviene de un problema industrial sobre medidas de cromatografía de gases en una torre de destilación. El conjunto de datos contiene 5000 registros y 25 variables de entrada potenciales. La variable de respuesta es la concentración de propileno en la parte superior de la torre de destilación. Los datos de *Tower* se pueden descargar en <http://www.symbolicregression.com/?q=towerProblem>.
- El conjunto de datos *Vladislavleva-F8*, denotado como *VF8* a partir de este momento, define las variables de salida como funciones de  $x_1$  y  $x_2$ , de la siguiente manera:

$$F_8(x_1, x_2) = \frac{(x_1 - 3)^4 + (x_2 - 3)^3 - (x_2 - 3)}{(x_2 - 2)^4 + 10}$$

Para todos los *benchmarks* se han estudiado las dos funciones objetivo,  $R^2$  y *RMSE*, con dos tipos de inicialización para la primera población del algoritmo: inicialización aleatoria (*random*) e inicialización

*sensible*, que controla el tamaño de los fenotipos de los individuos de la generación inicial evitando que sean excesivamente grandes. De cada experimento se han realizado diez ejecuciones diferentes utilizando los mismos parámetros de entrada para el algoritmo de GE, que se muestran en la Tabla I. Los experimentos se ejecutaron en un ordenador equipado con un procesador Intel Core i7 a 2.9 GHz con 16 GB de RAM sobre sistema operativo Mac OSX 10.13.5.

Parámetro	Valor
Población	300 individuos
Generaciones	1700
Selección	Torneo (2 indiv.)
Prob. Cruce	80 %
Prob. Mutación	2 %
Función objetivo	$R^2$ ó <i>RMSE</i>
Inicialización	<i>random</i> , <i>sensible</i>
Longitud del Cromosoma	256
Número máximo de wraps	5

Tabla I: Valores de los parámetros utilizados en los experimentos.

El primer aspecto a analizar es la evolución de las ejecuciones a través del análisis del mejor individuo. Para ello, se tomó el valor de la función objetivo cada 100 generaciones, en cada una de las ejecuciones. Dado que se lanzaron 10 ejecuciones para cada experimento, los datos que se muestran se corresponden con el promedio de las 10 ejecuciones.

En primer lugar, las figuras 2 y 3 muestran la evolución para el caso de la inicialización con una población aleatoria (*random*), utilizando como función objetivo  $R^2$  y *RMSE* respectivamente. Dado que el algoritmo trata de minimizar, el valor que se presenta en el caso de  $R^2$  es en realidad  $1 - R^2$ , y así será para el resto de experimentos. Como se puede apreciar, el problema más difícil de resolver es *Tower*. Por otro lado, la evolución tiene una tendencia hacia abajo menos pronunciada a partir de la generación 500 aproximadamente en ambas gráficas.

Las figuras 4 y 5 muestran la convergencia para el caso de la inicialización *sensible*, utilizando como función objetivo  $1 - R^2$  y *RMSE* respectivamente. De nuevo, el comportamiento es similar en ambos casos, siendo *Tower* el problema más difícil. En el caso de la comparativa entre los diferentes tipos de inicialización, se puede ver que entre las figuras 2 y 4, que utilizan la misma función objetivo,  $1 - R^2$ , y diferente inicialización, la inicialización *sensible* alcanza mejores valores promedio a lo largo de toda la ejecución del algoritmo en los tres *benchmarks*.

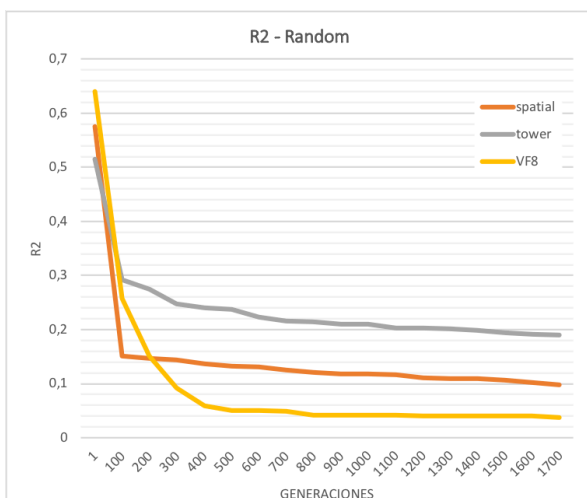


Figura 2: Valores promedio de la función objetivo:  $1 - R^2$ . Inicialización *random*.

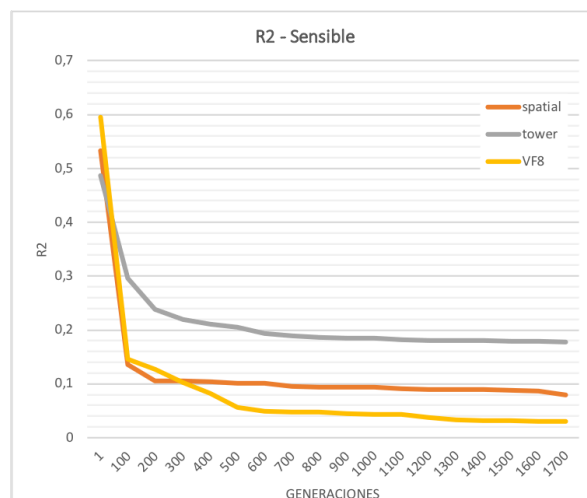


Figura 4: Valores promedio de la función objetivo:  $1 - R^2$ . Inicialización *sensible*.

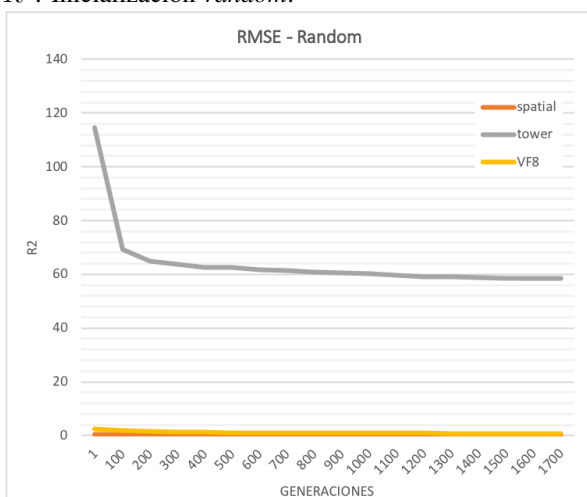


Figura 3: Valores promedio de la función objetivo: *RMSE*. Inicialización *random*.

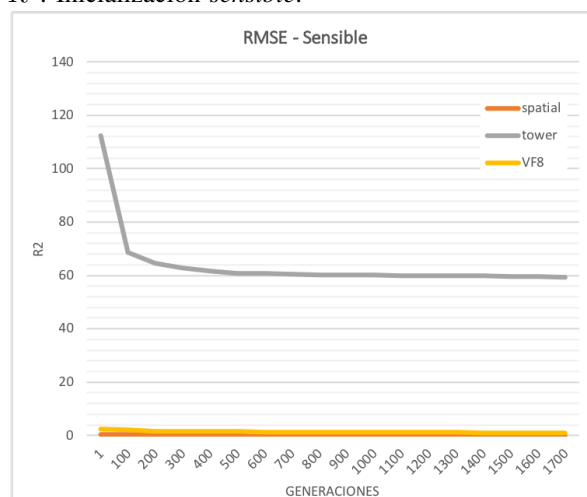


Figura 5: Valores promedio de la función objetivo: *RMSE*. Inicialización *sensible*.

En el caso de la optimización con *RMSE*, la inicialización *sensible* no aporta valores mejores en la evolución. A pesar de que la escala no permite ver las diferencias gráficamente en las figuras 3 y 5, sí que se aprecia que al principio los valores promedio son mejores, pero según avanza la evolución, los experimentos con inicialización aleatoria obtienen mejores valores promedio.

Independientemente de las funciones objetivo utilizadas, se han calculado los valores de  $R^2$  y *RMSE* para todas las soluciones obtenidas en los experimentos. La figuras 6 y 7 muestran, respectivamente, los valores promedio de  $R^2$  y *RMSE* para cada uno de los *benchmarks* en los diferentes casos. Se denota el experimento con optimización a través de  $R^2$  e inicialización aleatoria

como *R2-Random*; optimización a través de  $R^2$  e inicialización *sensible* como *R2-Sensible* y, de manera análoga, para *RMSE* y las dos inicializaciones estudiadas como *RMSE-Random* y *RMSE-Sensible*.

En la Figura 6 se puede apreciar cómo los valores de *R2-Random* y *R2-Sensible* son mayores que en la optimización con *RMSE*. En el caso, por ejemplo, de *Tower*, la diferencia es muy significativa, llegando a estar por encima del 30%. Los valores concretos de  $R^2$  se muestran en la Tabla III. Se puede apreciar también que la inicialización *sensible* mejora los resultados hasta en un 2% para los experimentos donde la función objetivo es  $1 - R^2$ .

La Figura 7 muestra los valores de *RMSE* para las mismas soluciones que se han utilizado en la

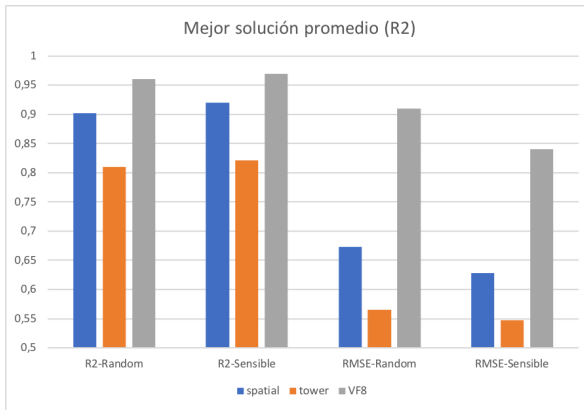


Figura 6: Valores promedio de  $R^2$  para los diferentes experimentos. Mejor cuanto más alto.

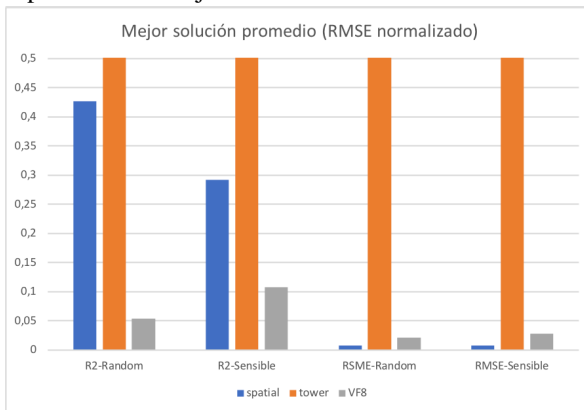


Figura 7: Valores promedio de  $RMSE$  normalizados para los diferentes experimentos. Mejor cuanto más bajo.

	spatial	tower	VF8
R2-Random	0,901794689	0,80986102	0,960807366
R2-Sensible	0,920002539	0,821656958	0,969546579
RMSE-Random	0,672605164	0,56536177	0,910126819
RMSE-Sensible	0,627583903	0,547273778	0,840721671

Tabla II: Valores promedio de  $R^2$ .

Figura 6. En este caso, dado que los valores presentan diferentes órdenes de magnitud, se han normalizado respecto a la décima parte del valor máximo. Los valores concretos de  $RMSE$  se presentan en la Tabla III. Como se puede apreciar, los valores de  $RMSE$  son mucho mejores en el caso de los algoritmos que utilizan esta métrica como función objetivo, llegando a obtener un orden de magnitud de diferencia. Resulta significativo señalar que la inicialización *sensible* no mejora los resultados en este caso. Al contrario, se llega a peores resultados en los tres problemas bajo estudio.

Estos resultados son coherentes con lo que se espera de algoritmo de optimización, que consigue la mejora

	spatial	tower	VF8
R2-Random	15,58366595	348,2669206	1,942964282
R2-Sensible	10,67370612	365,2646472	3,94804835
RMSE-Random	0,272769052	58,39444643	0,787039727
RMSE-Sensible	0,298058917	59,40281873	1,035522363

Tabla III: Valores promedio de  $RMSE$ .

en los valores de la función objetivo. Sin embargo, también demuestran que las funciones objetivo bajo estudio no son equivalentes.

Por un lado, la función objetivo  $R^2$  muestra la correlación entre dos series de valores. Una elevada correlación indica que la tendencia de las dos series es similar, independientemente del valor concreto del error entre el valor observado y el valor esperado. Un modelo con elevado valor de  $R^2$  puede conseguir buenas predicciones sobre valores no conocidos.

Por otro lado, la función objetivo  $RMSE$  muestra el error o diferencia entre una serie de predicciones y los datos esperados correspondientes. Al minimizar  $RMSE$  se reduce el error, por lo que la serie de predicciones se acerca a la serie esperada. La consecuencia es que se obtienen modelos cuya precisión es alta, pero cuya capacidad de predicción puede ser mala, debido al efecto de sobreajuste, también llamado *overfitting*.

Con objeto de evaluar la calidad de las soluciones obtenidas a través de una métrica diferente, se propone en este artículo la utilización de la métrica conocida como similitud del coseno. Esta medida interpreta ambas series de datos, observada y esperada, como vectores  $O$  y  $E$  respectivamente. De esta manera, se calcula el coseno del ángulo que forman ambos vectores utilizando la siguiente expresión:

$$\text{Coseno}(\theta) = \frac{\sum_i O_i E_i}{\sqrt{\sum_i O_i^2} \sqrt{\sum_i E_i^2}}$$

Si el valor del coseno es cercano a 1, entonces el ángulo que forman los vectores es cercano a  $0^\circ$ , por lo que se tratará de series similares. Si el valor es cercano a -1, entonces el ángulo es cercano a  $180^\circ$ , lo que significa que las series tienen tendencia similar, pero en sentido opuesto. En caso de que el valor del coseno esté cerca de 0, se trata de series perpendiculares, lo que se considera como la menor similitud posible. Sobre las soluciones obtenidas en los experimentos se ha calculado la similitud de coseno. La Tabla IV muestra los valores obtenidos para los experimentos donde la inicialización es aleatoria. Para cada *benchmark* se muestran los valores de las optimizaciones donde la función objetivo es  $1 - R^2$  ( $R_2$ ) y también los valores donde la función



de coste ha sido *RMSE*. Los valores presentados corresponden a las 10 ejecuciones realizadas. Como se aprecia en la tabla, los valores para *RMSE* son más altos y más consistentes, tal y como indican los valores promedio que se muestra en la última fila. En el caso de  $R^2$ , existen valores altos, cercanos al 1, pero también existen valores negativos y alguno cercano al 0, 5.

Run	spatial		tower		VF8	
	R2	RSME	R2	RSME	R2	RSME
1	0,9393	0,9931	0,9514	0,9881	0,9255	0,9758
2	0,9931	0,9666	0,96	0,9857	0,9865	0,9732
3	0,9807	0,9877	0,9541	0,9856	0,9782	0,9534
4	-0,7232	0,9962	0,9846	0,9886	0,9634	0,9736
5	0,9643	0,9897	0,9754	0,9838	0,8801	0,9434
6	0,5044	0,9698	-0,9682	0,9844	0,6949	0,9871
7	0,9608	0,9902	0,9655	0,9874	0,905	0,961
8	0,946	0,9723	-0,9826	0,9871	-0,9157	0,9535
9	0,9622	0,9905	0,9784	0,9874	0,9879	0,9487
10	0,9588	0,9914	0,9721	0,985	0,9824	0,9265
Avg.	0,7486	0,9847	0,5791	0,9863	0,7388	0,9596

Tabla IV: Similitud de coseno para los experimentos con inicialización *random*. La última fila muestra el valor promedio.

La Tabla V muestra las mismas estadísticas, pero para los experimentos donde la inicialización es *sensible*. De nuevo, la tendencia es la misma, aunque los resultados para los experimentos donde la función de coste es  $1 - R^2$  son peores en este caso. Es destacable el conjunto de resultados de *spatial*, que obtuvo el mejor promedio de  $R^2$  para la optimización *R2-Sensible* (ver Tabla II), mientras que los valores de similitud de coseno de ese mismo experimento son los peores, con dos valores por debajo de 0, 29.

Run	spatial		tower		VF8	
	R2	RSME	R2	RSME	R2	RSME
1	0,9966	0,9886	0,9391	0,9866	0,8819	0,9733
2	0,9705	0,9911	-0,965	0,9862	-0,9872	0,9673
3	0,9708	0,9914	0,9793	0,9844	-0,9889	0,9456
4	0,2871	0,982	-0,9777	0,987	0,9486	0,8021
5	0,9883	0,9743	0,9576	0,9863	0,6819	0,9566
6	0,9519	0,9762	0,9868	0,9856	0,9682	0,845
7	0,9112	0,9897	0,9611	0,9853	0,9178	0,8424
8	0,9634	0,9725	0,9687	0,9863	0,9484	0,953
9	0,957	0,9739	-0,9472	0,9863	0,9791	0,9485
10	0,1608	0,991	0,9707	0,9844	0,9199	0,9714
Avg.	0,8158	0,9831	0,3873	0,9858	0,527	0,9205

Tabla V: Similitud de coseno para los experimentos con inicialización *sensible*.

### V. CONCLUSIONES Y TRABAJO FUTURO

A la vista de los experimentos realizados en este artículo, se puede afirmar que  $R^2$  y *RMSE* no son funciones que lleven a soluciones equivalentes en todos los casos. Muy al contrario, pueden obtener resultados realmente diversos. Sin embargo, este es un estudio muy preliminar que se deberá extender en el futuro.

### AGRADECIMIENTOS

Este trabajo ha sido financiado por el Ministerio de Economía y Competitividad de España a través de los proyectos con referencias TIN2014-54806-R y TIN2015-65460-C2.

### REFERENCIAS

- [1] D. L. J. Alexander, A. Tropsha, and David A. Winkler. Beware of  $r^2$ : Simple, unambiguous assessment of the prediction accuracy of qsar and qspr models. *Journal of chemical information and modeling*, 55.7:1316–1322, 2015.
- [2] J. M. Colmenar and J. I. Hidalgo. Análisis de algoritmos de evolución gramatical en problemas de regresión simbólica. In *XII Metaheurísticas, Algoritmos Evolutivos y Bioinspirados (MAEB 2017)*, pages 901–910, Barcelona, Spain, 2017. Universitat Pompeu Fabra.
- [3] Ian Dempsey, Michael O’Neill, and Anthony Brabazon. Live trading with grammatical evolution. In *GECCO 2004 Workshop Proceedings*, Seattle, Washington, USA, 26-30 June 2004.
- [4] J. Ignacio Hidalgo, J. Manuel Colmenar, José L. Risco-Martin, Alfredo Cuesta-Infante, Esther Maqueda, Marta Botella, and José Antonio Rubio. Modeling glycemia in humans by means of grammatical evolution. *Applied Soft Computing*, (20):40–53, 2014.
- [5] José Ignacio Hidalgo, J. Manuel Colmenar, José Luis Risco-Martín, Alfredo Cuesta-Infante, Esther Maqueda, Marta Botella, and José Antonio Rubio. Modeling glycemia in humans by means of grammatical evolution. *Appl. Soft Comput.*, 20:40–53, 2014.
- [6] J. R. Koza. *Genetic Programming*. The MIT Press, Cambridge, Massachusetts, 1992.
- [7] Michael O’Neill and Conor Ryan. Grammatical evolution. *IEEE Trans. Evolutionary Computation*, 5(4):349–358, 2001.
- [8] Ludo Pagie and Paulien Hogeweg. Evolutionary consequences of coevolving targets. *Evolutionary Computation*, 5:401–418, 1998.
- [9] E. J. Vladislavleva, G. F. Smits, and D. den Hertog. Order of nonlinearity as a complexity measure for models generated by symbolic regression via pareto genetic programming. *IEEE Transactions on Evolutionary Computation*, 13(2):333–349, April 2009.
- [10] David R. White, James McDermott, Mauro Castelli, Luca Manzoni, Brian W. Goldman, Gabriel Kronberger, Wojciech Jaskowski, Una-May O’Reilly, and Sean Luke. Better GP benchmarks: community survey results and proposals. *Genetic Programming and Evolvable Machines*, 14(1):3–29, Mar 2013.