

# Monitorización y Análisis de Métricas Poblacionales en un Método de Predicción de Incendios Forestales

María Laura Tardivo<sup>1,2,3</sup>, Paola Caymes-Scutari<sup>1,2</sup>, Germán Bianchini<sup>1</sup>, Miguel Méndez-Garabetti<sup>1,2</sup>

<sup>1</sup>Laboratorio de Investigación en Cómputo Paralelo/Distribuido (LICPaD) – Universidad Tecnológica Nacional (UTN-FRM) – Mendoza – Argentina

<sup>2</sup>Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)

<sup>3</sup>Departamento de Computación – Fac. de Cs. Exactas- Físico- Químicas y Naturales- Universidad Nacional de Río Cuarto (UNRC) – Río Cuarto, Córdoba – Argentina

lauratardivo@dc.unrc.edu.ar,  
{pcaymesscutari,gbianchini}@frm.utn.edu.ar,  
miguelmendezgarabetti@gmail.com

**Resumen.** *La incertidumbre sobre el valor de los parámetros de entrada que alimentan los modelos de predicción de incendios forestales es un problema crucial, que puede producir graves consecuencias si la salida de los modelos proporcionan predicciones erróneas. ESSIM-DE(r) es un método general de predicción y reducción de incertidumbre basado en Análisis Estadístico, la metaheurística Evolución Diferencial y Cómputo de Alto Rendimiento, y ha sido estudiado durante los últimos años con el fin de mejorar su rendimiento. Este trabajo presenta la incorporación del proceso de sintonización dinámica mediante el cual se perfecciona la estrategia evolutiva del método, en base a la monitorización y análisis de la distribución poblacional y de los movimientos efectivos del optimizador.*

## 1. Introducción

Año tras año los incendios forestales afectan numerosas áreas forestales, causando importantes pérdidas materiales, en la flora y en la fauna. Los incendios son aún más propensos a desarrollarse en temporada estival, donde las temperaturas elevadas, las bajas precipitaciones y la sequía prolongada constituyen el contexto ideal para el desarrollo de estos desastres naturales tan severos. Entre Noviembre de 2017 y Febrero de 2018 más de 600.000 hectáreas fueron arrasadas por el fuego y más de 1.000 animales murieron a causa de los incendios forestales en la provincia de La Pampa, Argentina [La Nacion 2018].

En este contexto, es preciso remarcar la importancia de contar con herramientas computacionales que permitan colaborar en los planes de prevención y predicción de incendios. Sin embargo, predecir el comportamiento de estos desastres naturales no es una tarea sencilla. Por un lado, existe una complejidad inherente al fenómeno físico en sí mismo, un sistema natural de interacciones complejas que debe ser simulado en un entorno discreto y computacional. Por otro lado, el avance de un incendio está condicionado por un gran número de variables, como las características del viento (velocidad y dirección), la topografía del terreno (pendiente, inclinación) o las características del material combustible predominante (la humedad o el tipo de material), etc. Algunas de estas variables poseen valores que cambian dinámicamente durante el desarrollo del incendio, mientras que otras poseen valores que suelen ser

desconocidos o difíciles de cuantificar. Esta falta de precisión sobre el valor de las variables se denomina incertidumbre, e impacta negativamente sobre las predicciones.

En los últimos años se han desarrollado, estudiado y mejorado una serie de métodos que se enfocan en reducir este impacto negativo que provoca la incertidumbre. Tal es el caso del método *Evolutionary Statistical System with Island Model and Differential Evolution with population Reinitialization*, ESSIM-DE(r), el cual utiliza Estadística para obtener el patrón de comportamiento de la línea de fuego, el algoritmo evolutivo denominado Evolución Diferencial con reinicio de poblaciones, para orientar la búsqueda hacia mejores soluciones, y Computación de Alto Desempeño con Modelo de Islas y un esquema de procesos de doble jerarquía, lo que permite explorar un amplio espacio de búsqueda y a la vez acelerar el proceso de predicción. ESSIM-DE(r) pertenece a la clase de métodos denominados Métodos Conducidos por Datos con Múltiples Soluciones Solapadas (*Data Driven Methods with Multiple Overlapped Solutions, DDM-MOS*), los cuales obtienen predicciones de la línea de fuego en base a la agregación de múltiples soluciones. El método ESSIM-DE(r) ha sido aplicado satisfactoriamente en la predicción de incendios forestales y ha logrado una buena calidad de predicción comparado contra otros DDM-MOS [Tardivo *et al*, 2018].

ESSIM-DE(r) ha sido estudiado durante el último tiempo con el fin de mejorar aquellos aspectos que pueden ser limitantes de mejor rendimiento, tanto en términos de calidad como de tiempo en la obtención de predicciones. En tal sentido, nos hemos focalizado en dotar a ESSIM-DE(r) de estrategias automáticas de sintonización, las cuales permitan calibrar en tiempo de ejecución algún aspecto crítico del método, con dos fines principales: por un lado, se pretende obtener mejor calidad de las predicciones de manera transparente al usuario final, y por otro lado, se desea que esta sintonización sea realizada independientemente del contexto de ejecución considerado.

La sintonización dinámica y automática [Naono *et al* 2010] se representa en la Figura 2. En la etapa de **instrumentación** se incluyen directivas en el código fuente para registrar ciertas métricas de interés. Dichas métricas corresponden a un modelo del comportamiento de la aplicación, y definen conocimiento específico acerca de la misma. En la etapa de **monitorización** se registran las métricas, para su posterior análisis y sintonización. En la etapa de **análisis** se procesan los valores registrados y, en base este conocimiento específico, se determinan las acciones de sintonización necesarias para mejorar el rendimiento de la aplicación. Posteriormente, en la etapa de **sintonización** se aplican las acciones definidas tendientes a mejorar los aspectos críticos. Todas estas etapas se realizan en tiempo de ejecución.



**Figura 2. Proceso de sintonización.**

Es importante enfatizar que la incorporación de estrategias de sintonización dinámicas y automáticas puede mejorar aquellos aspectos limitantes de mejor rendimiento, y a la vez abstraer al usuario no experto de los detalles y las dificultades tecnológicas de métodos complejos como ESSIM-DE(r). En la literatura se han propuesto algunas estrategias de sintonización aplicadas a métodos paralelos/evolutivos, principalmente relacionadas con aquellos aspectos afines al modelo de cómputo paralelo

[Caymes Scutari *et al*, 2016]. En este trabajo nos hemos focalizado en definir un modelo de rendimiento que permita calibrar dinámicamente un aspecto particular del funcionamiento de ESSIM-DE(r). El mecanismo de sintonización que proponemos permite controlar dinámicamente un parámetro clave del método, con el fin de lograr un rendimiento acorde a las condiciones y al contexto de ejecución considerado. El primer paso requerido para definir una estrategia de sintonización dinámica lo constituye la definición de las métricas que se registrarán en la etapa de monitorización. En este caso, hemos prestado especial interés en una de las etapas más destacadas de ESSIM-DE(r), constituida por el proceso evolutivo llevado a cabo por el algoritmo Evolución Diferencial [Storn *et al* 2005]. Esta metaheurística es un optimizador poblacional (basado en múltiples soluciones) y consiste en un proceso iterativo en el cual se evoluciona una población de individuos, cada uno de los cuales representa una solución al problema. En el caso particular de los incendios forestales, cada individuo representa una combinación de valores para las variables que determinan el avance del fuego, como la velocidad de viento, la dirección y pendiente del terreno, el tipo de material combustible, la humedad del material combustible vivo, etc. Un posible problema de rendimiento asociado a Evolución Diferencial lo constituye el estancamiento en óptimos locales y/o la convergencia prematura hacia un óptimo local [Lampinen y Zelinka, 2000, Yang *et al* 2015]. En este trabajo se establece un criterio mediante el cual es posible sintonizar la cantidad de generaciones por las cuales han de evolucionar las poblaciones a fin de detectar de manera anticipada una tendencia al estancamiento y/o a la convergencia prematura, y así evitar que estos dos fenómenos afecten de manera negativa el rendimiento del método, tanto en la calidad de las predicciones como en el tiempo de ejecución. Se presentan dos métricas poblacionales que han sido definidas para cuantificar dichos problemas, las cuales son registradas y analizadas mediante dos nuevas etapas incorporadas al proceso de predicción de ESSIM-DE(r), y a partir de las cuales se definen acciones de sintonización para mejorar estos aspectos.

Este trabajo representa el soporte experimental para sustentar la definición de un modelo de rendimiento que permita sintonizar a ESSIM-DE(r) de un modo transparente al usuario final. El trabajo se estructura de la siguiente forma. La sección 2 introduce el método ESSIM-DE(r), describiendo las diferentes etapas del proceso de predicción. También se incluye una breve descripción al algoritmo Evolución Diferencial, del operador de reinicio poblacional de ESSIM-DE(r), y se presentan los problemas del estancamiento y la convergencia prematura. La sección 3 describe dos métricas definidas para identificar si la población posee una tendencia hacia dichos problemas y describe una mejora del método ESSIM-DE(r), la cual incorpora la capacidad de monitorización y análisis de dichas métricas, y las acciones de sintonización en base al análisis del modelo de rendimiento definido. La sección 4 presenta los resultados obtenidos con tres casos de quemaduras controladas reales. Los resultados son analizados desde el punto de vista de la calidad de la predicción y respecto del tiempo de ejecución del método. Finalmente, la sección 5 presenta las conclusiones obtenidas y la proyección del trabajo a futuro.

## **2. Predicción de Incendios con ESSIM-DE(r)**

Para realizar la predicción de la propagación del fuego, en ESSIM-DE(r) el desarrollo total del incendio es dividido en diferentes instantes de tiempo discretos, denominados

pasos de simulación. En cada uno de estos instantes operan en ESSIM-DE(r) cuatro etapas principales: la **Etapa de Optimización**, la **Etapa Estadística**, la **Etapa de Calibración**, y la **Etapa de Predicción**. ESSIM-DE(r) es un método general de reducción de incertidumbre con un esquema paralelo de Múltiples Poblaciones y Migración, o Modelo de Islas. Por lo tanto, cada una de estas etapas es llevada a cabo por un grupo de procesos, organizados en dos niveles, como se puede observar en el diagrama general de ESSIM-DE(r) de la Figura 4. En una jerarquía superior se dispone de un proceso denominado *Monitor*, que es el encargado de enviar información de inicialización a cada isla, recolectar los datos procesados en la etapa final de la simulación y determinar los valores de salida, identificando cuál es la isla con mejor aproximación al incendio real. Se puede observar que el componente *Monitor* está relacionado con cada componente *Isla* a través de la comunicación antes mencionada. Por su parte, el nivel inferior de jerarquía responde a un esquema de isla. El proceso *Master* de cada isla coordina las acciones llevadas a cabo con la población de su isla y delega cierto cómputo a procesos *workers*. Además, las islas intercambian individuos entre sí, en una etapa de migración llevada a cabo por los *Masters* de cada isla. La **Etapa de Optimización** se concreta en el *Master* y los *workers* de cada isla mediante dos sub-etapas:  $EO_M$  (Etapa de Optimización en *Master*) y  $EO_W$  (Etapa de Optimización en *workers*). Estas etapas permiten evolucionar una población de individuos en base al optimizador Evolución Diferencial.

## 2.1. Evolución Diferencial

El algoritmo Evolución Diferencial (*Differential Evolution, DE*) [Storn *et al* 2005] es un optimizador estocástico basado en múltiples soluciones, denominadas individuos, los cuales están organizados en una población. DE comienza generando una población inicial de individuos con valores aleatorios dentro de los rangos del problema. Cada individuo está representado por un vector de números reales y codifica una posible solución. Los individuos evolucionan a través de las generaciones (o iteraciones) por medio de los operadores de mutación, cruzamiento y selección.

El gráfico de la Figura 3 describe, en rasgos generales, la entrada y la salida de cada etapa evolutiva de DE. En cada generación  $g$  se recorre la población y por cada individuo  $X_{i,g}$  el proceso de mutación consiste en seleccionar al azar tres individuos de la población diferentes entre sí ( $X_{r1,g}$ ,  $X_{r2,g}$  y  $X_{r3,g}$ ). Estos tres individuos son utilizados en un operador de recombinación que realiza diferencia de vectores para generar un nuevo vector  $V_{i,g}$ , denominado *vector mutante*.

En la etapa de cruzamiento, el vector mutante es combinado con el individuo  $X_{i,g}$  para generar un nuevo vector  $U_{i,g}$ , denominado *vector trial*. Este vector posee características de ambos progenitores. La herencia es regulada por un parámetro del algoritmo denominado probabilidad de cruzamiento y por un tipo de cruzamiento, el cual determina qué porciones del progenitor serán heredadas. Dependiendo de la selección de los individuos que participan en la mutación y del tipo de cruzamiento, en DE se definen diferentes estrategias de mutación. Una de las más utilizadas se denomina '*DE/rand/1/bin*', indicando que uno de los individuos participantes en la mutación se seleccionará de manera aleatoria y el tipo de cruzamiento es binomial (para mayor detalle en las estrategias de mutación y los tipos de cruzamiento puede consultar Storn *et al*, 2005).

Finalmente, en la etapa de selección se determina el mejor candidato entre el individuo corriente  $X_{i,g}$  y el vector trial  $U_{i,g}$ . Aquel que tenga mejor valor respecto de la función a optimizar será el que sobrevivirá a la generación siguiente. Es decir, siendo  $f$  la función a optimizar,  $U_{i,g}$  reemplaza a  $X_{i,g}$  si y solo si  $f(U_{i,g}) < f(X_{i,g})$  (para problemas de minimización); caso contrario, no ocurre reemplazo o selección.

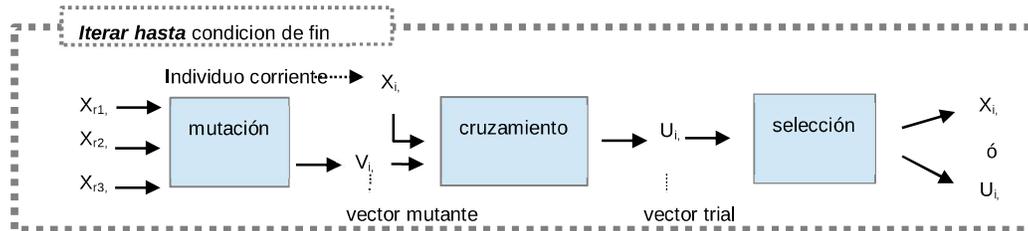


Figura 3. Esquema general del algoritmo Evolución Diferencial

## 2.2. Etapas de ESSIM-DE(r)

El proceso *Master* es el encargado de inicializar la población y aplicar los operadores evolutivos de mutación y cruzamiento de DE para generar nuevos individuos (etapa  $EO_M$ ). Para determinar si los nuevos individuos generados permanecerán en la generación siguiente, el *Master* distribuye los individuos entre ciertos procesos *workers*, quienes evalúan el desempeño de cada uno. En la Etapa de Optimización llevada a cabo por los procesos *workers* ( $EO_W$ ) se realiza la evaluación de la aptitud de los individuos mediante dos subetapas internas denominadas Simulación de Incendios (*SI*) y Evaluación de aptitud con Evolución Diferencial ( $DE_{apt}$ ). *SI* debe ser alimentado con la línea de fuego real del incendio en el instante de tiempo  $t_{i-1}$  ( $LFR_{i-1}$ ), junto con un escenario de quema, o combinación de las variables que describen el contexto del incendio (*esc*).

Cuando *SI* concluye con la simulación del individuo, el mapa de quema resultante de la simulación es introducido en la etapa  $DE_{apt}$  para comparar el mapa obtenido de la simulación contra el mapa real del instante de tiempo  $t_i$  ( $LFR_i$ ). Aquí se computa la diferencia entre ambos mapas, y se calcula el valor de aptitud para cada individuo. La comparación se realiza mediante una función denominada función de *fitness*, la cual se basa en el índice Jaccard [Real y Vargas 1996] y determina la coincidencia de celdas entre ambos mapas, computada como lo indica la expresión (1), en donde A representa el conjunto de celdas en el mapa real sin el subconjunto de celdas quemadas antes de iniciar el proceso de predicción, y B representa el conjunto de celdas en el mapa predicho sin el subconjunto de celdas quemadas antes de iniciar la predicción. Por lo tanto, el valor de *fitness* estará comprendido en el rango  $[0,1]$ , donde un valor igual a 1 representa una predicción perfecta, y un valor igual a cero indica el máximo error.

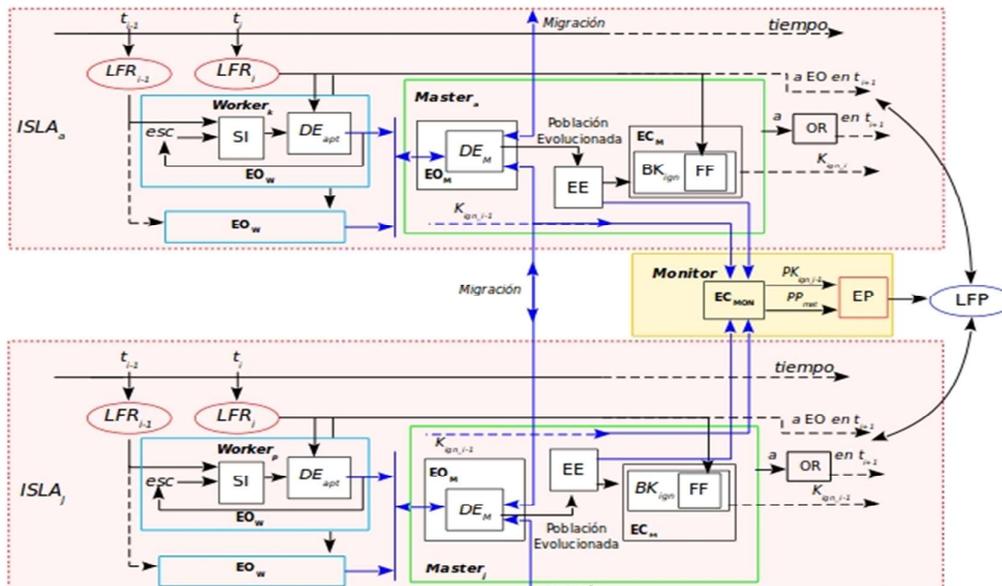
$$fitness = \frac{A \cap B}{A \cup B} \quad (1)$$

La función de *fitness* en ESSIM-DE(r) se utiliza con dos propósitos principales. Como hemos mencionado, por un lado se utiliza a nivel  $EO_W$  para evaluar la aptitud de cada individuo, comparando el mapa obtenido por la simulación de cada uno de ellos

respecto de la línea de fuego real del instante considerado  $t_i$  (**LFR<sub>i</sub>**). Por otro lado, se utiliza a nivel de **EC<sub>M</sub>** y **EC<sub>MON</sub>**, como función para evaluar la calidad de una predicción: una vez que ESSIM-DE(r) obtiene la predicción de la línea de fuego, en base a la agregación de todos los mapas de quema generados en la isla, y una vez que se conoce el estado real del incendio, se pueden comparar ambos mapas para analizar la calidad de la predicción utilizando la función de *fitness* de la expresión (1).

A medida que los *workers* van obteniendo los resultados de la etapa **EO<sub>w</sub>**, envían los mapas de quema obtenidos y su respectivo valor de aptitud al **Master**, quien determina si los individuos que han evolucionado sobrevivirán a la siguiente generación. Para ello, verifica si la aptitud calculada por los *workers* es mayor a la aptitud del individuo original, en cuyo caso, el nuevo individuo generado reemplaza al individuo original. Este proceso se realiza para cada individuo de la población, en todas las islas. El **Master** también se encarga de efectuar la migración grupal de individuos. En la etapa de **migración** se selecciona un conjunto de individuos de la población actual para ser enviados a otra isla. Entre los individuos seleccionados se incluyen los dos mejores de la población (aquellos dos que tengan los mejores valores de aptitud) y el resto del conjunto es seleccionado al azar. La comunicación entre islas sigue un orden topológico de anillo. Los individuos que arriban a una isla reemplazan a los peores individuos de la población destino.

Una vez que la población completa las diferentes generaciones evolutivas, es introducida en la Etapa Estadística (EE). En esta etapa se construye una matriz de probabilidades, cuyas celdas indican la probabilidad de ignición, es decir, la probabilidad de que cada celda sea alcanzada por el fuego respecto de todos los posibles escenarios de la población. Dicha probabilidad se calcula utilizando la expresión (2), donde  $n$  es el número total de escenarios y  $n_c$  es el número de escenarios en el cual la celda  $c$  fue alcanzada por el fuego. Para comprender mejor este proceso, consideremos la Figura 5. En este ejemplo se cuenta con cuatro escenarios diferentes ( $n=4$ ). Cada escenario es introducido al simulador de incendios y se obtiene un mapa de quema. Todos ellos son combinados en una matriz que incluye la probabilidad de ignición para cada celda  $y$ , por lo tanto, considera la contribución de cada escenario respecto del número total de escenarios.



**Figura 4. Diagrama general de ESSIM-DE(r).** SI: Simulador de Incendios,  $De_{apt}$ : Evolución Diferencial (evaluación de aptitud),  $EO_M$ : Etapa de Optimización (Master),  $EO_w$ : Etapa de Optimización (Worker),  $esc$ : escenario,  $LFR_t$ : línea de fuego Real del instante  $t$ ,  $EE$ : Etapa Estadística,  $BK_{ign}$ : Búsqueda del valor clave de ignición,  $FF$ : Función de *Fitness*,  $K_{ign,t}$ : Valor clave de ignición del instante  $t$ ,  $EC_{MON}$ : Etapa de Calibración (Monitor),  $EP$ : Etapa de Predicción,  $LFP$ : Línea de fuego predicha,  $OR$ : Operador de Reinicio poblacional.

A partir del análisis de la matriz de probabilidades es posible identificar cuál es la tendencia de la línea de fuego en una Etapa de Calibración ( $EC_M$ ). Esto se puede lograr seleccionando el conjunto de celdas con una probabilidad menor o igual a cierto valor. Para ello, se define un mapa de probabilidad  $P_K$  como conjunto de celdas cuyo valor  $P_{ign}$  es mayor o igual que  $K$ . En la Figura 5 se pueden observar los diferentes mapas de probabilidad  $P_K$  obtenidos para el ejemplo. El siguiente paso consiste en comparar el mapa real con cada mapa de probabilidades  $P_K$ . El objetivo es encontrar un valor particular de  $P_{ign}$ , cuyo mapa de probabilidad asociado sea el que mejor represente el avance real del fuego. Para ello se realiza la búsqueda de lo que se denomina valor clave de ignición  $K_{ign}$  ( $BK_{ign}$ ). Este valor será utilizado para realizar las predicciones en el siguiente instante de tiempo. La evaluación de la aptitud de cada mapa de probabilidad se lleva a cabo en la etapa que implementa la Función de *Fitness* ( $FF$ ), en  $EC_M$ .

$$P_{ign}(c) = \frac{n_c}{c} \quad (2)$$

La Etapa de Predicción ( $EP$ ) se lleva a cabo en el Monitor en la etapa  $EC_{MON}$ . Para ello, la salida de  $EE$  se envía a  $EC_{MON}$  junto con los  $j$  valores de  $K_{ign,i-1}$  calculados en el instante de tiempo anterior por las  $j$  islas. En la Etapa  $EC_{MON}$  el proceso *Monitor* selecciona el mejor valor de  $K_{ign}$  entre los obtenidos por cada isla. El valor hallado  $K_{ign}$  es utilizado por el monitor para realizar la predicción ( $LFP$ ). Como se puede observar, para obtener la línea de fuego de predicción se hace necesario contar

con el valor clave de ignición del instante de tiempo anterior. Por este motivo, en ESSIM-DE(r), así como en todos los DDM-MOS, no se pueden realizar predicciones en el primer instante de tiempo, puesto que en el primer paso se requiere calcular el primer valor de  $K_{ign}$  a ser usado en el segundo paso. El funcionamiento detallado de los niveles de ESSIM-DE(r) se resume en dos Etapas de Optimización ( $EO_w$  y  $EO_M$ ), una Etapa Estadística ( $EE$ ), dos etapas de calibración ( $EC_M$  y  $EC_{MON}$ ) y una Etapa de Predicción ( $EP$ ).

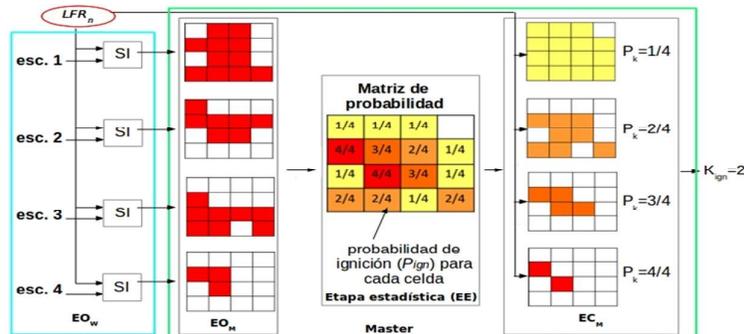


Figura 5. Etapa Estadística y Etapa de Calibración en una isla

### 2.3. Operador de reinicio poblacional y problemas de rendimiento

Una vez que se obtuvo la predicción para el paso de simulación considerado, ESSIM-DE(r) avanza hacia el siguiente paso de predicción. Es importante destacar que ESSIM-DE(r) cuenta con un operador denominado **reinicio poblacional (OR)**. Este operador previene el estancamiento global de DE, al generar un nuevo espacio de soluciones independientes de las anteriores al inicio de cada paso de simulación. Por lo tanto, el proceso de predicción comienza con una población completamente distribuida al inicio de cada instante de tiempo en los cuales se divide la duración del incendio total. El operador de reinicio de poblaciones permite perfeccionar los individuos de acuerdo paso considerado, y se ha logrado mejorar la calidad de las predicciones para la mayoría de los casos de prueba utilizados en la experimentación [Tardivo *et al* 2018].

También es muy importante remarcar que el proceso **Master** de cada isla es el que determina el fin de la Etapa de Optimización, es decir, es quien decide cuándo la población frena su evolución para continuar con las siguientes etapas del proceso de predicción: la Etapa Estadística ( $EE$ ), la Etapas de Calibración (llevada a cabo en la  $EC_M$  y  $EC_{MON}$ ) y la Etapa de Predicción ( $EP$ ), para luego avanzar hacia el siguiente paso de predicción. En ESSIM-DE(r) la condición de fin del ciclo evolutivo está determinada por un parámetro del método, el cual regula la cantidad de iteraciones evolutivas máximas que debe realizar el **Master** en la Etapa de Optimización. Este parámetro (también denominado límite de iteraciones o cota máxima de iteraciones) tiene una doble influencia en el proceso de optimización, puesto que limita la cantidad de generaciones evolutivas por las cuales evolucionará cada población, y en consecuencia, determina el tiempo de ejecución del proceso evolutivo.

En ESSIM-DE(r), la configuración de este parámetro tiene el mismo valor para todos los pasos de simulación. Sin embargo, esta configuración presenta varias desventajas. En general, no es factible conocer anticipadamente cuántas generaciones

deben ser consideradas adecuadas como cota máxima de iteraciones. Si se utiliza un valor bajo como cota, puede suceder que no se logre evolucionar suficientemente la población, quedando zonas potencialmente significativas del espacio de búsqueda sin explorar y, en consecuencia, perjudicando la calidad de las predicciones. De manera opuesta, si se utiliza un valor muy alto la población evoluciona demasiadas iteraciones, provocando pérdida de la diversidad de la población, la consecuente pérdida de la capacidad del optimizador para producir mejores individuos, y un incremento significativo del tiempo de cómputo total. En el contexto de la predicción de incendios, este último problema es crucial, puesto que se pretende obtener predicciones a corto plazo.

Por todos estos motivos, la definición de las métricas de monitorización para el modelo de rendimiento definidas en este trabajo están asociadas a esta condición de fin de ciclo evolutivo que fuera definida para ESSIM-DE(r). Se propone que el valor de cota máxima de iteraciones pueda ser computado dinámicamente para cada paso de predicción. Esta hipótesis se fundamenta principalmente debido a que el proceso de optimización está condicionado por múltiples factores: depende de la velocidad de convergencia del método (condicionada por la probabilidad de cruzamiento y el factor de mutación), de la diversidad de la población en los diferentes pasos de simulación, de la frecuencia de migración (debido a su influencia en la diversidad de la población), del tipo de reemplazo, entre otros. En consecuencia, sería una ventaja contar con un mecanismo de sintonización dinámica para la condición del fin de ciclo del proceso evolutivo, de manera tal que en tiempo de ejecución se auto-determine la cantidad de iteraciones que se deben realizar en cada paso de predicción.

#### **4. Definición de Métricas Poblacionales**

Como ha sido mencionado en la sección anterior, hemos prestado especial interés en uno de los parámetros claves de ESSIM-DE(r). Se desea identificar una nueva condición para establecer dinámicamente el fin del ciclo evolutivo. La meta consiste en determinar algún criterio por el cual se pueda detectar en la Etapa de Optimización cuándo la población ha evolucionado por suficientes generaciones, sin llegar a un estado de estancamiento o de convergencia prematura. El estancamiento es la situación en la cual el optimizador no es capaz de generar nunca más una nueva solución mejor que la anterior, pero aún la población no ha convergido. Esto significa que, aún cuando la diversidad de la población no sea pobre, el optimizador es incapaz de encontrar mejores soluciones. Por su parte, la convergencia prematura es la situación en la cual la población converge a un óptimo local (en nuestro caso, un máximo local), debido a la pérdida de diversidad [Lampinen y Zelinka 2000]. El problema del estancamiento y de la convergencia prematura empeora cuando se utilizan algunos mecanismos de control adaptativos para el tamaño de la población [Yang *et al* 2015]: incrementar el tamaño de la población decreta la posibilidad de encontrar direcciones de búsqueda correctas, mientras que decrementar el tamaño aumenta las posibilidades para una convergencia prematura y un estancamiento.

El problema del estancamiento de la población también depende de los movimientos efectivos del optimizador Evolución Diferencial. Cuando se genera un nuevo individuo se produce un movimiento en el espacio de búsqueda que representan los individuos de la población. Este movimiento se considera efectivo si el nuevo individuo generado posee mejor valor de aptitud respecto de su antecesor. De todos los

posibles movimientos que se realizan en la población, algunos son efectivos, mientras que otros no lo son, y por lo tanto estos últimos implican un esfuerzo computacional en vano.

Como es posible notar, estos problemas son determinantes para lograr efectividad del optimizador. Por lo tanto, la detección anticipada de una tendencia hacia estos problemas puede ser una buena medida para determinar el momento en que la población es apta para pasar a la siguiente etapa.

### 3.1. Diversidad poblacional y movimientos efectivos

Para abordar los problemas del estancamiento de la población y de la convergencia prematura se propone cuantificar dos métricas diferentes para la población:

- Movimientos efectivos (métrica **ME**): cuantifica el porcentaje de individuos que luego de un ciclo evolutivo han sido mejorados, es decir, cuyo valor de aptitud es mejor que el su predecesor.
- Diversidad poblacional (métrica **DP**): cuantifica la dispersión de la población, es decir, la variabilidad de la distribución de la población.

Los movimientos efectivos, o métrica **ME**, pueden ser computados de manera sencilla, contabilizando aquellos individuos que mejoran a sus antecesores. Respecto de la métrica **DP**, existen diferentes enfoques que han sido utilizados en la literatura como medida de la distribución en Estrategias Evolutivas. Uno ellos es el enfoque basado en la cuantificación genotípica, es decir, la medición de la dispersión respecto de las variables de los individuos (dimensión, o genotipo) que codifican a cada individuo. Dentro de esta categoría encontramos las mediciones de distancia en el espacio. Ursem (2009) analiza cada dimensión de los individuos, y se computa el promedio de los valores de cada dimensión. La dispersión está asociada a la distancia euclidiana de un individuo respecto a este valor promedio. Herrera y Lozano (2003) utilizan la distancia relativa a cada dimensión respecto del mejor individuo de la población. Por su parte, Yang *et al* (2013) computan la distancia de las dimensiones de todos los individuos, tomados de a dos. Otros enfoques utilizan métricas estadísticas, como la desviación estándar, también computada a nivel de las dimensiones del individuo, e identifican que una dimensión está estancada cuando el valor de desviación estándar computado alcanza un cierto valor de umbral [Yang *et al* 2015].

Los trabajos antes mencionados realizan el cómputo de la distribución poblacional con un enfoque basado en el genotipo, es decir, en los valores de las variables representando las dimensiones del problema considerado. ESSIM-DE(r) es un DDM-MOS, y como tal basa sus predicciones en la contribución de múltiples escenarios, sin la existencia de requisito alguno respecto de las características particulares de las dimensiones o variables que componen a los individuos. Por lo tanto, el enfoque elegido en este trabajo fue el cómputo de la distribución poblacional considerando la aptitud de los individuos de la población, es decir, en base al fenotipo de los individuos.

La métrica estadística elegida es el Rango Intercuartil (*Inter Quartile Range*, *IQR*) [Healey 2007] de los valores de aptitud de todos los individuos de la población, el cual representa una métrica de dispersión, y se define como la distancia desde el tercer al primer cuartil, como indica la expresión (3).

$$IQR = Q_3 - Q_1 \quad (3)$$

Como es posible notar, el cómputo del Rango Intercuartil es una medida que considera la distribución del 50% de los datos. A diferencia del cómputo de otras métricas, como el promedio de valores de aptitud, o la desviación estándar de los valores de aptitud, el cómputo de  $IQR$  es menos sensible a los valores atípicos (*outlayers*). Por lo tanto, es considerado un estimador más robusto que los antes mencionados y, como tal, una buena medida para considerar la distribución de los datos [Healey 2007].

### 3.2. Detección de pérdida de la distribución poblacional.

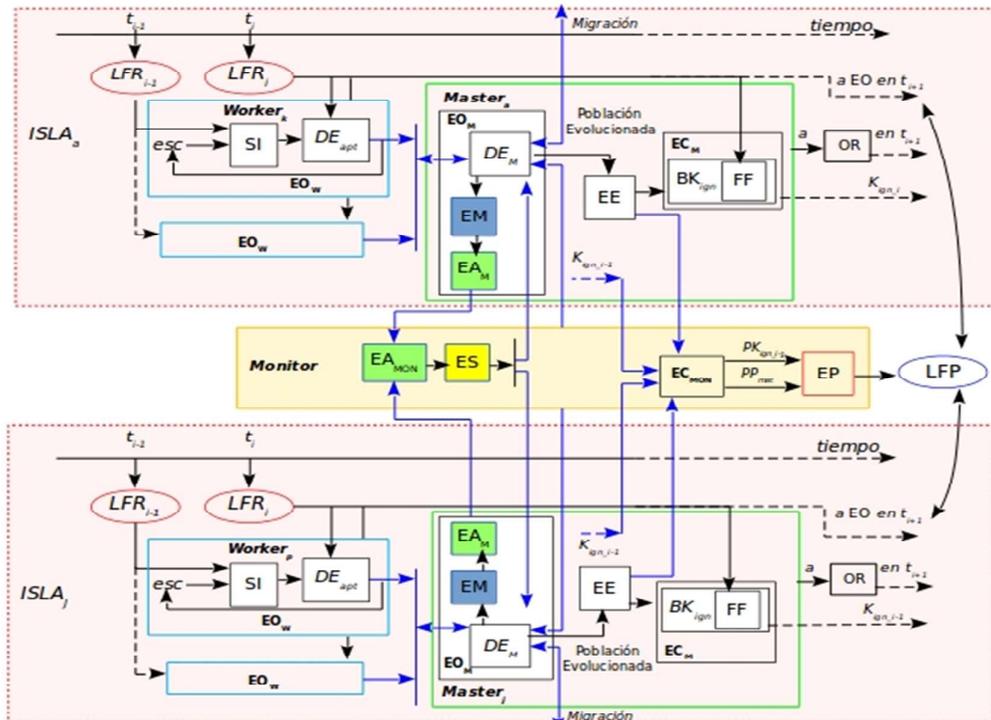
Las métricas establecidas cuantifican, por un lado, el Rango Intercuartil  $IQR$  computado a partir de los valores de aptitud de los individuos de la población, y por otro lado, los movimientos efectivos del optimizador. En este punto, cabe recordar que la convergencia prematura es la situación en la cual la población converge a un máximo local, debido a la pérdida de diversidad. La métrica  $IQR$  permite cuantificar la distribución fenotípica de la población en base a la medición de la dispersión de los valores de aptitud de los individuos, y por lo tanto puede servir como indicador de convergencia. Por su parte, el estancamiento es la situación en la cual, aún existiendo diversidad en la población, el optimizador es incapaz de encontrar mejores soluciones. La métrica  $IQR$  no es suficiente para cuantificar el estancamiento, y es por ello que para detectar este problema se incluye la medición de los movimientos efectivos del optimizador a través de la métrica  $ME$ . En consecuencia, se puede considerar que la población está estancada y/o converge a un máximo local si se cumple la condición de la expresión (4).

$$IQR = 0 \wedge ME = 0 \quad (4)$$

Como es posible observar, un valor  $IQR$  igual a cero implica que el 50% de los individuos de la población tienen el mismo valor de aptitud. Paralelamente, considerar que el optimizador no produzca ningún movimiento efectivo ( $ME=0$ ) implica que ha perdido por completo su capacidad de generar mejores soluciones candidatas. Desde la literatura [Yang *et al* 2015] se afirma, y hemos corroborado empíricamente, que identificar total estancamiento y/o convergencia requiere un tiempo de cómputo elevado. Por lo tanto, es posible suavizar la condición de la expresión (4), utilizando valores de umbral para cada métrica, como lo indica la expresión (5), donde los valores  $umbral\_IQR$  y  $umbral\_ME$  deben ser establecidos en base a la experimentación.

$$IQR \geq umbral\_IQR \wedge ME \geq umbral\_ME \quad (5)$$

Como es posible notar, el valor  $umbral\_IQR$  pertenece al rango  $[0,1]$ , puesto que se computa como diferencia de elementos de la muestra, es decir, valores de aptitud. Por su parte, el  $umbral\_ME$  pertenece al rango  $[0, NP]$ , siendo  $NP$  el tamaño de la población. Para ambas métricas, los valores de umbral a establecer serán considerados como porcentaje, de manera tal que el valor  $umbral\_IQR$  es ponderado al porcentaje respecto del máximo valor de aptitud (valor igual a 1), y el valor  $umbral\_ME$  representa el porcentaje de los individuos de la población que han sido efectivamente mejorados en una determinada generación. En caso de que la condición de la expresión (5) se cumpla en alguna isla, se considera que la población debe finalizar el ciclo evolutivo, y se puede proceder con el resto de las etapas de ESSIM-DE(r).



**Figura 7. Diagrama general de ESSIM-DE(r)-IQR-ME.** EM: Etapa de Monitorización. EA<sub>M</sub>: Etapa de Análisis (Master), EA<sub>MON</sub>: Etapa de Análisis (Monitor), ES: Etapa de Sintonización, SI: Simulador de Incendios, DE<sub>apl</sub>:Evolución Diferencial (evaluación de aptitud). EO<sub>M</sub>: Etapa de Optimización (Master), EO<sub>W</sub>: Etapa de Optimización (Worker), esc: escenario, LFR<sub>t</sub>: línea de fuego Real del instante t, EE: Etapa Estadística, BK<sub>ign</sub>: Búsqueda del valor clave de ignición, FF: Función de Fitness, K<sub>ign\_t</sub>: Valor clave de ignición del instante t, EC<sub>MON</sub>: Etapa de Calibración (Monitor), EP: Etapa de Predicción, LFP: Línea de fuego predicha.

### 3.3. Monitorización y análisis de métricas poblacionales.

El gráfico general de ESSIM-DE(r) con las capacidades de sintonización dinámica se presenta en la Figura 7. Utilizaremos la sigla ESSIM-DE(r)-IQR-ME para nombrar a ESSIM-DE(r) con las capacidades de sintonización dinámica. La monitorización de las métricas es llevada a cabo en el Master, en una nueva sub-etapa de la etapa Etapa de Optimización del Master (EO<sub>M</sub>), denominada **Etapa de Monitorización (EM)**, y se concreta en dos fases:

- Métrica ME (Movimientos Efectivos): El proceso Master actualizará la métrica ME cada vez que los procesos worker envíen la evaluación de la aptitud de un individuo. En la etapa EO<sub>M</sub>, el Master analiza si el nuevo individuo generado es mejor que su ancestro. El resultado de dicho análisis es introducido a la etapa EM en la cual se incrementa un contador local para la métrica ME.
- Métrica DP (Distribución Poblacional): Para realizar el cómputo de IQR, se requieren registrar los valores de aptitud de todos los individuos de la población. Por lo tanto, una vez que finaliza una generación evolutiva, se envían dichos valores a la etapa EM.

Como el cómputo de ambas métricas (ME y DP) se realiza en cada isla, la **Etapa de Análisis** requiere de la asistencia de los *Masters* de cada isla y del *Monitor*. Los valores registrados son enviados desde EM hacia la Etapa de Análisis llevada a cabo en el *Master* ( $EA_M$ ). Para que se calcule la métrica  $IQR$ , el *Master* ordena los valores de aptitud en orden creciente y computa la expresión (3). Al finalizar cada iteración, cada *Master* envía las métricas de su isla computadas en la  $EA_M$  hacia el proceso *Monitor*, quien realiza una agregación de todas las métricas de las islas en su Etapa de Análisis ( $EA_{MON}$ ). Los valores recopilados en la  $EA_{MON}$  son enviados a la Etapa de Sintonización ( $ES$ ), en la cual se determina si existe alguna isla con tendencia a estancamiento y/o convergencia prematura. Si se evalúa como verdadera la expresión (5) con los valores computados de alguna de las islas, el *Monitor* determina que la población, de acuerdo a sus características de diversidad y calidad, es suficientemente apta como para frenar la evolución, realizar la predicción, e iniciar con el siguiente paso. Luego, se aplicará el operador de reinicio poblacional (OR) para comenzar el proceso evolutivo del siguiente paso con una nueva población (nuevo espacio de búsqueda).

#### 4. Experimentación y Resultados Obtenidos

Para poder validar la propuesta de sintonización automática, con las nuevas métricas que miden la distribución poblacional y los movimientos efectivos del optimizador, se realizó un estudio experimental comparando los valores obtenidos con ESSIM-DE(r), respecto a ESSIM-DE(r) incorporando las capacidades de sintonización, al cual denominaremos ESSIM-DE(r)-IQR-ME. Los valores de referencia que se establecieron en las primeras pruebas fueron  $umbral\_IQR=2.5\%$  y  $umbral\_ME=20\%$ . Esto significa que, cuando la diferencia entre el tercer y el primer cuartil esté por debajo de 0.025 y cuando haya menos del 20% de movimientos efectivos del optimizador, se considera que la población tiene una tendencia al estancamiento y/o a la convergencia, y por lo tanto el ciclo evolutivo finaliza. La configuración para el resto de los parámetros ha sido establecida en base a aquellos valores utilizados en computaciones previas con el método ESSIM-DE(r), y se incluyen en la Tabla 1.

Para realizar los experimentos se han utilizado tres casos reales de estudio correspondientes a incendios controlados realizados en diferentes terrenos ubicados en Portugal (precisamente en Serra de Lousã, Gestosa), todos pertenecientes al proyecto SPREAD [Viegas 2004]. Cada caso posee una determinada duración expresada en minutos, una pendiente en grados y una superficie específica en metros cuadrados. El período de duración total de cada incendio ha sido dividido en lapsos de tiempo denominados pasos de simulación. Los tres casos de quema considerados se describen en la Tabla 2. Para cada uno de ellos se computaron los valores de *fitness* y de tiempos, para un promedio de 10 semillas. Los experimentos fueron realizados en un *cluster* de computadoras, utilizando un total de 41 unidades de procesamiento Intel Q9550 2.83GHz, con memoria RAM de 4GB DDR3 1333Mz.

**Tabla 1. Configuración de los parámetros del método**

Parámetro	Valor
Tamaño de cada población	200 individuos
Porcentaje de individuos a migrar	20%
Frecuencia de migración	en cada iteración
Límite de iteraciones evolutivas	30 iteraciones
Estrategia de mutación y cruzamiento	DE/rand/1/bin
Probabilidad de cruzamiento	0.45
Factor de mutación	0.9
Cantidad de islas	5
Cantidad de <i>workers</i> por isla	7

**Tabla 2. Casos de quema: dimensión, pendiente y tiempo de inicio, fin e incremento.**

Caso	Ancho (m)	Alto (m)	Pend. (gr)	T.ini (min)	T. fin (min)	Incr.(min)
A	89	109	21	2.0	14.0	2.0
B	75	126	19	3.0	9.0	1.0
C	60	90	6	2.0	10.0	2.0

**Tabla 3. Tiempos de ejecución en promedio.**

Caso	ESSIM-DE(r)	ESSIM-DE(r)-IQR-ME	reducción
A	[0:48:14]	[0:27:22]	43.2%
B	[0:25:53]	[0:22:54]	11.5%
C	[0:45:8]	[0:14:46]	67.3%

#### 4.1. Resultados

Los resultados obtenidos son presentados en las Figuras 8, 9 y 10 representando la experimentación con los casos A, B y C de la Tabla 2, respectivamente. En cada una se incluye un gráfico con los valores de *fitness* promedio obtenidos por cada experimento y en cada paso de predicción.

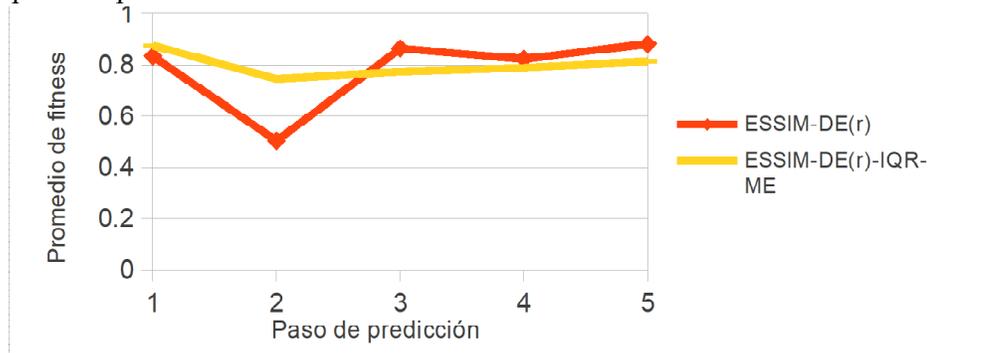


Figura 8. Experimento A.

En la Figura 8 se muestran los resultados obtenidos con el experimento A. El incendio real se inicia en el minuto 2 y finaliza en el minuto 14, con un incremento de 2 minutos. Por lo tanto cuenta con un paso de simulación (minuto 2 al minuto 4) y 5 pasos de predicción (minutos 4-6, 6-8, 8-10, 10-12 y 12-14). Como es posible observar de la Figura 8, ESSIM-DE(r)-IQR-ME supera a ESSIM-DE(r) en los dos primeros pasos de predicción, obteniendo valores promedio de *fitness* cercanos a 0.9 en el paso 1 y a 0.75 en el paso 2 (recordemos que una predicción perfecta es equivalente a obtener un valor de *fitness* 1.0). En los pasos de predicción 3, 4 y 5 ESSIM-DE(r) obtiene mejores resultados respecto del método con sintonización. Sin embargo, los valores obtenidos por ESSIM-DE(r)-IQR-ME son cercanos a 0.8 en la mayoría de los pasos de predicción, mientras que ESSIM-DE(r) evidencia un bajo rendimiento respecto al *fitness* en el paso 2. En el contexto de la predicción de incendios, resulta fundamental analizar el tiempo de ejecución en promedio de ambos métodos. Para ello, se incluye la Tabla 3, con los tiempos de ejecución en promedio para cada caso. Se puede observar para el Experimento A que ESSIM-DE(r)-IQR-ME obtiene una reducción de un 43% aproximadamente. Esta reducción del tiempo está asociada a una detección temprana de la tendencia de la población al estancamiento y/o convergencia prematura. Esto significa que gracias al proceso de sintonización, el ciclo evolutivo es interrumpido en menor cantidad de iteraciones respecto de ESSIM-DE(r), lo cual permite reducir el tiempo de ejecución global.

Los resultados obtenidos para el Experimento B se describen en la Figura 9. En este caso, el incendio consta de cinco pasos de predicción. Se puede observar que ESSIM-DE(r)-IQR-ME mejora los promedios de *fitness* en el paso 2, y mantiene promedios similares en los pasos 1, 3 y 5. En el paso de predicción 6, ESSIM-DE(r) obtiene mejor desempeño. Respecto a los tiempos de ejecución en promedio, se puede observar de la Tabla 3 que ESSIM-DE(r)-IQR-ME obtiene una reducción de un 11.5% aproximadamente. En este caso, los resultados de *fitness* y de tiempos de ejecución son similares para ESSIM-DE(r) y ESSIM-DE(r)-IQR-ME.

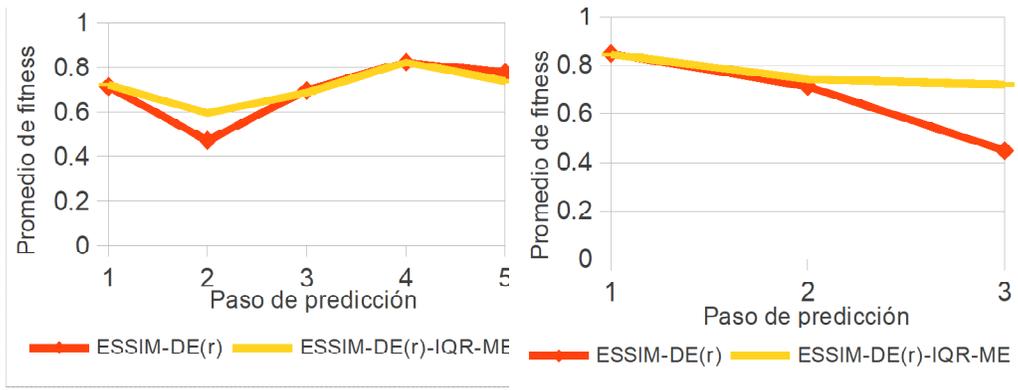


Figura 9. Experimento B

Figura 10. Experimento C

La Figura 10 muestra los resultados obtenidos para el Experimento C. En el paso de predicción 1, ESSIM-DE(r) y ESSIM-DE(r)-IQR-ME obtienen promedios de *fitness* similares, pero a medida que avanza la simulación ESSIM-DE(r)-IQR-ME mejora los resultados obtenidos respecto de ESSIM-DE(r). Se puede notar una marcada diferencia en el paso de predicción 3, en el cual ESSIM-DE(r)-IQR-ME alcanza un valor promedio de *fitness* superior a 0.7. También es posible remarcar una importante reducción del tiempo de ejecución de la versión con sintonización dinámica ESSIM-DE(r)-IQR-ME, logrando una reducción de un 67% menos respecto de ESSIM-DE(r), aproximadamente. En este experimento, el método con el proceso de sintonización ha demostrado mejor rendimiento, tanto en calidad de predicción como en tiempos de ejecución promedio.

## 5. Conclusiones

En este trabajo se presentó una estrategia de sintonización dinámica definida para ESSIM-DE(r), un método general de reducción de incertidumbre aplicado a la predicción de incendios forestales. Las acciones de sintonización permiten determinar dinámicamente el límite de iteraciones evolutivas con el fin de detectar anticipadamente una tendencia al estancamiento y a la convergencia prematura en la Etapa de Optimización de ESSIM-DE(r). Las métricas del modelo de rendimiento definidas fueron la medida estadística de dispersión denominada Rango Intercuartil (IQR) y los movimientos efectivos del optimizador (ME). El esquema paralelo de múltiples islas de ESSIM-DE(r) permite que estas métricas sean computadas en cada población. El proceso Monitor de ESSIM-DE(r) es quien determina en la Etapa de Sintonización si alguna isla posee una tendencia al estancamiento y/o a la convergencia prematura y determina el fin de la Etapa de Optimización. La experimentación fue realizada con tres casos de quemas reales controladas. Los resultados obtenidos indican que la estrategia definida logra mejorar la calidad de las predicciones, para algunos pasos de predicción, y a un costo temporal menor comparado con el método original. La reducción del tiempo de ejecución ronda entre un 12% a un 67% para los casos de quema considerados. Esta reducción de tiempos está asociada a la detección temprana a los problemas del estancamiento y convergencia, evitando que el optimizador continúe con el ciclo evolutivo, y por lo tanto, reduciendo el tiempo de ejecución global. Como trabajo a futuro se plantea, en primer

lugar, realizar una calibración estática para identificar si los valores utilizados para los umbrales de las métricas definidas son adecuados, o pueden ser mejorados. En segundo lugar, se plantea establecer un criterio para auto-ajustar ambas métricas en tiempo de ejecución, de manera tal de evitar la configuración *off-line* de estos valores de umbral y simplificar la configuración de ESSIM-DE(r).

## References

- La Nacion (2018) <https://www.lanacion.com.ar/2099226-por-que-se-repiten-los-incendios-forestales-en-verano>. Accedido el 10 Julio de 2018.
- Caymes Scutari, P. , Bianchini, G. , Sikora, A. y Margalef, T. “Environment for Automatic Development and Tuning of Parallel Applications”. En *2016 Int. Conf. on High Perf. Computing & Simulation (HPCS)*, Innsbruck, 2016, pp. 743-750.
- Healey, J. F. (2007) “The Essentials of Statistics: A Tool for Social Research”, Thomson/Wadsworth.
- Herrera, F. y Lozano, M. (1996) “Adaptation of Genetic Algorithm Parameters Based on Fuzzy Logic Controllers”, *Gen. Alg. and Soft Comp.* Physica-Verlag, p. 95-125.
- Lampinen , J. y Zelinka, I. (2000) “On the Stagnation of the Differential Evolution algorithm”, VI Int. Conf. Soft Computing (Mendel 2000), Rep. Checa, p. 76-83.
- Naono, K., Teranishi, K., Cavazos, J. y Suda, R. (2010) “Software Automatic Tuning: From Concepts to State-of-the-Art Results”, Springer, New York.
- Real, R. y Vargas, J.M. (1996) “The Probabilistic Basis of Jaccard's Index of Similarity”, *Journal of Systematic Biology*, 45, 3, 1996, p. 380-385.
- Storn, R., Price, K. y Lampinen J. (2005), “Differential Evolution: A Practical Approach to Global Optimization”, Springer, New York.
- Tardivo, M.L., Caymes-Scutari,P., Méndez-Garabetti, M. y Bianchini, G. (2018) “Optimization for an Uncertainty Reduction Method Applied to Forest Fires Spread Prediction”, En *Computer Science – Cacic 2017*, Springer, p. 13-23.
- Ursem, R. (2002) “Diversity-Guided Evolutionary Algorithms”, En *Parallel Problem Solving from Nature - PPSN VII*, Springer Berlin Heidelberg, p. 462-471.
- Yang, M., Cai, Z., Li, C. y Guan, J. (2013) “An Improved Adaptive Differential Evolution Algorithm with Population Adaptation” En *Proc. 15th Annual Conf. on Genetic and Evolutionary Comput. (GECCO '13)*, New York, USA, p. 145-152.
- Yang, M., Li, C., Cai, Z., y Guan, J. (2015) “Differential Evolution With Auto-Enhanced Population Diversity” En *IEEE Trans. on Cybernetics*, 45, 2, p. 302-315.
- Viegas, D. X. (2004), “Project Spread – Forest Fire Spread Prevention and Mitigation”, <http://www.algosystems.gr/spread/>. Accedido el 10 de Julio de 2018.