

# MÉTODO ESTADÍSTICO-EVOLUTIVO PARA LA REDUCCIÓN DE INCERTIDUMBRE EN PROCESOS DE PREDICCIÓN

Germán Bianchini<sup>†</sup>, Paola Caymes-Scutari<sup>†‡</sup>

<sup>†</sup> *Laboratorio de Investigación en Cómputo Paralelo/Distribuido (LICPaD) Departamento de Ingeniería en Sistemas de Información, Facultad Regional Mendoza - Universidad Tecnológica Nacional. (M5502AJE) Mendoza, Argentina. [gbianchini@frm.utn.edu.ar](mailto:gbianchini@frm.utn.edu.ar), [pcaymesscutari@frm.utn.edu.ar](mailto:pcaymesscutari@frm.utn.edu.ar)*

<sup>‡</sup> *Consejo Nacional de Investigaciones Científicas y Técnicas - Argentina*

Resumen: Existen diversos factores que dificultan y limitan el modelado y la implementación de fenómenos ambientales, y que impactan en la calidad de la predicción arrojada por la evaluación del modelo en función de los parámetros de entrada. Por su parte, existen diferentes métodos para la reducción de incertidumbre, los cuales permiten reducir los efectos negativos provocados por los aspectos no representables o cuantificables ya sea matemática o computacionalmente. En este trabajo se describe un método de reducción de incertidumbre basado en la estadística, la computación evolutiva y el paralelismo, denominado *Evolutionary-Statistical System* o ESS. El método es iterativo y dirigido por datos con solución múltiple solapada. Su potencia se basa en aumentar la predicción realizada por el modelo implementado en el simulador con la orientación de la búsqueda y la calibración de los resultados.

Palabras claves: *estadística, incertidumbre, algoritmos evolutivos*  
2000 AMS Subjects Classification: 65C50 - 60H35 - 68U20

## 1. INTRODUCCIÓN

Uno de los pilares más importantes de la Ciencia Computacional lo constituye la Matemática, a través de las diferentes ramas que permiten modelar diversos aspectos y fenómenos de la naturaleza. El modelado matemático, la simulación, la visualización, la programación y la computación de alto rendimiento, propician el estudio de problemas que anteriormente resultaban muy complejos, demandantes de tiempo y espacio (en términos de memoria y almacenamiento), e incluso muy peligrosos por los riesgos humanos y/o ambientales que conllevan [8]. Sin embargo, aún cuando la formulación del modelo y el algoritmo de implementación estén libres de errores, es muy difícil que el simulador ofrezca una salida siempre fiel a la situación real que se desea representar, dada la incertidumbre introducida tanto por las expresiones que pueden no responder a un modelo estrictamente matemático como por la representación numérica de la arquitectura de computadoras, ya que ambas presentan limitaciones intrínsecas a la dificultad o impedimento de medir, cuantificar y/o representar todos los aspectos o parámetros involucrados en el sistema bajo estudio. Existen diferentes métodos para la reducción de incertidumbre, los cuales permiten reducir los efectos negativos provocados por los aspectos no representables matemática ni computacionalmente. De esta manera, la reducción de incertidumbre constituye un proceso muy importante para lograr predicciones más acordes a la realidad, que permitan constituir herramientas confiables para asistir a los sistemas de toma de decisiones. El presente artículo se centra en un método de reducción de incertidumbre, denominado ESS (*Evolutionary-Statistical System* o Sistema Estadístico-Evolutivo), el cual descansa sobre tres pilares: la estadística, la computación evolutiva [5] y la computación paralela. La estadística ofrece la posibilidad de estudiar tendencias en el comportamiento del sistema bajo diversas condiciones. La computación evolutiva permite orientar la búsqueda hacia una predicción más precisa, y la computación paralela ofrece la infraestructura para explotar y desplegar el espacio de búsqueda, en tiempos humanamente razonables. Si bien el método ESS puede aplicarse sobre diversos sistemas de predicción y/o prevención que conlleven simulación de propagación, en este artículo se comenta su utilidad para la predicción de incendios forestales.

## 2. EL MÉTODO ESTADÍSTICO-EVOLUTIVO

El sistema ESS constituye un método conducido por datos de múltiples soluciones solapadas, donde cada solución se representa a través de un mapa basado en celdas. Por un lado, ante la incertidumbre existente en los parámetros de entrada del modelo, ESS requiere que para cada parámetro se indique un dominio, de forma tal que pueda variarse el valor del parámetro dentro de dicho dominio. Esta característica resulta muy importante para parámetros dinámicos. De esta manera, las posibles combinaciones en el conjunto de valores de entrada del modelo resultan numerosas, de las cuales algunas

podrán describir la realidad de forma más fiel que otras. Cada interpretación particular de los parámetros del modelo se denomina ‘escenario’, y a su vez puede considerarse un ‘individuo’ en el paradigma evolutivo. Por otro lado, ESS es un método iterativo en el sentido de que la predicción del fenómeno se realiza cada cierto intervalo de tiempo considerando el estado anterior y actual de la propagación. La observación del estado real se utiliza para calibrar la predicción y lograr predicciones más ajustadas para los pasos de simulación siguientes. En cada iteración ESS presenta dos fases generales: la fase de optimización (OS u *Optimization Stage*) que permite orientar la búsqueda hacia los escenarios que resultan más fieles a la realidad (implementada mediante un algoritmo evolutivo paralelo), y la fase de Calibración (CS o *Calibration Stage*) encargada de establecer la tendencia y proveer una predicción (fase implementada estadísticamente). A continuación se describen los tres componentes en los que se basa ESS.

## 2.1 COMPONENTE PARALELO

Dado que por su naturaleza operativa el método requiere una gran cantidad de operaciones, el mismo resulta muy demandante en tiempo. Por esta razón, se aplica un esquema de cómputo paralelo de forma tal de reducir los tiempos de ejecución. Gracias a que el procesamiento de cada escenario/individuo es independiente del de los demás escenarios, en cada iteración del método se aplica el patrón paralelo Master/Worker [6]: un proceso Master establece las combinaciones de parámetros y las envía a los Workers. Los Workers, paralelamente, llevan a cabo las simulaciones asignadas y la evaluación de la calidad del escenario, retornando los resultados al Master, el cual agrega todos los resultados individuales. Este procedimiento se realiza hasta que todos los individuos de la población hayan sido tratados.

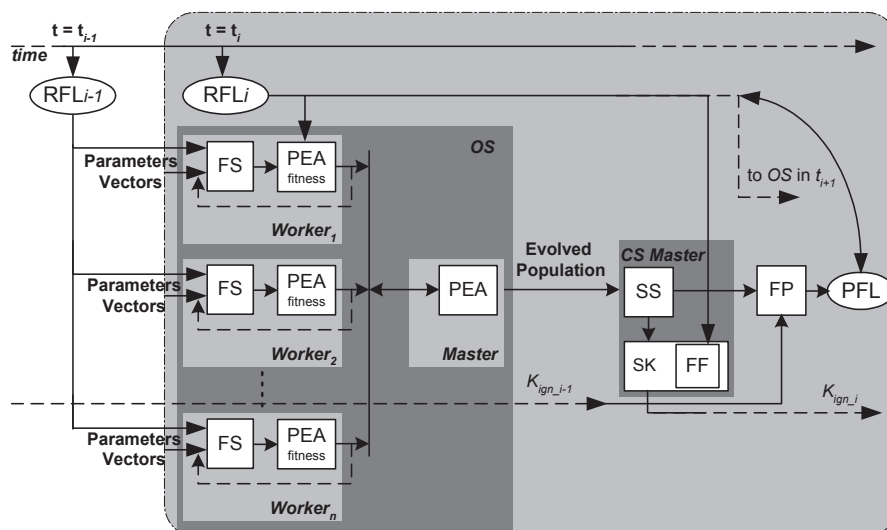


Figura 1. Diagrama de ESS (FS: Fire Simulator; PEA: Parallel Evolutionary Algorithm; OS: Optimization stage; SS: Statistical System; SK: Search  $K_{ign}$ ; FF: Fitness Function; CS: Calibration stage; FP: Fire Prediction; PFL: Predicted Fire Line, RFLX: Real Fire Line on time  $X$ ).

## 2.2 COMPONENTE EVOLUTIVO

En una versión primigenia del método, el mismo se basaba en el concepto de experimento factorial, por lo que era de esperar que cierto porcentaje de los escenarios generados no contribuyesen significativamente al resultado global, ya sea por cierto nivel de redundancia o porque la combinación de parámetros definía escenarios muy alejados de la realidad, lo cual terminaba degradando los resultados arrojados por el método. Dicho problema se solventó mediante la incorporación de un algoritmo evolutivo en la etapa de calibración, que a través de su intrínseca capacidad para ponderar mejor a los escenarios o individuos que mejor se adaptan, orienta la predicción estadística ajustándola a la realidad. Puede observarse en la figura 1 que la etapa de optimización OS itera hasta que la población alcanza una cierta calidad. Para cada individuo se realiza en paralelo la simulación y el cálculo del *fitness*. Es por ello que se distinguen dos tipos diferentes de etapas PEA: ‘PEA fitness’ implementa el rol del Worker en el proceso evolutivo (es decir la simulación del modelo y evaluación de la función de *fitness*), y ‘PEA Master’

implementa el rol evolutivo del Master (aplicación de operadores para evolucionar la población). El conjunto de escenarios resultantes (o población) constituyen la entrada de la fase estadística (cuadro SS). En esta fase se utiliza una función de fitness (FF) para evaluar el mapa de probabilidades.

### 2.2.1 FUNCIÓN DE FITNESS

La función de fitness constituye una herramienta que permite evaluar los individuos, ponderar y comparar la predicción resultante con la situación real. Dado que el simulador utiliza una aproximación basada en celdas, FF está definida como un cociente que considera las celdas del mapa de predicción obtenido a partir del conjunto de parámetros de cada individuo. La Ecuación (1) describe dicha expresión:

$$Fitness = \frac{\#celdas \cap \#celdasIgnición}{\#celdas \cup \#celdasIgnición} \quad (1)$$

donde  $\#celdas \cap$  representa la cantidad de celdas en la intersección entre los resultados de la simulación y el mapa real,  $\#celdas \cup$  es el número de celdas en la unión de los resultados de la simulación y la situación real, y  $\#celdasIgnición$  representa el número de celdas quemadas antes de comenzar la simulación. Un valor de fitness igual a uno corresponde a una predicción perfecta, dado que puede interpretarse el resultado como que el área predicha coincide con el área realmente quemada. Por otro lado, un valor de fitness igual a cero indica el máximo error, dado que en este caso el resultado no tendría ningún grado de coincidencia con la realidad.

### 2.3 COMPONENTE ESTADÍSTICO

La estadística constituye uno de los aspectos más importantes de la metodología dado que permite determinar la tendencia en la evolución o propagación del fenómeno bajo estudio. ESS toma en consideración toda la población de escenarios, donde cada escenario está determinado por un conjunto de valores de entrada para los parámetros del modelo implementado por el simulador. Los posibles escenarios corresponden a los individuos creados por el algoritmo evolutivo considerando un cierto dominio para cada parámetro, y el modelo se evalúa para cada conjunto de valores. En base a los mismos, ESS determina un patrón de comportamiento del modelo sin entrar en análisis específicos de cada escenario. La evaluación del modelo arroja para cada escenario un mapa de ignición basado en celdas. Un mapa de ignición está constituido por una matriz en la que se indica para cada elemento (o celda) si sufrió ignición o no. Los resultados (es decir los mapas de ignición) son entonces combinados o agregados para determinar la tendencia en el comportamiento del modelo. Dicho patrón de comportamiento se utiliza para predecir el siguiente paso del incendio.

El mapa de probabilidades que constituye la salida de dicha fase posee un doble propósito: por un lado, los mapas de probabilidad se utilizan como entrada de la fase SK (Search  $K_{ign}$ ) para determinar el valor de  $K_{ign}$  actual (es decir el número clave de igniciones necesario para realizar la predicción), y que será utilizado en el siguiente paso de predicción bajo la asunción de una cierta localidad espacial y temporal del fenómeno. Por otro lado, la salida de SS constituye la entrada de la etapa Fire Prediction (FP). FP posee la responsabilidad de generar el mapa de predicción tomando en cuenta el valor  $K_{ign}$  evaluado en el período previo.

#### 2.3.1 VALOR DE $K_{ign}$

El valor  $K_{ign}$  (por *Key Ignition value*) establece la probabilidad de ignición de las celdas que deben considerarse en el mapa estadístico que permite una mejor aproximación a la línea de fuego real. Para más detalles pueden consultarse [2] y [3]. La figura 1 muestra un diagrama general del sistema completo y la participación del valor  $K_{ign}$ . En general, para cada  $i$  desde 2 hasta  $n$ , en el instante  $t_i$  se solapan la etapa de predicción para el tiempo  $t_i$  y la etapa de calibración para obtener el  $K_{ign}$  a utilizar en el instante  $t_{i+1}$ .

El periodo de tiempo entre  $t_0$  y  $t_1$  constituye un caso particular, dado que el proceso de predicción carece de elementos para realizar la predicción; es por ello que en ese instante el método sólo genera el valor de  $K_{ign}$  que permita comenzar la cadena de predicciones.

### 3. CASO DE APLICACIÓN

La simulación de la propagación de incendios forestales constituye un desafío desde el punto de vista computacional, dada la complejidad que involucran los modelos, la necesidad de métodos numéricos y la eficiente administración de los recursos para obtener los resultados. En este contexto, el método presentado en este trabajo constituye una importante herramienta para la prevención y predicción de incendios forestales, dado que provee información más completa acerca del posible comportamiento del fuego y de las zonas que corren mayor peligro de ignición. El peligro depende de factores estáticos como la pendiente del terreno o el tipo particular de vegetación, pero también depende de factores dinámicos como la humedad de la vegetación y la velocidad y dirección del viento. Al producirse un incendio, es muy difícil – si no imposible– conocer de antemano las condiciones específicas bajo las que se inició el fuego, con lo cual se dificulta bastante la tarea de prever la velocidad de propagación del fuego o incluso la intensidad de la llama a fin de predecir el comportamiento del incendio. Por ello, resulta primordial reducir las fuentes de incertidumbre que atenúan la precisión de las predicciones. El sistema ha sido evaluado tomando en consideración registros de quemas reales controladas, y ha sido comparado tanto con el caso clásico de predicción como con otros métodos alternativos a la predicción clásica como son BBOF [1] en sus esquemas de Búsqueda Tabú, Simulated Annealing y Algoritmos Genéticos, así como también GLUE [7], que se basa en el método de Monte Carlo y finalmente S<sup>2</sup>F<sup>2</sup>M, que representa la versión inicial del método ESS. Se obtuvieron resultados positivos, los cuales pueden consultarse en [4].

### 4. CONCLUSIONES

A través del uso de métodos heurísticos es posible facilitar la búsqueda de soluciones a determinados problemas, como por ejemplo el caso de la incertidumbre en los parámetros de entrada de un modelo. No obstante, como hemos podido observar tras la comparación de diversas metodologías a las que se ha hecho referencia en la sección 3, el uso de tales heurísticas tampoco garantiza el descubrimiento de tales soluciones. En vista de tal situación, en el método ESS se ha conjugado la aplicación de diversas herramientas modernas como el Cómputo de Alto Rendimiento y el Paralelismo junto con el uso de Computación Evolutiva (como heurística) y Análisis Estadístico. A través de este último concepto, es posible llevar a cabo el proceso que mejora la salida del método, ya que la interpretación de la muestra que representa la población de posibles soluciones en su conjunto, ofrece en general más información que la que puede brindar cualquier individuo en particular.

### AGRADECIMIENTOS

Este trabajo ha sido financiado por CONICET a través del proyecto PIP 11220090100709, por la UTN mediante el proyecto UTN1194 y por la ANPCyT mediante el proyecto PICT PRH 2008-00242.

### REFERENCIAS

- [1] B. ABDALHAQ, *A methodology to enhance the Prediction of Forest Fire Propagation*. Ph. D Thesis. Universitat Autònoma de Barcelona (Spain), 2004.
- [2] G. BIANCHINI, M. DENHAM, A. CORTÉS, T. MARGALEF, E. LUQUE, *Wildland Fire Growth Prediction Method Based on Multiple Overlapping Solution*, Journal of Computational Science, Vol. 1-4 (2010) , pp. 229-237.
- [3] G. BIANCHINI, A. CORTÉS, T. MARGALEF, E. LUQUE, *Improved prediction methods for Wildfires using High Performance Computing: A comparison*, LNCS 3991, Part I (2006), pp. 539-546.
- [4] G. BIANCHINI, M. MENDEZ GARABETTI, P. CAYMES SCUTARI, *Evolutionary-Statistical System for Uncertainty Reduction Problems in Wildfires*. XVIII CACIC - XII WPDP. ISBN: 978-987-1648-34-4, pp.230-238, 2012.
- [5] D.E. GOLDBERG, *Genetic and evolutionary algorithms, Come of age*, Communications of the ACM, Vol 37-3 (1994), pp. 113-119.
- [6] T. MATTSON, B. SANDERS, B. MASSINGILL, *Patterns for Parallel Programming*, Addison-Wesley, 2004.
- [7] J. PIÑOL, R. SALVADOR, K. BEVEN, *Model Calibration and uncertainty prediction of fire spread*. Forest Fire Research & Wildland Fire Safety. On CD-ROM, Millpress, 2002.
- [8] A. SHIFLET, G. SHIFLET, *Introduction to Computational Science – Modeling and Simulation for the Sciences*, Princeton University Press, 2006.