

Miguel Méndez Garabetti

**Método de reducción de incertidumbre
basado en algoritmos evolutivos
y paralelismo orientado a la predicción
y prevención de desastres naturales**

**TESIS DOCTORAL EN CIENCIAS INFORMÁTICAS
PREMIO DR. RAÚL GALLARD | Año 2021**

**Método de Reducción de Incertidumbre
Basado en Algoritmos Evolutivos y
Paralelismo Orientado a la Predicción
y Prevención de Desastres Naturales**

Dr. Miguel Méndez-Garabetti

UNIVERSIDAD NACIONAL DE SAN LUIS

TESIS DOCTORAL EN CIENCIAS INFORMÁTICAS

**Método de Reducción de Incertidumbre
Basado en Algoritmos Evolutivos y
Paralelismo Orientado a la Predicción
y Prevención de Desastres Naturales**

Dr. Miguel Méndez-Garabetti

Director: Dr. Germán Bianchini

Co-Directora: Dra. Verónica Gil Costa

Co-Directora beca CONICET: Dra. Paola Caymes-Scutari

*La Plata, Argentina
Abril de 2021*


EDITORIAL DE LA UNLP

Méndez-Garabetti, Miguel

Método de reducción de incertidumbre basado en algoritmos evolutivos y paralelismo orientado a la predicción y prevención de desastres naturales / Miguel Méndez-Garabetti. - 1a ed. - La Plata: EDULP, 2022.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-8475-41-7

1. Tecnología Informática. I. Título.

CDD 004.02

Método de Reducción de Incertidumbre Basado en Algoritmos Evolutivos y Paralelismo Orientado a la Predicción y Prevención de Desastres Naturales

Dr. Miguel Méndez-Garabetti



EDITORIAL DE LA UNIVERSIDAD NACIONAL DE LA PLATA (EDULP)

48 N° 551-599 4° Piso/ La Plata B1900AMX / Buenos Aires, Argentina

+54 221 44-7150

edulp.editorial@gmail.com

www.editorial.unlp.edu.ar

Edulp integra la Red de Editoriales de las Universidades Nacionales (REUN)

ISBN 978-987-8475-41-7

Queda hecho el depósito que marca la Ley 11.723

© 2022 - Edulp

Impreso en Argentina

Agradecimientos

A Germán, Paola, Verónica y Laura.

Al CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas), por la beca doctoral otorgada, y todas las instituciones que me permitieron llevar a cabo esta tesis: UTN-FRM (Universidad Tecnológica Nacional, Facultad Regional Mendoza), LICPaD, Laboratorio de Investigación en Cómputo Paralelo/Distribuido, UNSL-FCFMyN (Universidad Nacional de San Luis, Facultad de Ciencias Físico-Matemáticas y Naturales), CELFI (Centro Latinoamericano de Formación Interdisciplinaria), entre otros.

Dedicado a

Luz, Aitana, Malena y mis padres.

Dr. Miguel Méndez-Garabetti
La Plata, 2021

Resumen

La presente tesis doctoral aborda la problemática de la incertidumbre existente en todo sistema de predicción, focalizando en el desarrollo de métodos de reducción de incertidumbre aplicados a la predicción de fenómenos naturales. Debido a que estos fenómenos suelen causar gran impacto en las comunidades, la flora y la fauna, el ecosistema, entre otros, los sistemas de predicción deben proporcionar respuesta en el menor tiempo posible. Por estos motivos, los métodos propuestos han sido desarrollados utilizando capacidades de alto rendimiento. El primer método desarrollado en esta tesis (ESS-IM), comenzó con el objetivo de lograr una mejora a una metodología previamente desarrollada denominada ESS (Sistema Estadístico Evolutivo). Específicamente se trabajó en el incremento del paralelismo de la metaheurística interna, incorporando una arquitectura basada en modelo de islas bajo un esquema de migración. Este desarrollo logró incrementar la capacidad de búsqueda de la metaheurística interna, impactando de forma directa en un incremento en la calidad de predicción del método. En la validación, ESS-IM fue aplicado en una serie de casos de quemas controladas e incendios forestales. Es importante destacar que, en forma conjunta, al desarrollo de la tesis, se llevaron a cabo diferentes investigaciones complementarias, tales como: estudios de sintonización de parámetros, desarrollo de un sistema de generación de mapas de incendios forestales a partir de imágenes satelitales, entre otros. Finalmente, en la última etapa de la tesis, se implementó una versión híbrida basada en metaheurísticas evolutivas bajo una estrategia colaborativa basada en islas. El método HESS-IM, se implementó de forma heterogénea (a nivel de hardware), logrando que los resultados obtenidos incrementen la calidad de predicción y eficiencia del método.

Contenido

Capítulo 1. Introducción	1
1.1. Planteamiento	1
1.2. Predicción.....	4
1.2.1. Sistemas de Predicción	5
1.3. Incertidumbre	6
1.3.1. Reducción de Incertidumbre.....	8
1.4. Desastres Naturales	9
1.4.1. Los Incendios Forestales como Desastre Natural.....	11
1.5. Predicción de Incendios Forestales bajo Incertidumbre	14
1.5.1. Predicción Clásica	16
1.6. Computación de Alto Rendimiento	17
1.7. Objetivos de la Tesis	19
1.8. Organización de la Tesis	20
1.9. Contribución.....	21
Capítulo 2. Metaheurísticas	23
2.1. Introducción.....	23
2.2. Optimización	24
2.2.1. Optimización Multiobjetivo	25
2.2.2. Optimización Combinatoria	26
2.2.3. Optimización Multimodal	26
2.2.4. Optimización Dinámica.....	27
2.3. Metaheurísticas.....	28
2.3.1. Clasificación de las Metaheurísticas	29
2.4. Metaheurísticas Paralelas	39
2.4.1. Paralelización de Metaheurísticas Poblacionales Evolutivas	40
2.5. Síntesis	45
Capítulo 3. Predicción de Incendios mediante Métodos Guiados por Datos	46

3.1. Introducción.....	46
3.1.1. Predicción.....	47
3.1.2. Simulación.....	48
3.2. Métodos de Predicción Guiados por Datos	48
3.2.1. Urgent Computing.....	50
3.2.2. Métodos Guiados por Datos	51
3.3. Síntesis	53
Capítulo 4. Métodos Propuestos	55
4.1. Introducción.....	55
4.2. Descripción de los Métodos	57
4.2.1. Componente Simulación	58
4.2.2. Componente Estadístico	64
4.2.3. Componente de Optimización Metaheurística	68
4.2.4. Componente de Computación de Alto Rendimiento	70
4.3. Integración.....	74
4.3.1. ESS.....	75
4.3.2. ESS-IM.....	78
4.3.3. HESS-IM.....	86
4.4. Síntesis	92
Capítulo 5. Evaluación Experimental y Análisis de Resultados	93
5.1. Introducción.....	93
5.1.1. Casos de quemas controladas	93
5.1.2. Casos de incendios forestales	97
5.2. Experimentos.....	101
5.2.1. Comparación A): S ² F ² M, ESS y ESS-IM.....	101
5.2.2. Comparación B): ESS y ESS-IM	108
5.2.3. Comparación C): ESS-IM y HESS-IM	113
5.3. Síntesis	125
Capítulo 6. Calibración de Métodos	126
6.1. Introducción.....	126
6.1.1. Tipos de calibración	126

6.2. Calibración de ESS-IM	127
6.2.1. Parámetros de los AEP	128
6.2.2. Calibración homogénea	130
6.2.3. Calibración heterogénea	137
6.3. Calibración de HESS-IM.....	143
6.3.1. Resultados	144
6.4. Síntesis	146
Capítulo 7. Imágenes Satelitales y Redes de Sensores Inalámbricos....	147
7.1. Introducción.....	147
7.1.1. Alerta Temprana y Monitoreo	148
7.1.2. Predicción de Incendios Forestales mediante Información Satelital.....	150
7.2. Detección de área quemada	151
7.2.1. Tipos de imágenes necesarias para estimar a partir del uso de bandas e índices espectrales	153
7.2.2. Obtención de imágenes	154
7.2.3. Proceso de Automatización	155
7.3. Generación de casos de prueba.....	159
7.3.1. Caso 1	160
7.3.2. Caso 2	165
7.4. Propuesta: detección de incendios con red de sensores inalámbricos	168
7.5. Síntesis	171
Capítulo 8. Conclusiones y Líneas de Trabajo Futuro	173
8.1. Introducción.....	173
8.2. Cumplimiento de los objetivos planteados.....	178
8.3. Líneas de trabajo futuro.....	180
Bibliografía	182

Introducción

Este capítulo introduce el problema abordado en la presente tesis, los objetivos planteados en el plan de trabajo doctoral, como también aquellos que surgieron con el avance de la labor. Finalmente, se presenta la organización de la tesis y las contribuciones realizadas producto de esta investigación.

1.1. Planteamiento

Los modelos se han convertido en uno de los principales instrumentos de la ciencia moderna, éstos suelen utilizarse en la representación de gran cantidad de sistemas físicos [1], inclusive en el ámbito de la ciencia ambiental para efectuar predicciones. La modelización puede ser vista como el establecimiento de relaciones semánticas entre la teoría y los fenómenos u objetos. Generalmente, los modelos representan alguna situación real de manera incompleta e inexacta; esto permite realizar estudios de sistemas complejos de forma aproximada. Usualmente, los modelos deben ser alimentados con parámetros de entrada, los cuales brindan la información necesaria con la cual es posible determinar las condiciones iniciales del fenómeno bajo estudio, y de esta manera estimar sus posibles variaciones en el tiempo. Esto se conoce como la salida del modelo, la cual representa la evolución del sistema a lo largo del tiempo.

Cuando se lleva a cabo el estudio de un fenómeno complejo mediante la simplificación de la realidad, usando modelos, éstos pueden verse afectados debido a la calidad de la información con que dicho modelo es alimentado. Es decir, el mero hecho de simplificar un fenómeno físico para poder representarlo mediante un modelo implica pérdida de precisión, exactitud y certeza, respecto del comportamiento real del fenómeno. Si junto a esto se considera que dicho modelo “simplificado” es alimentado con información incompleta o de baja calidad, evidentemente tendremos resultados poco satisfactorios y distantes de la realidad.

Esta problemática, que surge a partir de la falta de precisión o exactitud en los parámetros de entrada de un modelo, se conoce como incertidumbre. Como bien se menciona en [2]: “cuando se analizan sistemas complejos, los resultados calculados incluyen incertidumbres debidas a los propios modelos informáticos, como así también a las existentes en los valores de las variables de entrada o los parámetros utilizados en el análisis. Debido a esto, es importante no solo cuantificar las incertidumbres de salida sino también identificar las variables de entrada dominantes que contribuyen a las incertidumbres de salida”. En base a esto, se puede decir que en cualquier ámbito donde se realice algún tipo de predicción existirá mayor o menor grado de incertidumbre (tanto se trate de un modelo económico, meteorológico o natural) y con ello el esfuerzo por intentar minimizar sus efectos negativos. Dicho en otras palabras, ante la presencia de incertidumbre (principal fuente de imprecisión en todo sistema de predicción), es necesario intentar eliminarla, o en el mejor de los casos lograr reducirla.

En la literatura se utiliza el término Reducción de Incertidumbre o Métodos de Reducción de Incertidumbre para hacer referencia a aquellas técnicas o herramientas que de alguna manera permiten reducir los efectos causados por la falta de precisión en cualquier variable o factor de un sistema de predicción. El concepto de Método de Reducción de Incertidumbre (MRI o Uncertainty Reduction Method, URM) ha sido acuñado en el trabajo de Ishigami y Homma en 1990 [2].

El concepto de MRI se utiliza en diversas áreas de la ciencia, por ejemplo, en [3] se propone un nuevo MRI para evaluar la precisión de predicción de las características de la neutrónica para el proceso de fabricación de combustibles. En [4] se presenta un MRI con integración de calibración Bayesiana y validación de un modelo de forma adaptativa aplicado a modelos de estructura y material. Dicho método se utiliza para predecir la respuesta estructural en diferentes condiciones de carga. También en [5] se propone un MRI epistémico basado en la propagación de incertidumbre y en la optimización anidada en dos etapas para la reducción de la incertidumbre. En [6] se presenta un enfoque de reducción de incertidumbre basado en análisis de sensibilidad de un sistema de ingeniería multidisciplinario descompuesto en múltiples subsistemas. En [7] se presenta un sistema basado en metaheurísticas para reducir la incertidumbre de segmentación en imágenes de resonancia magnética cerebral. En [8] se presenta el desarrollo de un MRI para la evaluación del impacto del cambio climático en la disponibilidad de recursos hídricos (la disponibilidad de agua en una región es uno de los factores más importantes para determinar la calidad de la vida humana, la cual puede

modificarse drásticamente debido al impacto del cambio climático) en [9] se presenta un MRI para hacer frente a la incertidumbre de la demanda de unidades en los sistemas de planificación y control de fabricación. Por su parte, en [10] se presenta la adaptación de una estrategia de gestión de la incertidumbre computacional en el marco de un sistema de predicción y prevención de inundaciones costeras. Esta estrategia consiste en determinar formas significativas de propagación de la incertidumbre y la ubicación de bloques de acción¹ para reducir o evaluar la incertidumbre, por citar sólo algunos ejemplos.

En el contexto de la presente tesis se define un Método de Reducción de Incertidumbre como un sistema computacional que tiene como propósito mejorar la calidad de salida de un modelo determinado, de forma independiente a los datos con que dicho modelo haya sido alimentado. Esto se logra mediante algún tipo de pre-procesamiento y/o pos-procesamiento de los datos de entrada y/o salida del modelo. Las funciones llevadas a cabo por un MRI no deben interferir con las operaciones internas del modelo bajo análisis, es decir, este último debe ser tratado como una caja negra.

Una representación gráfica de un MRI puede observarse en la Fig. 1.1. En ella se puede apreciar que las funcionalidades del método conforman un contexto que interactúa con las entradas y salidas del modelo, sin operar de forma directa con los componentes internos. Los MRI cobran vital importancia en el desarrollo de sistemas de predicción de fenómenos naturales, ya que éstos se ven fuertemente afectados por la presencia de incertidumbres [11].

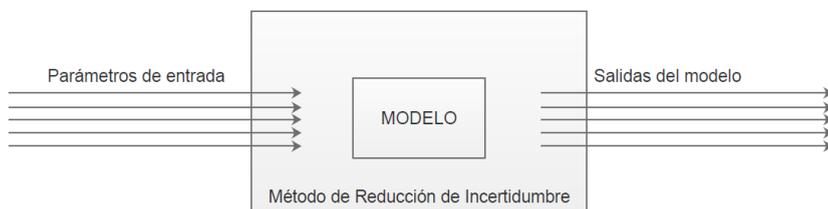


Figura 1.1: Relación Modelo y Método de Reducción de Incertidumbre.

Tales sistemas de predicción pueden ser utilizados para diferentes propósitos, tales como:

¹ Los bloques de acción involucran varios procedimientos: calibración de modelos, asimilación de datos, pronósticos, y diversas técnicas para la evaluación de la incertidumbre.

- anticipar ocurrencia de desastres naturales,
- estimar riesgos de desastres naturales,
- pronosticar posibles comportamientos de los fenómenos naturales,
- planificar acciones de evacuación ante desastres naturales, etc.

La predicción de cualquier fenómeno natural es una tarea compleja tanto por la cantidad de parámetros de entrada que los modelos suelen utilizar, como también por la complejidad inherente al fenómeno que se intenta modelar, y más aún debido a la imposibilidad de poder cuantificar la totalidad de las variables en tiempo real. Debido a estos problemas es necesario el desarrollo y perfeccionamiento de métodos de reducción de incertidumbre que permitan mejorar la calidad de las predicciones generadas por los modelos diseñados.

Diferentes técnicas y herramientas han sido utilizadas en el desarrollo y mejora de métodos de reducción de incertidumbre, entre ellas se puede destacar: herramientas estadísticas, técnicas de optimización como las metaheurísticas/heurísticas, plataformas de computación paralela o de alto rendimiento, técnicas de inyección de datos en tiempo real, inteligencia artificial, entre otras.

Este trabajo aborda el desarrollo de lo que aquí se llama “Método General de Reducción de Incertidumbre” (MGRI) para ser aplicado a sistemas de predicción y prevención de desastres naturales en general. Los métodos desarrollados centran sus bases en la combinación de diferentes herramientas tales como: análisis estadístico, técnicas y/o estrategias de optimización aproximadas, puntualmente metaheurísticas híbridas y diferentes tecnologías de paralelismo computacional.

1.2. Predicción

Según [12], la Ciencia Aplicada se orienta a conocer y comprender mediante inferencias y predicciones. Actualmente la predicción es utilizada en diversas áreas de la ciencia y la industria para tomar decisiones y desarrollar estrategias. La predicción es de por sí una tarea compleja y, por ende, ya sea que se trate de fenómenos naturales o eventos económicos, tales sistemas o métodos que realicen de alguna manera predicción deben ser diseñados especialmente para la aplicación que se pretende desarrollar de forma conjunta con usuarios y analistas [13]. En este sentido, la predicción debe permitir además ser

utilizada para elaborar diseños que contribuyan a la solución de problemas y evaluar la validez del conocimiento disponible.

Según la Real Academia Española [14], predecir consiste en anunciar por revelación, conocimiento fundado, intuición o conjetura algo que ha de suceder. Evidentemente, desde el contexto de la ciencia, el acto de predecir se relaciona con el hecho de contar con el debido “conocimiento fundado”, que permita realizar tales predicciones. La palabra predecir se utiliza en Ciencias Económicas, Ciencias Sociales, Ciencias Ambientales, Ciencias de la Computación, entre otras. Inclusive existe cierto debate por la diferencia entre “predecir” y “pronosticar”, que, si bien tienen diferencias sutiles, según el contexto de aplicación, en la mayoría de los trabajos de la literatura científica ambos términos se utilizan a menudo como sinónimos.

Debido a que la predicción es un concepto central en la ciencia, y dado que su significado difiere de forma sutil en diferentes áreas y disciplinas, resulta importante definir el alcance del concepto de predicción y de sistemas de predicción utilizado en el contexto de esta tesis.

1.2.1. Sistemas de Predicción

En este trabajo se habla de predicción y se enmarca dentro del concepto de “Sistema de Predicción”. En general, un sistema de predicción es un software que, a partir del conocimiento de ciertas variables iniciales, y mediante la ejecución de diferentes operaciones, permite realizar estimaciones futuras acerca del comportamiento de un fenómeno determinado [15].

El término sistema de predicción se utiliza con gran frecuencia en diversas áreas de la ciencia. Por ejemplo, en [16] se describe un sistema para problemas relacionados con la predicción de las características de las olas oceánicas. En [17] se propone un sistema de predicción de la tendencia del precio de las acciones, el cual puede predecir tanto el movimiento del precio de éstas como su intervalo de crecimiento. También en [18] se presenta el concepto de sistema de predicción en el contexto de la educación, puntualmente se trata de un sistema de alerta temprana para identificar a estudiantes en riesgo, con el cual pueden predecir la calificación final de los estudiantes. Otro estudio [19], propone un sistema de predicción de posición utilizando datos de ubicación adquiridos por dispositivos de posicionamiento integrados en teléfonos móviles y automóviles. En [20] se propone un sistema de predicción de retardo de trenes basado en datos de redes ferroviarias a

gran escala haciendo uso de big data. En [21], se comenta sobre un sistema de predicción, en el ámbito de la salud, que, al considerar la gravedad de lesiones de trauma severo, puede predecir si es probable que los pacientes sobrevivan o no. En [22] se propone un sistema de predicción y diagnóstico de cáncer de mama considerando que la alta precisión en la predicción del cáncer es importante para mejorar la calidad del tratamiento y la tasa de supervivencia de los pacientes.

Según el fenómeno con el que se esté trabajando, el modelo utilizado, la cantidad, frecuencia y calidad de los datos de entrada con que se cuente, se estará ante la presencia de mayor o menor grado de incertidumbre. Es decir, siempre que se hable de predecir o pronosticar, se estará ante situaciones de incertidumbre. Esta incertidumbre puede manifestarse de diferentes maneras, tales como:

- falta de precisión en la información de entrada,
- grado de detalle del modelo respecto de la realidad,
- frecuencia de acceso a información actualizada sobre el fenómeno real,
- tipo de implementación computacional que se esté utilizando, entre otras.

Tal como se puede observar, el concepto de Sistema de Predicción abarca innumerables áreas del conocimiento y siempre tendrá estrecha relación con el grado de incertidumbre con el que se esté lidiando.

1.3. Incertidumbre

Con el advenimiento del modelado multiescalar y multifísica, encontrar formas prácticas pero rigurosas de interpretar la incertidumbre y caracterizar su impacto en la evaluación de resultados probables se ha vuelto cada vez más difícil. Desde el punto de vista del modelado computacional, se puede decir que existen dos tipos de incertidumbres. Por un lado, las incertidumbres aleatorias, se refieren a fenómenos físicos que son aleatorios por naturaleza, y por el otro, las incertidumbres epistémicas a los parámetros de un modelo computacional para el cual hay una falta de conocimiento. Aquí también entran los errores de modelado, que no se pueden describir en términos de los parámetros del modelo computacional [23]. En este trabajo se adopta un concepto de “Incertidumbre” que incluye ambos tipos.

Como se ha mencionado, la incertidumbre está presente en todo sistema de predicción, por lo tanto, es importante:

- identificar las principales fuentes de incertidumbre.
- cuantificar de algún modo la cantidad de incertidumbre.
- tomar acciones que traten de reducir de alguna manera esta incertidumbre.

Además, en [24] Schoemaker plantea una estrecha relación entre incertidumbre y complejidad, la cual es representada gráficamente en la Fig. 1.2, ubicando en el eje de abscisas la incertidumbre y en el eje de ordenadas la complejidad. “Esta última se refiere al número de variables y al grado en que están interrelacionadas. En cambio, la incertidumbre se refiere al grado de conocimiento disponible sobre la variable objetivo, ya sea simple o compleja” [25].

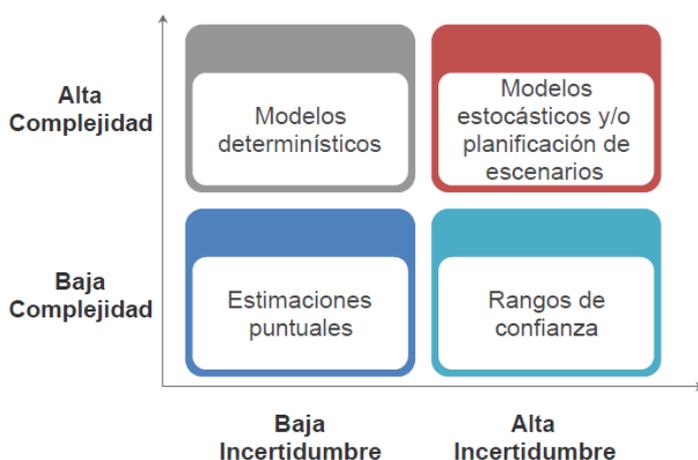


Figura 1.2: Complejidad e incertidumbre [24].

Incluso Schoemaker destaca que la complejidad no reconocida puede aparecer como incertidumbre, pero que son conceptualmente distintas. Evidentemente, la incertidumbre puede manifestarse de diversas formas haciendo compleja su identificación, cuantificación y reducción. Sin embargo, independientemente de cómo la incertidumbre se presente en un modelo determinado, el efecto de ésta se representa como una falta de precisión o error en la salida del modelo, por lo tanto, cualquier acción que mejore la calidad de salida estará reduciendo de alguna manera la incertidumbre del sistema.

1.3.1. Reducción de Incertidumbre

Tal como se ha mencionado, la reducción de incertidumbre corresponde al proceso que permite minimizar los efectos negativos causados por la falta de precisión en las variables de entrada de un modelo determinado, permitiendo optimizar el funcionamiento de éste e impactando de forma positiva en la calidad de salida del mismo. Este proceso debe llevarse a cabo sin realizar modificaciones en el funcionamiento interno de dicho modelo.

La reducción de incertidumbre puede lograrse de diferentes maneras en función del fenómeno con el que se esté tratando, el modelo utilizado, la implementación realizada, etc. Por ejemplo, en la predicción meteorológica, la reducción de incertidumbre puede llevarse a cabo mediante la medición de variables atmosféricas en tiempo real [26, 27].

El concepto ilustrado en la Fig. 1.1 representa esta idea, en donde el modelo (que representa al fenómeno) es tratado como una caja negra, y el MRI interactúa con las entradas y salidas del modelo realizando procesamiento de datos antes del ingreso y/o posterior al egreso del mismo. Un esquema más detallado de este concepto puede observarse en la Fig. 1.3 donde se consideran las interfaces (de entrada y salida) entre el modelo y el MRI, además se observa que el MRI es representado con una arquitectura modular compuesta de tres partes principales:

- MRI_E módulo entrada: responsable de la inicialización del sistema, lectura de parámetros y entrega de los conjuntos de datos requeridos por el modelo a la interfaz de entrada.
- MRI_S módulo salida: responsable de recibir los datos provistos por el modelo y adaptados por la interfaz de salida, analizarlos y determinar si los mismos deben ser evaluados nuevamente, o si están en condiciones de ser tratados como datos finales.
- MRI_C módulo central: este módulo se encarga de generar la salida final del sistema en base a los resultados entregados por el módulo de salida.

Las interfaces de entrada y salida son las responsables de efectuar las adaptaciones de formatos correspondientes entre los datos de entrada y los aceptados por el modelo, y de forma análoga en la salida.

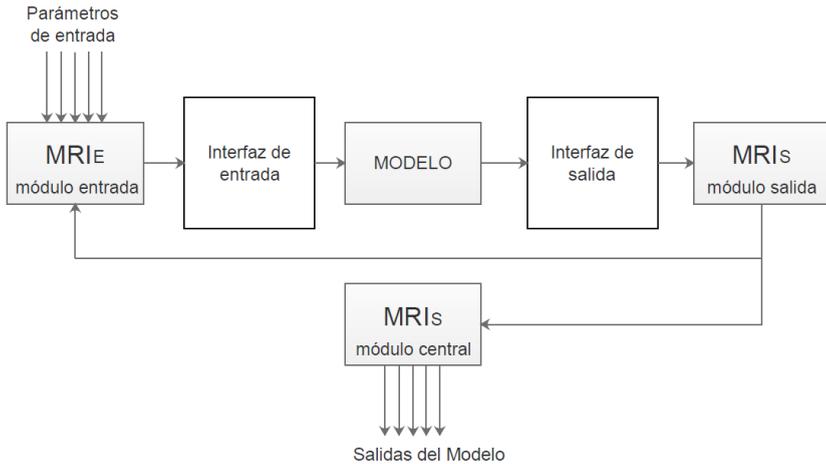


Figura 1.3: Relación modelo y método de reducción de incertidumbre.

1.4. Desastres Naturales

Un desastre natural es una situación negativa que altera la estabilidad y las condiciones de vida en un medioambiente dado. Los desastres naturales son causados por diferentes tipos de amenazas, las cuales pueden ser de origen natural o antrópico (este último a su vez puede ser intencional o no intencional) [28]. Los desastres naturales han causado la pérdida de millones de vidas en las últimas décadas, desplazando cada año a unos 26 millones de personas a la pobreza y ocasionando daños económicos del orden de los 520 millones de dólares anuales [29]. Una clasificación general, en función del tipo de origen de los desastres naturales, se lista a continuación [30, 31]:

- de origen natural.
- de origen antropogénico no intencional.
- de origen antropogénico intencional.

A su vez, los desastres de origen natural pueden ser generados a partir de fenómenos de diferente índole, tales como [31]:

- de origen atmosférico: ciclones tropicales, tormentas, tornados, relámpagos, granizadas, tormentas de viento, tormentas de hielo, tormentas de nieve, ventiscas, ondas frías, olas de calor, avalanchas, niebla, escarcha, entre otros.
- de origen geológico: temblores/terremotos, volcanes, tsunamis, derrumbes, hundimiento, entre otros.

- de origen hidrológico: inundaciones, deslizamientos, sequías, incendios forestales.
- de origen extraterrestre: asteroides, meteoritos.
- de origen biológico: enfermedades, epidemias, pandemias, superpoblación, hambruna.

Por su parte, los desastres naturales de origen antropogénico de carácter no intencional pueden ser originados por las siguientes actividades [31]:

- industria
- minería
- actividad nuclear
- transporte
- construcción, entre otros.

Finalmente, los desastres naturales de origen antropogénico intencional pueden originarse por cualquier tipo de acción humana intencional que ponga en riesgo o dañe el ambiente o las personas (generalmente se incluye en esta clasificación la desobediencia civil y las acciones bélicas). Todos estos tipos de eventos pueden ser de ocurrencia rápida o lenta, con consecuencias dramáticas en lo social, económico, ambiental y salud, afectando la calidad de vida de las poblaciones.

En consecuencia, un desastre natural puede ser definido como un impacto rápido e instantáneo del entorno natural en el sistema socio-económico [32], o como un desequilibrio repentino entre las fuerzas liberadas por el sistema natural y las fuerzas del sistema social. La gravedad de tales desequilibrios depende de la relación entre la magnitud del evento natural y la tolerancia de los asentamientos humanos a tal evento [33].

Por consiguiente, el riesgo de los desastres naturales no sólo depende de la posibilidad de que ocurran eventos o fenómenos naturales – cualquiera sea su origen–, sino también de las condiciones de vulnerabilidad que favorecen y/o facilitan que se desencadenen desastres cuando se presentan dichos fenómenos. Según la División de Medio Ambiente, Desarrollo Rural y Gestión de Riesgo de Desastres del BID [34] “la vulnerabilidad está íntimamente ligada a los procesos sociales que se desarrollan en las áreas propensas y usualmente tiene que ver con la fragilidad, la susceptibilidad o la falta de resiliencia de la población ante amenazas de diferente índole”. También indica que, “los desastres son eventos con componente ambiental, pero también social (socio-ambientales) cuya materialización es el resultado de la construcción social del riesgo”.

Dada esta estrecha relación y las dramáticas consecuencias que pueden verse generadas a partir de ello, se considera de gran importancia trabajar para lograr reducir las condiciones que benefician o facilitan la ocurrencia de los fenómenos naturales o aquellas que aumentan su impacto negativo. Esto debe tenerse en cuenta tanto para la reconstrucción posterior a la ocurrencia de un desastre, y también en la formulación de políticas públicas en pos de evitar tales daños.

En consecuencia, según lo expuesto en los párrafos anteriores, se deja evidencia de la necesidad de tomar acciones que permitan minimizar los efectos negativos causados por los desastres naturales. Tales acciones involucran a diferentes actores, y las medidas son de las más variadas: desde medidas económico-sociales, como creación de planes de ordenamiento territorial, desarrollo de sistemas integrados de gestión de riesgos, mayor inversión en estudios e investigaciones en sistemas de alerta y respuesta temprana, mejora de la conservación de los espacios naturales, incorporación y ampliación de la educación ambiental sobre riesgos ambientales, desarrollo de sistemas de predicción, entre otros.

1.4.1. Los Incendios Forestales como Desastre Natural

Tal como ya se ha comentado, la presente tesis se ha enfocado en el desarrollo de métodos generales de reducción de incertidumbre que puedan ser aplicados en la predicción de diferentes fenómenos naturales. Puntualmente se ha utilizado como caso de estudio a los incendios forestales dado que, desde el contexto local, éste es uno de los fenómenos naturales que ocurren con mayor frecuencia.

El fuego ha sido una herramienta fundamental en el desarrollo de las civilizaciones. Hough, en su trabajo [35], menciona que el hombre ha sido el único mamífero en superar el miedo al fuego, siendo esta cualidad una característica humana de gran importancia. Hough, en el mismo trabajo, deja abierto el interrogante sobre si la no adopción del fuego –por parte del ser humano– no hubiera dejado a la humanidad en decadencia destinada a la extinción. Posteriormente destaca que: “La única fuerza esencial disponible en la naturaleza unida al ingenio del hombre se convirtió en una combinación de potencialidades inimaginables”.

Cuando el fuego, por falta de control, se extiende sobre una superficie, se lo conoce como incendio. Cuando los incendios consumen la vegetación de áreas forestales, se denominan incendios forestales [36]. Cuando se habla de incendios forestales, comúnmente se connota con

los aspectos negativos que éstos poseen. Sin embargo, éstos juegan un rol fundamental en los cambios naturales que ocurren en los ecosistemas de nuestro planeta. Los efectos del fuego, sobre bosques, praderas, etc., propician la diversidad de la vida vegetal y animal. Incluso ciertos tipos de plantas no tienen posibilidad de reproducirse sin el fuego. El fuego inicia procesos naturales que permiten convertir la materia orgánica en nutrientes que posteriormente la lluvia entrega al suelo, proporcionando un semillero fértil rejuvenecido para las plantas [37]. Según Bond [38] “el fuego es un proceso generalizado en el sistema terrestre y juega un papel clave en la composición y distribución del ecosistema”.

Sin embargo, los incendios forestales fuera de control suelen generar grandes daños, destacándose entre ellos: pérdidas de vidas humanas, daños a la flora, la fauna y al suelo [39]. Las pérdidas económicas son difíciles de estimar debido a la gran cantidad de sectores que pueden verse afectados a corto, mediano o largo plazo, pero en general los recursos ambientales de toda zona afectada suelen ser considerablemente damnificados.

Los incendios forestales afectan millones de hectáreas año tras año haciendo de estos fenómenos uno de los principales agentes de perturbación de los ecosistemas del planeta. Además, ocurren con mayor frecuencia en época estival, donde se incrementa la sequedad, las temperaturas y/o el viento, lo que en conjunto facilita la propagación del fuego destruyendo así extensas superficies. Debido a las pérdidas y daños que pueden producir, desde los equipos de lucha contra incendios se trabaja constantemente en el desarrollo de herramientas, sistemas, estrategias, etc., que permitan reducir los efectos negativos que éstos ocasionan. Entre estos se pueden encontrar: estrategias de evacuación, tecnologías para la extinción del fuego, cortafuegos, simuladores de incendios, sistemas de predicción, etc. Estas herramientas suelen utilizarse en distintas etapas del proceso de lucha contra incendios siendo las mismas: prevención, predicción, detección y monitorización.

Según el último informe publicado por el Ministerio de Ambiente y Desarrollo Sustentable de la Presidencia de la Nación Argentina, año 2016, en su informe de Estadísticas de Incendios Forestales de la República Argentina [40], durante 2016 se consumieron 1.072.642 hectáreas. De este total un 54,74 % corresponde a superficie de pastizal, un 28,04 % de arbustal, un 16,47 % de bosque nativo y el resto corresponde a bosque cultivado y superficie sin determinar. A modo de ejemplo se pueden citar algunos incendios ocurridos a nivel local, regional y global durante los últimos años:

- Brasil 2019, incendio forestal ocurrido en la Amazonía Brasileira, durante el mes de agosto de 2019 consumiendo más de 2,5 millones de hectáreas. [41].
- Argentina 2018, incendio forestal ocurrido en los departamentos de General Alvear y San Rafael del sur de la provincia de Mendoza. Este incendio consumió más de 90.000 hectáreas. El mismo se propagó a una velocidad de 15 Km por hora debido a las ráfagas de viento de 60 a 70 km por hora [42].
- Argentina 2017, incendio forestal desatado en Cerro Mogote (provincia de Córdoba) el cual se extendió por zona de montañas y quebradas, afectando a más de 200 hectáreas durante agosto de 2017 [43].
- Argentina 2016 (Fig. 1.4), en agosto de 2016 en la provincia de San Luis, cerca de 10.000 hectáreas fueron consumidas por las llamas y más de 100 personas debieron ser evacuadas [44].
- Chile 2014, incendio ocurrido en Valparaíso (abril de 2014) donde el fuego consumió más de 2.900 viviendas, causando la pérdida de más de 15 vidas, alrededor de 500 heridos y más de 12.000 damnificados y evacuados (ver Fig. 1.5).
- Portugal 2017, entre el 17 y el 24 de junio de 2017, se produjo en Pedrógão Grande un gran incendio que afectó al centro de Portugal causando 64 muertos y más de 135 personas heridas (ver Fig. 1.6).



Figura 1.4: Incendio forestal ocurrido en San Luis, Argentina, en agosto de 2016.



Figura 1.5: Incendio forestal ocurrido en Valparaíso, Chile, en abril de 2014.



Figura 1.6: Incendio forestal ocurrido en Portugal en junio de 2017.

1.5. Predicción de Incendios Forestales bajo Incertidumbre

En la literatura, el término Predicción de Incendios Forestales puede referirse a dos conceptos con cierto grado de relación, pero a su vez muy diferentes. Uno de ellos consiste en poder predecir la ocurrencia de un incendio forestal en un área determinada bajo evaluación. El otro, en cambio, se refiere a la predicción del comportamiento de un incendio forestal ya iniciado. Ambos pueden ser utilizados para propósitos

cercanos, pero, es necesario diferenciarlos. En este trabajo, se hace referencia a la predicción del comportamiento, es decir, cuando se menciona la predicción de incendios, se estará hablando sobre el proceso que consiste en intentar pronosticar el comportamiento de la propagación de un incendio forestal ya iniciado, con la idea de conocer las zonas que presentan mayor riesgo de ser alcanzadas por el fuego y tomar decisiones o medidas de acuerdo a ello.

La simulación de la propagación de incendios forestales es una tarea compleja desde el punto de vista computacional, tanto por la complejidad involucrada en los modelos utilizados, como así también por la incertidumbre que afecta a este proceso. Más allá de la imprevisibilidad del comportamiento de los incendios forestales, la incertidumbre proviene de diferentes fuentes tales como: datos inexactos y/o faltantes, comprensión científica incompleta de la respuesta ecológica al fuego y de la respuesta del comportamiento del fuego a los tratamientos de mitigación, entre otros [11].

Según [45] un incendio forestal es una compleja combinación de energía liberada en un proceso de combustión en donde, dicha energía, al ser transportada hasta un combustible cercano (el cual aún no ha sido quemado), genera la posterior ignición de dicho combustible. Cada uno de estos procesos ocurren a diferentes escalas: a) la liberación de energía: ocurre en el dominio de la química y se produce a escala molecular y b) el transporte de energía: ocurre en el dominio de la física y se produce en escalas que van desde milímetros hasta kilómetros. La interacción de éstos procesos, en el amplio rango de escalas temporales y espaciales involucradas en incendios forestales, hace que el modelado de los incendios sea una tarea compleja [45].

Los modelos de comportamiento de fuego, utilizados para el desarrollo de sistemas de predicción, pueden clasificarse en modelos físicos, empíricos y semi-empíricos o semi-físicos. Los modelos físicos se basan en leyes que rigen la mecánica de fluidos y la transferencia de calor. Los empíricos se obtienen a partir de datos experimentales y no incluyen procesos físicos, por ello se aconseja utilizarlos cuidadosamente, ya que su uso se limita a las situaciones específicas en que fueron experimentados [46]. Los modelos semi-físicos y semi-empíricos se caracterizan por ser una mezcla de los dos primeros.

El modelo de Rothermel (1972) [47] es uno de los más utilizados; éste es un modelo semi-físico ya que combina técnicas físicas y correlaciones empíricas obtenidas a partir de datos experimentales. La mayoría de los simuladores de comportamiento del fuego emplean en sus cálculos este modelo.

1.5.1. Predicción Clásica

El concepto general de predicción del comportamiento de incendios forestales puede verse representado gráficamente en la Fig. 1.7, donde un Simulador de Incendios (SI), tratado como una caja negra, es alimentado con ciertos parámetros de entrada (PE) que representan condiciones actuales del incendio, tales como: velocidad y dirección del viento, vegetación, características de la superficie, entre otros. El simulador necesita, además de los PE, la Línea de Fuego Real (LFR) del incendio en el instante de tiempo t_n (LFR_n). Una vez que el simulador es ejecutado, éste proporcionará en su salida la predicción de la línea de fuego o Línea de Fuego Predicha (LFP) para el instante de tiempo posterior, t_{n+1} . Por supuesto, se espera que dicha predicción coincida con el estado real del incendio para dicho instante de tiempo (LFR_{n+1}). Este enfoque también es conocido como predicción de un paso adelante [48].

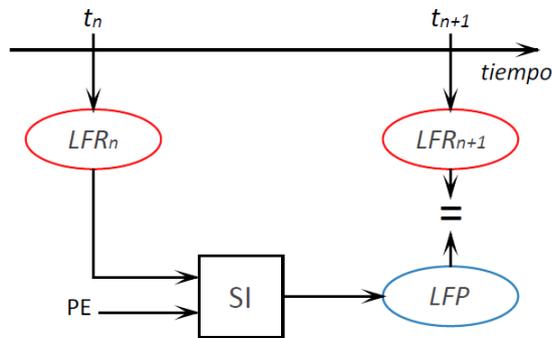


Figura 1.7: Predicción clásica. (LFR_x : Estado de la Línea de Fuego Real en tiempo x , SI: Simulador de propagación de Incendios forestales, LFP: Estado de la Línea de Fuego Propuesta por el modelo).

Sin embargo, debido a la complejidad e incertidumbre del modelo de comportamiento del fuego, la incertidumbre en los parámetros de entrada, y que la predicción está basada en una única simulación, esta metodología no suele ofrecer una predicción lo suficientemente aproximada a la realidad como para ser considerada una herramienta confiable para la toma de decisiones. Algunos ejemplos de predicción clásica son [49, 50, 51, 52, 53, 54].

Las limitaciones presentes en el enfoque de predicción clásica hacen evidente la necesidad de incorporar técnicas de reducción de incertidumbre que permitan mejorar el rendimiento (en calidad y tiempo de respuesta) de los sistemas de predicción actuales. En este

contexto, los desarrollos de los MRI cobran vital importancia ya que han demostrado, tras ser aplicados en diferentes áreas de la ciencia, efectividad a la hora de minimizar los efectos negativos causados por las diferentes incertidumbres permitiendo mejorar la calidad de salida de los modelos tratados. Además, debido a que estos sistemas suelen trabajar con grandes conjuntos de datos sobre los cuales es necesario realizar complejas operaciones, éstos sistemas suelen requerir de sistemas de computación de alto rendimiento, los cuales permiten reducir el tiempo de procesamiento al utilizar múltiples unidades de cómputo en simultáneo.

1.6. Computación de Alto Rendimiento

Resulta evidente que desde el origen de la computación ha habido un constante y arduo esfuerzo por conseguir mayores prestaciones computacionales. Este desarrollo ha generado grandes logros, materializados en avances significativos en el hardware y software de computadoras, haciendo cada vez más rápidos y potentes los sistemas de procesamiento. El mayor poder computacional ha permitido abordar problemas más grandes y complejos, aunque, para una gran cantidad de éstos, aún sigue siendo necesario contar con mayor capacidad de procesamiento. Este tipo de requerimientos ha desembocado en la utilización de grupos de procesadores de forma conjunta. Esta práctica se conoce como Computación de Alto Rendimiento (HPC, High Performance Computing) [55], concepto también asociado al de Computación Paralela [56]. En términos generales, se puede decir que la computación de alto rendimiento consiste en la utilización de determinada cantidad de elementos de procesamiento en la resolución de un problema de forma cooperativa, tratando de minimizar el tiempo de procesamiento y dar así una respuesta lo más rápido posible.

En HPC es importante distinguir entre el software paralelo y el hardware paralelo. El software se enfoca en las técnicas que permiten el tratamiento “concurrente” de los datos y acciones entre n procesos que trabajan en la solución de un problema común. Y el hardware se refiere a una arquitectura con múltiples unidades de procesamiento que tiene capacidad para soportar procesamiento paralelo [57].

Las aplicaciones HPC están diseñadas para resolver problemas complejos que requieren una gran cantidad de cálculos. Normalmente, los problemas requieren muchos más recursos que los disponibles en las estaciones de trabajo normales, o su ejecución es más lenta que la que necesita el usuario (por ejemplo, la predicción de fenómenos

naturales como el clima o la predicción del comportamiento de incendios forestales, en donde el resultado generado por el sistema de predicción posee una validez temporal). En este tipo de aplicaciones se vuelve necesario la utilización de sistemas de cómputo de alto rendimiento. En este contexto se puede identificar el concepto de Computación Urgente (UG, Urgent Computing) [58].

Los entornos HPC son sistemas de procesamiento paralelo, lo que significa que se debe contar con un hardware compuesto por varias unidades de procesamiento que permitan realizar dicho procesamiento paralelo. La forma en que se distribuyen los elementos de procesamiento y de memoria dan origen a diferentes arquitecturas paralelas. Una de las clasificaciones más relevantes, cuando de arquitecturas se habla, consiste en la taxonomía de Flynn [59, 60], la cual considera: a) el flujo de instrucciones y b) el flujo de datos:

- una instrucción y un dato (SISD, single instruction, single data): esta clasificación involucra a las computadoras compuestas por una única CPU, es decir, aquellas que pueden llevar a cabo un único flujo de instrucciones sobre un único flujo de datos, por lo tanto, computadoras que ejecutan las instrucciones de forma secuencial.
- una instrucción y múltiples datos (SIMD, single instruction, multiple data): esta clasificación incluye a las computadoras capaces de procesar un único flujo de instrucciones sobre múltiples flujos de datos (distintos datos). Evidentemente esta arquitectura cuenta con varias CPUs.
- múltiples instrucciones y un dato (MISD, multiple instruction, single data): en esta arquitectura muchas unidades de procesamiento realizan diferentes operaciones sobre los mismos datos. Prácticamente no existen computadoras que implementen este tipo de arquitecturas.
- múltiples instrucciones y múltiples datos (MIMD, multiple instruction, multiple data): aquí las computadoras poseen unidades de procesamiento independientes con sus propios flujos de instrucciones y de datos. En esta clasificación se encuentran dos tipos de arquitecturas: las de memoria compartida, y las de memoria distribuida. La mayoría de los sistemas de computación de alto rendimiento se encuentran dentro de la categoría MIMD, ya que inclusive a las SIMD se los considera como un caso especial de MIMD.

Como se ha comentado, de la arquitectura MIMD deriva otra taxonomía, la cual considera la memoria y la forma de comunicación entre los procesadores: a) de memoria compartida: donde todas las

unidades de procesamiento comparten la memoria de la arquitectura paralela, y b) de memoria distribuida: que se corresponde con las compuestas por varios procesadores y sus respectivas memorias conectadas mediante una red de interconexión, la cual se utiliza para efectuar el envío explícito de mensajes entre nodos. Es importante mencionar que también existen arquitecturas híbridas, las cuales permiten explotar el uso del hardware paralelo realizando paralelismo de memoria compartida dentro de cada nodo, y paralelismo de memoria distribuida entre la totalidad de los nodos.

Durante el desarrollo de la presente tesis se trabajó, en primer lugar, con una arquitectura de memoria distribuida basada en clusters de computadoras heterogéneas mediante la librería de paso de mensajes (MPI, Message Passing Interface) [61]. Posteriormente, con el propósito de incrementar el nivel de paralelismo de los métodos desarrollados, se realizó una implementación basada en memoria compartida distribuida, lo que se conoce como computación heterogénea. Dicha implementación consta del uso de paralelismo mediante unidades de procesamiento gráfico de propósito general (GPGPU, General Purpose Graphics Processing Units) mediante CUDA (Compute Unified Device Architecture), la arquitectura de cálculo paralelo de NVIDIA.

1.7. Objetivos de la Tesis

La presente tesis contribuye al desarrollo de métodos de reducción de incertidumbre capaces de asistir y participar en el proceso de resolución de problemas de alta magnitud y complejidad computacional, tal como la predicción y prevención de desastres naturales. Este tipo de problemas puede ser clasificado en la categoría de Grand Challenge Problems [62], es decir, problemas de alta complejidad, y a la vez importantes (desde el punto de vista social y económico) donde se requieren varios órdenes de magnitud de recursos para resolverlos.

Los métodos desarrollados, si bien han sido concebidos como generales (ya que podrían ser aplicados a diferentes fenómenos naturales con características de propagación), han sido aplicados a la predicción del comportamiento de incendios forestales usando metaheurísticas híbridas paralelas, análisis estadístico y computación de alto rendimiento. En el desarrollo de tales métodos también se involucran diferentes aspectos relacionados con: modelos y paradigmas de computación paralela, hibridación de diferentes tecnologías de paralelización (MPI, OpenMP y CUDA), simuladores de incendios forestales, entre otros.

De acuerdo a lo expuesto y con el propósito de alcanzar el objetivo general, se propone abordar en este trabajo los siguientes objetivos específicos:

- Diseño, implementación y evaluación de métodos de reducción de incertidumbre aplicados a la predicción y prevención de desastres naturales. Este objetivo implica un estudio profundo de la literatura existente en este campo y la definición de los casos de estudio de análisis.
- Inclusión de capacidad de operar con múltiples metaheurísticas. La hibridación de metaheurísticas supone un incremento en la capacidad de búsqueda de toda técnica de optimización. Debido a esto se propone dotar al método con la capacidad de operar de forma descentralizada con múltiples metaheurísticas de forma colaborativa.
- Identificación de los fenómenos naturales a trabajar, para así obtener datos de casos reales de acuerdo al modelo seleccionado. Adaptación de datos reales para su aplicación en los métodos desarrollados.
- Validación de los resultados obtenidos realizando una experimentación detallada aplicada a casos reales y contrastada con otras metodologías.

1.8. Organización de la Tesis

El presente trabajo de tesis se encuentra organizado en 8 capítulos. A continuación, se brinda una breve descripción de cada uno de ellos:

- Capítulo 1 - Introducción: Este capítulo ofrece una breve introducción a la problemática abordada en la presente tesis, su justificación, objetivos y finalmente se proporciona la organización de la tesis.
- Capítulo 2 - Metaheurísticas: En este capítulo se presentan los conceptos relacionados de las técnicas de optimización y metaheurísticas.
- Capítulo 3 - Predicción de Incendios mediante Métodos Guiados por Datos: Describe el concepto de método de predicción guiado por los datos, junto con un análisis del estado del arte.
- Capítulo 4 - Métodos Propuestos: En este capítulo se describe detallada mente la propuesta de los métodos ESS-IM y HESS-IM. Para ello, el mencionado capítulo comienza describiendo los métodos previos tales como: S²F²M y ESS.

- Capítulo 5 - Evaluación Experimental y Análisis de Resultados: Este capítulo describe las experimentaciones realizadas con el propósito de evaluar la validez de los métodos desarrollados.
- Capítulo 6 - Calibración de Métodos: Este capítulo presenta algunos resultados obtenidos luego de realizar un análisis de calibración de tipo estática los métodos propuestos. Se evidencia la necesidad de realizar este tipo de análisis de forma permanente.
- Capítulo 7 - Imágenes Satelitales y Redes de Sensores Inalámbricos: El método presentado en esta tesis trabaja sin inyección de datos en tiempo real. En este capítulo se aborda brevemente la posible interacción del método propuesto junto a tecnologías que permitan mejorar la calidad de los datos de entrada, tales como: Redes de Sensores Inalámbricos e Imágenes Satelitales.
- Capítulo 8 - Conclusiones y Trabajo Futuro: Presenta un resumen del trabajo y las conclusiones dando una mirada hacia los trabajos, retos y desarrollos futuros.

1.9. Contribución

En esta sección se presentan sintéticamente los principales aportes de esta tesis doctoral. Es importante mencionar que en la Sección 8 se listan, junto a las conclusiones, las principales publicaciones derivadas de estas contribuciones. En este trabajo se presentan avances en el desarrollo de métodos de reducción de incertidumbre en la categoría de Métodos Guiados por Datos con Solución Múltiple Solapada (DDM MOS, Data Driven Methods with Multiple Overlapping Solutions) [63], clasificación a la que pertenecen los métodos: S²F²M y ESS. Puntualmente en este trabajo se logra incrementar dos de las características principales de este tipo de metodologías, que son: a) la calidad de predicción, y b) el tiempo de respuesta. A continuación, brevemente se mencionan las principales contribuciones, y que posteriormente son desarrolladas en los capítulos siguientes:

- ESS-IM: Desarrollo del método de reducción de incertidumbre denominado Sistema Estadístico Evolutivo con Modelo de Islas (Evolutionary Statistical System with Island Model, ESS-IM). Este método ha sido aplicado a la predicción del comportamiento de incendios forestales, logrando incrementar considerablemente la calidad de predicción, respecto de

metodologías previas. Dicha mejora es lograda mediante el incremento del paralelismo de la metaheurística interna.

- HESS-IM: Desarrollo del método de reducción de incertidumbre denominado Sistema Estadístico Híbrido Evolutivo con Modelo de Islas (Hybrid Evolutionary Statistical System with Island Model, ESS-IM). Este desarrollo, también aplicado a la predicción del comportamiento de incendios forestales, ha permitido incrementar tanto calidad de predicción como tiempo de respuesta. Esto ha sido posible mediante la utilización de un esquema híbrido a nivel de la técnica de optimización basada en metaheurísticas, y uno heterogéneo a nivel de utilización de hardware. Para ello el método implementa un framework capaz de operar con diferentes metaheurísticas de forma colaborativa (con formando así una metaheurística híbrida) y el paralelismo necesario ha sido implementado basado en una arquitectura heterogénea CPU-GPU.

Metaheurísticas

En este capítulo presenta uno de los conceptos fundamentales relacionados a la contribución de esta tesis, puntualmente se trata de las estrategias o métodos de optimización utilizados: las metaheurísticas, y los enfoques relacionados al paralelismo.

2.1. Introducción

Cada día, ingenieros, científicos, tomadores de decisiones en general, se enfrentan a problemas de creciente complejidad en diversos sectores técnicos, por ejemplo, en la investigación operativa, el diseño de sistemas mecánicos, el procesamiento de imágenes, la electrónica, entre otros. El problema a resolver a menudo puede ser expresado como un problema de optimización, donde se define una función objetivo (o varias funciones), que se pretende minimizar o maximizar con respecto a todos los parámetros implicados. La definición del problema de optimización suele complementarse con información en forma de restricciones. Todos los parámetros de las soluciones adoptadas deben satisfacer estas restricciones, de lo contrario estas soluciones no son alcanzables.

En este capítulo se presentan los fundamentos necesarios de los métodos o algo ritmos de optimización conocidos como: metaheurísticas o meta-heurísticas. Estos, si bien se utilizan desde los años 40, donde no se estudiaban formalmente, han estado disponibles como tal a partir de los años 80 [64], y desde entonces han sido utilizados para la resolución de problemas de gran diversidad. Todo Algoritmo metaheurístico tiene un objetivo común: resolver, de la mejor forma posible, los problemas de optimización duros [65].

En la gran mayoría de estos problemas, no suele ser necesario o posible obtener una solución óptima, siendo aceptable en cambio una aproximación cercana a la mejor solución del problema [66]. La obtención de este tipo de soluciones, “aceptables” en relación a la complejidad del problema, puede ser generada mediante la utilización de técnicas de optimización aproximadas. En [66], los métodos de optimización se clasifican como exactos o aproximados según el grado

de complejidad asociada al problema tratado. Dicha clasificación se presenta en la Fig. 2.1. Los primeros se caracterizan por lograr la obtención de una solución exacta del problema tratado, con la desventaja de no ser aplicable a gran cantidad de problemas de la vida real (ya sea por su tamaño o complejidad). En cambio, los métodos aproximados permiten resolver problemas de optimización de alta complejidad con rasgos de incertidumbre y dinamicidad en plazos razonables de tiempo [67]. Esto ha favorecido a que cobren gran importancia en la comunidad científica en las últimas décadas.

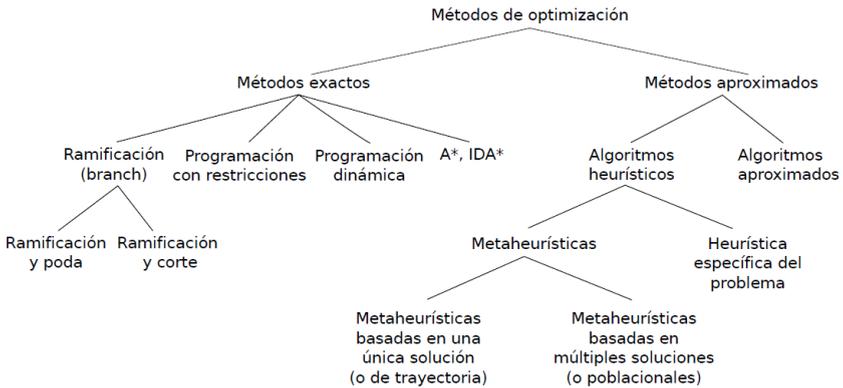


Figura 2.1: Clasificación Clásica de los Métodos de Optimización.

2.2. Optimización

Optimización es el proceso por el cual se intenta encontrar la mejor solución posible para resolver un determinado problema. Tal como se ha mencionado anteriormente, la optimización es necesaria para cualquier problema relacionado con la toma de decisiones, ya sea en ingeniería, economía, industrias, entre otros. La toma de decisiones implica elegir de la mejor manera posible entre un gran número de diferentes alternativas [65]. Existen diferentes criterios para clasificar los problemas de optimización. Desde el punto de vista del tipo de variables involucradas, éstos se pueden clasificar en: problemas discretos y problemas continuos. Por supuesto que en la práctica también se pueden encontrar problemas mixtos, que comprenden simultáneamente variables discretas y continuas.

En la literatura se mencionan dos tipos de problemas de optimización dura o NP-duros [68]. Por un lado, tenemos aquellos problemas de optimización discretos, para los cuales no se conoce un algoritmo

exacto con complejidad polinómica que permita resolverlo. Por otro lado, aquellos problemas de optimización con variables continuas, para las cuales no hay conocimiento de un algoritmo que permita localizar la mejor solución posible en un número finito de cálculos. En la historia se registran muchos esfuerzos llevados a cabo de forma aislada para resolver estos dos tipos de problemas encontrando diferentes limitaciones y dificultades en cada uno de los casos. Por ejemplo, en la optimización continua, la convexidad en la función objetivo, y en la optimización discreta, las heurísticas específicas diseñadas para la resolución de un problema dado.

Existe un gran abanico para clasificar a los problemas y técnicas de optimización. Por ejemplo, en el área de la Investigación Operativa (IO) existen cuatro tipos de problemas de optimización: optimización con restricciones, optimización multiobjetivo, optimización combinatoria y optimización multimodal [69]. Sin embargo, también es importante mencionar la optimización dinámica [70], que mantiene una estrecha relación con los problemas de optimización planteados en este trabajo. A continuación, se brinda una breve descripción de las definiciones más cercanas a este trabajo.

2.2.1. Optimización Multiobjetivo

Los problemas de optimización multiobjetivo son aquellos en los que es necesario realizar la optimización simultánea en más de una función objetivo, las cuales generalmente suelen estar en conflicto. Un problema de optimización multiobjetivo puede definirse como:

$$\begin{aligned}
 & \text{optimizar } F(x) = [(f_1(x), f_2(x), \dots, f_k(x))], \\
 & \text{sujeto a: } g_i(x) \leq 0, \quad i = \{1, \dots, m\} \quad (2.1) \\
 & \quad h_j(x) = 0, \quad j = \{1, \dots, p\} \\
 & \quad x = (x_1, \dots, x_n) \in \Omega
 \end{aligned}$$

Es importante tener en cuenta que las restricciones: a) $g_i(x) \leq 0$, de desigualdad y b) $h_j(x) = 0$, de igualdad, representan restricciones que deben cumplirse mientras se minimiza o maximiza $F(x)$. Además, Ω contiene todas las posibles soluciones x que pueden usarse para satisfacer una evaluación de $F(x)$. De esta manera, un problema multiobjetivo consta de k objetivos reflejados en las funciones objetivo $k, m + p$ restricciones sobre las funciones objetivo y n variables de

decisión. Las funciones objetivo k pueden ser lineales o no lineales y de naturaleza continua o discreta, al igual que el vector de las variables de decisión x_i [71].

2.2.2. Optimización Combinatoria

La mayoría de los problemas prácticos tienen un número finito o contable de soluciones alternativas que pueden ser formulados como problemas de optimización combinatoria [72]. En este sentido se puede tomar como definición de optimización combinatoria la siguiente:

Definición 1 *La optimización combinatoria es el estudio matemático de encontrar una disposición óptima, agrupando, ordenando o seleccionando objetos discretos normalmente finitos en números [73].*

Todo problema de optimización combinatoria puede ser definido por un par (S, f) donde $S \neq \emptyset$ representa el conjunto de soluciones factibles dentro del espacio de soluciones Ω ($S \subseteq \Omega$), y $f : \Omega \rightarrow \mathbb{R}$ es la función a optimizar o "función objetivo". La función objetivo, asigna a cada $s \in S$ un valor de costo real mediante el cual se puede establecer su calidad. Además, ya sea un problema de maximización o minimización, la función f permite definir una relación de orden entre las distintas soluciones. De esta manera, resolver un determinado problema de optimización combinatoria consiste en encontrar una solución óptima $s^* \in S$ que satisfaga la siguiente desigualdad:

$$f(s^*) \leq f(s), \quad \forall s \in S \quad (2. 2)$$

2.2.3. Optimización Multimodal

El término unimodal hace referencia a funciones que tienen un solo extremo, mínimo o máximo, en ellas. Por otro lado, multimodal se refiere a aquellas con dos o más extremos. Dicho lo anterior, se puede definir una función multimodal como aquella con más de un máximo o mínimo (u óptimo) local y/o global. Como óptimo local nos referimos a aquellas soluciones del espacio de búsqueda que son mejores a sus vecinos, pero que no necesariamente son la mejor solución posible, llamado óptimo global. La tarea de la optimización multimodal consiste en la obtención de soluciones de óptimas locales y globales en el espacio de búsqueda de la función a optimizar.

2.2.4. Optimización Dinámica

Seguida a esta clasificación, también debemos considerar los problemas de optimización dinámica (DOPs, Dynamic Optimization Problems) [70]. Los DOPs pueden ser definidos como:

$$DOP = \left\{ \begin{array}{l} \text{optimizar} \\ \text{s.t.} \end{array} \begin{array}{l} f(x,t) \\ x \in F(t) \subseteq S, t \in T \end{array} \right\} \quad (2.3)$$

donde:

- $S \in \mathfrak{R}^n$, S es el espacio de búsqueda.
- t es el tiempo.
- $f : S \times T \rightarrow R$ es la función objetivo, la cual asigna un valor numérico ($f(x, t) \in \mathfrak{R}$) a toda posible solución ($x \in S$) en el instante de tiempo t .
- $F(t)$, es el conjunto de posibles soluciones $x \in F(t) \subseteq S$ en el instante de tiempo t .

De otra manera, se puede decir que, en un problema de este tipo, el contexto cambia con el tiempo (la función objetivo o las restricciones). El método más sencillo para resolver estos problemas es ignorar la dinámica, considerando cada cambio como la llegada de un nuevo problema de optimización, pero a menudo es poco práctico. Por lo tanto, el objetivo de los métodos para tratar con DOPs ya no es localizar una solución óptima estacionaria, sino seguir su movimiento a través del espacio de solución y tiempo lo más cerca posible. Los algoritmos clásicos han sido adaptados para hacer frente a estos escenarios dinámicos mejorando su capacidad para el seguimiento del óptimo móvil.

Si bien en la resolución de cualquier tipo de problema siempre se perseguirá la obtención del óptimo global, para una gran cantidad de problemas y aplicaciones suele ser aceptable y válido la obtención de un óptimo local de buena calidad. Este concepto se encuentra representado gráficamente en la Fig. 2.2 (tal como se puede observar el valor alcanzado por el óptimo local es próximo al óptimo global). Si el problema acepta soluciones sub-óptimas, la primera solución podrá ser considerada como una solución aceptable al problema tratado.

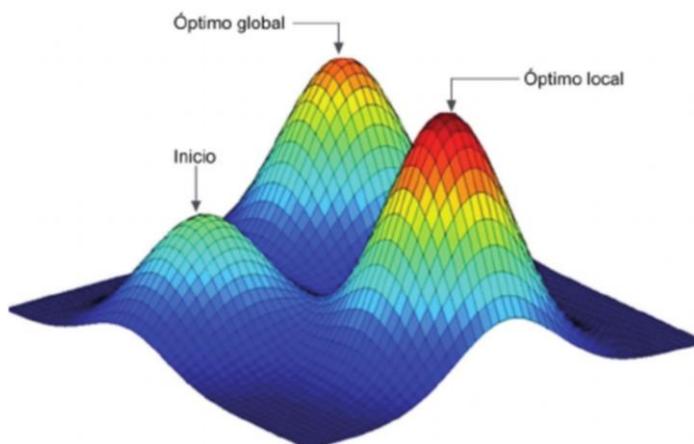


Figura 2.2: Óptimo local y óptimo global.

2.3. Metaheurísticas

En ciertos tipos de problemas de optimización puede resultar demasiado costoso al punto de resultar prácticamente inviable ser resuelto de forma exacta, por lo que algunos problemas o situaciones particulares, donde no es estrictamente necesario alcanzar una solución exacta, pueden ser resueltos mediante técnicas de optimización aproximadas. Esto permite solucionar el problema en cuestión en periodos considerablemente menores de tiempo. Los métodos de optimización aproximados brindan soluciones de alta calidad sin garantizar una solución óptima global, en contraste con los métodos de optimización exactos, los cuales sí aseguran encontrar soluciones óptimas a cambio de un alto costo computacional [66]. De los métodos de optimización aproximados, las metaheurísticas han cobrado importancia en los últimos veinte años debido a su capacidad de brindar soluciones satisfactorias a problemas de gran tamaño en plazos razonables de tiempo.

La palabra “heurística” se deriva del griego *heuriskein*, que significa “encontrar” o “descubrir”. El significado del término ha variado históricamente. Algunos han usado el término como un antónimo de “algorítmico”. Según [72], una metaheurística es un conjunto de conceptos que pueden ser usados para definir métodos heurísticos para ser aplicados a una amplia variedad de problemas. Es decir, una metaheurística puede ser vista como un marco general, para ser aplicado a diferentes problemas de optimización con mínimos cambios para ser adaptado a cada situación. Según [74] la *computación metaheurística* es vista como una metodología o tecnología de computación emergente

desarrollada en el campo de la computación genética, evolutiva, autónoma y basada en la búsqueda. Para tal afirmación, Wang se sustenta en los trabajos de [75, 76, 77, 66, 78, 79, 80, 81].

Definición 2 *La computación metaheurística es una metodología adaptativa y/o autónoma para la computación que aplica reglas heurísticas generales, algoritmos y procesos para resolver una categoría de problemas computacionales [74].*

2.3.1. Clasificación de las Metaheurísticas

Las metaheurísticas son algoritmos que recorren, mediante procedimientos iterativos, un espacio de búsqueda determinado con el propósito de obtener una solución aceptable al problema tratado. Las metaheurísticas pueden ser clasificadas en función de diferentes criterios [76]: por la cantidad de soluciones procesadas en cada ciclo de ejecución (iteración); la forma en que se generan nuevas soluciones en el espacio de búsqueda; el origen de inspiración del algoritmo; si operan con o sin memoria; con una o varias estructuras de vecindario, entre otras. Sin embargo, una de las clasificaciones más adoptadas las divide en metaheurísticas *basadas en trayectoria* y *basadas en población*. Las primeras manipulan un único elemento del espacio de búsqueda, mientras que las segundas trabajan sobre un conjunto de ellos (población). De esta manera, se puede observar que la diferencia principal entre ambas clasificaciones depende de la cantidad de soluciones candidatas utilizadas en cada iteración del algoritmo. A las metaheurísticas de trayectoria también se las conoce como metaheurísticas basadas en una única solución, a diferencia de las poblacionales las cuales operan con un conjunto de soluciones candidatas llamado población.

Respecto al modo de operación, las de trayectoria tienen la característica de mejorar de forma continua una única solución para explorar el espacio de búsqueda del problema. Este proceso se asemeja a una trayectoria que recorre el espacio de búsqueda, el cual pertenece al dominio de solución del problema [82]. Estas metaheurísticas se caracterizan por hallar rápidamente soluciones óptimas locales, por lo que usualmente se denominan métodos *orientados a la explotación*, promoviendo la intensificación en el espacio de búsqueda. Mientras que las metaheurísticas poblacionales o basadas en población, operan mejorando de forma iterativa un conjunto de soluciones candidatas. Las metaheurísticas poblacionales son muy adecuadas para resolver problemas de optimización de alta complejidad con rasgos de

incertidumbre y dinamicidad [83, 84], como sucede en los métodos de reducción de incertidumbre cuando son aplicados sobre un método en particular. Estas técnicas suelen llamarse métodos *orientados a la exploración* porque su capacidad principal depende de la diversificación en el espacio de búsqueda.

2.3.1.1. Metaheurísticas de Trayectoria

En general, las distintas metaheurísticas que caen bajo esta clasificación se caracterizan por comenzar con una única solución. La actualización de esta solución se realiza examinando el vecindario de dicha solución, generando así una trayectoria. La mayoría de estos algoritmos surgen como extensiones de los métodos simples de búsqueda local a los que se les añade algún mecanismo para escapar de los mínimos locales, debido a que su rendimiento es poco satisfactorio. Los criterios de parada deben ser más complejos que alcanzar un mínimo local; lo que habitualmente se utiliza es alcanzar un tiempo máximo de procesamiento o un número máximo de iteraciones, encontrar una solución *s* de determinada calidad o detectar que no se produjeron mejoras en las últimas iteraciones.

Algunos ejemplos de este tipo de metaheurísticas son:

- Enfriamiento Simulado (Simulated Annealing).
- Búsqueda Tabú (Tabu Search).
- Búsqueda Local Iterativa (Iterated Local Search).
- Búsqueda con Vecindario Variable (Variable Neighborhood Search).
- Búsqueda Local Guiada (Guided Local Search).

Durante el análisis de alternativas de hibridación, se contemplaron algunas estrategias basadas en metaheurísticas de trayectoria, pero que finalmente no fueron incorporadas a los métodos propuestos.

2.3.1.2. Metaheurísticas Poblacionales

Las metaheurísticas basadas en múltiples soluciones, también conocidas como poblacionales, poseen muchas características o conceptos en común. Éstas pueden verse como un proceso de mejora iterativa sobre una población de posibles soluciones. Primero, la población se inicializa generando una nueva población de soluciones candidatas al problema en cuestión. Posteriormente, las soluciones candidatas son alteradas, modificadas mediante diferentes operadores

que siguen algún comportamiento particular. Finalmente, esta nueva población se integra en la población actual utilizando algunos procedimientos de selección. Este proceso de búsqueda itera mejorando en cada iteración las soluciones candidatas, y se detiene cuando se cumple una condición dada (criterio de finalización). Algunos ejemplos de metaheurísticas poblacionales son:

- Metaheurísticas Evolutivas:
 - Algoritmos Evolutivos (Evolutionary Algorithms).
 - Algoritmos Genéticos (Genetics Algorithms).
 - Estrategias Evolutivas (Evolution Strategies).
 - Programación Evolutiva (Evolutionary Programming).
- Otras Metaheurísticas Evolutivas:
 - Estimación de Algoritmos Distribuidos (Estimation of Distribution Algorithms).
 - Evolución Diferencial (Differential Evolution).
 - Algoritmos Coevolutivos (Coevolutionary Algorithms).
- Metaheurísticas de Inteligencia de Enjambre:
 - Optimización por Cúmulo de Partículas (Particle Swarm Optimization).
 - Optimización por Colonia de Hormigas (Ant Colony Optimization Algorithms).

A continuación, se describen sintéticamente algunas metaheurísticas representativas de las distintas clases.

Algoritmos Evolutivos (Evolutionary Algorithms)

Los *Algoritmos Evolutivos* AE (Evolutionary Algorithms, EAs) se inspiran en la capacidad de adaptarse de los seres vivos y así evolucionar. Los EAs poseen un comportamiento totalmente estocástico, el cual modela la herencia genética y la supervivencia. Esta técnica de optimización, a diferencia de las metaheurísticas de trayectoria, opera sobre una *población* de soluciones candidatas, cada una de ellas denominadas *individuos*. De forma general la población es iniciada de forma aleatoria, pero también existen otras estrategias de inicialización. El esquema general de un AE está formado por tres pasos: la selección, la alteración o adaptación y el reemplazo. Estos pasos se repiten hasta que se alcanza alguna condición de finalización, normalmente después de un número dado de iteraciones o si se alcanza un determinado margen de error. Es importante mencionar que cada

individuo es evaluado mediante alguna función que permita cuantificar qué tan buena es la solución que dicho individuo representa para el problema tratado. Este valor se conoce como valor de aptitud o fitness.

Algoritmo 1: Pseudo-código de un Algoritmo Evolutivo.

```
Generar(P(0)) ; /*Población inicial*/
t = 0 ;
while no se cumpla Criterio_de_Finalización(P(t)) do
    Evaluar(P(t)) ;
    P'(t) = Selección(P(t)) ;
    P'(t) = Reproducción(P'(t)) ; Evaluar(P'(t)) ;
    P(t + 1) = Reemplazo(P(t), P'(t)) ;
    t = t + 1 ;
```

Output: Mejor individuo o mejor población encontrada.

El proceso comienza con la *selección* de algunos miembros de la población para someterlos a transformaciones por medio de determinados operadores, proceso que permite generar nuevas soluciones. Los operadores básicos de todo AE son: el operador de *cruciamiento*, el cual crea nuevos individuos al combinar varias partes de dos o más individuos, y el operador de *mutación*, el cual crea nuevos individuos al producir pequeños cambios en un único individuo de la población. Posteriormente, se conforma la nueva población, la cual deberá admitir o no a los nuevos individuos dando lugar a la siguiente generación del algoritmo. Esto corresponde con el último de los tres pasos, el cual *reemplaza* individuos en función de los criterios que hayan sido consignados previamente. Tras algún número de generaciones, el algoritmo converge; es de esperar que el mejor individuo de la población represente una solución cercana a la óptima.

Estos algoritmos establecen un equilibrio entre la exploración de buenas soluciones (fase de selección) y la exploración de nuevas zonas del espacio de búsqueda (fase de reproducción), basados sobre el hecho de que la política de reemplazo permite la aceptación de nuevas soluciones que no mejoran necesariamente las existentes. En la literatura se han propuesto diferentes algoritmos basados en este esquema general. Básicamente, estas propuestas se pueden clasificar en tres categorías que fueron desarrolladas de forma independiente:

- Algoritmos Genéticos o Genetic Algorithms (GA).
- Estrategias Evolutivas o Evolution Strategies (ES).
- Programación Evolutiva o Evolutionary Programming (EP).

Algoritmos Genéticos (Genetic Algorithms)

Los *Algoritmos Genéticos* (Genetic Algorithms, GAs) constituyen un método de búsqueda inspirado en la teoría de la evolución biológica (acorde a los mecanismos de selección y genética natural), para la resolución de problemas de optimización [79]. El proceso de búsqueda está compuesto por iteraciones llamadas generaciones donde un conjunto de soluciones candidatas, denominado población, evoluciona imitando los principios de selección natural y supervivencia del más apto [85] esperando que converja a una solución. En la situación más favorable la solución hallada será la óptima, pero normalmente suele definirse un umbral de lo que representa una solución aceptable.

Cada uno de los individuos de la población representa a una posible solución mediante un cromosoma, el cual puede ser representado mediante cadenas binarias, permutaciones, números enteros, números reales [86] (tradicionalmente, los GAs están asociadas con el uso de una representación binaria, pero hoy en día se pueden encontrar GA que usan otros tipos de representaciones [66]). Además, cada individuo posee un valor de aptitud o fitness, el cual cuantifica su capacidad de adaptación, es decir, de representar una buena solución para el problema tratado. En el Algoritmo 1 se puede observar un esquema básico de un GA donde t identifica el número de generaciones y $P(t)$ representa la población en la t –ésima generación. En cada iteración t , se escoge un conjunto de individuos de la población actual $P(t)$. El proceso de selección está diseñado para permitir que las buenas soluciones tengan mayor probabilidad de ser elegidas, para posteriormente aplicar entre los miembros de esta sub-población los operadores genéticos que permitan, de manera artificial, imitar el proceso de adaptación natural de: cruzamiento, mutación y reemplazo.

Puntualmente, el operador de selección lleva a cabo la elección de aquellos individuos que serán los progenitores de la nueva generación (está relacionado con la determinación de buenas soluciones) siendo sometidos a los operadores genéticos que mantienen la analogía con la metáfora de inspiración donde los más débiles perecen por selección natural. Este proceso debe favorecer a los más aptos, es decir, a aquellos individuos que posean un fitness mayor o menor en función del tipo de problema de que se trate (maximización o minimización respectivamente). Existen diversos métodos de muestreo para llevar a cabo la selección de individuos entre los que se destacan: muestreo estocástico por sorteo, muestreo estocástico universal o por ruleta, y muestreo estocástico por torneos.

Luego de la selección del conjunto de individuos (llamados criadores), éstos son sometidos al proceso de reproducción para generar nuevos individuos. Existen dos operadores de reproducción, por un lado, el operador de recombinación o cruce, el cual actúa sobre pares de individuos creando otro par de individuos descendientes. Este operador realiza un intercambio de información esperando que se exploten las mejores características presentes en la población actual. Por otro lado, el operador de alteración o mutación actúa sobre un solo individuo. La mutación consiste en realizar una pequeña modificación en alguno de los genes del individuo, de esta manera se introduce cierto grado de diversidad para intentar alcanzar todos los puntos del espacio de búsqueda permitiendo explorar nuevos dominios de búsqueda. Por último, se lleva a cabo el proceso de reemplazo sobre ciertos individuos de la población actual en comparación con los nuevos descendientes para determinar cuáles formarán parte de la población de la generación posterior. Del mismo modo que para el proceso de selección, aquí también existen distintos criterios para efectuar el reemplazo. Por ejemplo, una manera es reemplazar de forma inmediata los progenitores con los descendientes; otra opción es reemplazar a aquellos individuos de la población que posean características más similares a los descendientes o reemplazar aquellos individuos que posean los peores valores de fitness.

Evolución Diferencial (Differential Evolution)

El algoritmo de *Evolución Diferencial* (Differential Evolution, DE) [87], a gran des rasgos, cuenta con etapas semejantes a las que involucran los GAs pero con relevantes diferencias. En la etapa de inicialización de DE, la población de soluciones candidatas se genera de manera aleatoria y distribuida uniformemente dentro del rango del problema. De esta manera, la población queda constituida por Np vectores, cada uno con D dimensiones que conforman a la población Px , tal que se cumpla la Ecuación 2.4.

$$P_{x,g} = (x_{i,g}), i = 1, 2, \dots, Np, g = 1, 2, \dots, g_{max}, \quad (2.4)$$

$$x_{i,g} = (x_{j,i,g}), j = 1, 2, \dots, D.$$

$$v_{i,g} = x_{r0,g} + F(x_{r1,g} - x_{r2,g}), \quad (2.5)$$

donde $i = 1, 2, \dots, NP$ y $r1 \neq r2 \neq r3 \neq i$.

Donde g_{max} representa la cantidad máxima de generaciones. En DE, además de la población P_x , se manejan dos poblaciones adicionales denominadas P_v y P_u las cuales poseen las mismas características denotadas en la Ecuación 2.4. En cada una de estas poblaciones cada componente de los vectores representa una variable de decisión del problema. Posteriormente a la inicialización, los vectores son sometidos a procesos de mutación, recombinación y selección para cada generación. En términos generales, DE perturba la población mediante el concepto de diferencia de vectores, donde muta de forma aleatoria vectores de la población actual P_x para producir cada vector de la población intermedia P_v como se observa en la Ecuación 2.5. Existen varias estrategias para realizar el proceso de mutación. Éstas utilizan la notación $DE/x/y/z$ donde x indica el vector a ser mutado, éste puede ser *rand*, cuando es un vector de población elegido al azar, o *best* si se utiliza el vector de mejor aptitud de la población actual, y la cantidad de restas de vectores realizadas (1 o 2), y z denota el esquema de recombinación utilizado (bin: binomial o exp: exponencial). En la Ecuación 2.5 se puede observar la mutación correspondiente a la estrategia $DE/rand/1/bin$ donde el factor de escala F es un número real en el rango $[0, 1]$ [87]. Posteriormente, cada parte del vector de la población actual P_x es recombinado con su correspondiente del vector de la población intermedia P_v (vectores mutados), para producir la población de prueba P_u . En esta operación entra en juego la probabilidad de recombinación Cr , para el caso de utilizarse el crossover o recombinación binomial.

Por último, se efectúa la selección con el objetivo de mantener en la población de la generación siguiente a aquellos individuos que posean mejores características, por lo tanto, se evalúa para cada vector de prueba el valor de su función objetivo y se compara con su correspondiente de la población actual. Cada uno de estos operadores son procesados en cada generación tal como puede observarse en el Algoritmo 2, el cual describe al algoritmo clásico de DE.

Algoritmo 2: Pseudo-código de un Algoritmo de Evolución Diferencial DE/rand/1/bin.

Input : Parámetros: F (factor de escala), Cr (probabilidad de recombinación). Inicialización de la población (distribución aleatoria uniforme);

repeat

for ($i = 1, j \leq K, i ++$) **do**

 Mutar y Recombinar:

for ($j = 1, j \leq D, j ++$) **do**

if ($rand_j[0, 1] < Cr$) **or** ($j = j_{rand}$) **then**

$u_{ij} = v_{ij} = x_{r3j} + F \cdot (x_{r1j} - x_{r2j});$

else

$u_{ij} = x_{ij}$

 Reemplazo:

$$x_i(t+1) = \begin{cases} u_i(t+1), & \text{if } f(u_i(t+1)) \leq f(x_i(t)) \\ x_i(t), & \text{en otro caso.} \end{cases} \quad (2.6)$$

End for

until criterios de detención en: un número determinado de generaciones;

Output: La mejor población o solución encontrada.

Optimización por Cúmulo de Partículas (Particle Swarm Optimization)

La metaheurística conocida como *Optimización por Cúmulo de Partículas* (Particle Swarm Optimization, PSO) forma parte de la familia de algoritmos bioinspirados que mantienen analogía con distintos sistemas sociales o presentes en la naturaleza. Un ejemplo que puede citarse es la manera en que algunas especies comparten información para luego desplazarse a ubicaciones donde encuentren alimento, resguardo o mejor condición climática, donde el comportamiento de los mejores individuos afecta a toda la población [88]. Puntualmente PSO utiliza una población de partículas para explorar el espacio de búsqueda, de forma similar a las metaheurísticas anteriormente presentadas (cada miembro de la población representa una solución candidata del problema de optimización tratado). Las partículas poseen un valor de posición y de velocidad, los cuales son actualizados en cada iteración del algoritmo. En términos generales, PSO maneja la población de soluciones como un cúmulo de partículas donde cada partícula i consta de tres componentes. Por un lado x_i representa la posición de partícula dentro del espacio de búsqueda, v_i describe la velocidad de la partícula y $p_{i\text{best}}$ es la memoria de la partícula, lo cual representa la mejor posición encontrada hasta el momento. Por otro lado, el algoritmo maneja una componente social (o

global a la población) denominada p_gbest , que representa la mejor posición encontrada por alguna partícula en toda la población [89]. La velocidad de cada partícula y su posición son actualizadas en cada iteración del algoritmo, de acuerdo a las Ecuaciones 2.7 y 2.8 respectivamente.

$$v_i(t + 1) = w * v_i(t) + i_{local} + i_{global}, \quad (2.7)$$

donde $\begin{cases} i_{local} = c_1 u_{1i} [p_i best(t) - x_i(t)], \\ i_{global} = c_2 u_{2i} [p_g best(t) - x_i(t)], \end{cases}$

$$x_i(t + 1) = x_i(t) + v_i(t + 1) \quad (2.8)$$

Donde i_{local} (información local) representa el componente o factor cognitivo o personal e i_{global} (información global) el componente o factor social. Los factores c_1 y c_2 son pesos que regulan la contribución de la componente cognitiva, el primero, y de la componente social, el segundo. Los factores u_{1i} y u_{2i} corresponden a valores aleatorios distribuidos uniformemente en el rango $[0, 1]$. Y por último, el factor w equilibra las capacidades de búsqueda global y local [90], esto lo hace controlando la velocidad previa de la partícula a ser mantenida en el ciclo actual. Es importante mencionar que el factor de inercia w no forma parte del algoritmo básico PSO, y fue introducido en [91]. Un esquema de PSO puede verse en el Algoritmo 3.

Algoritmo 3: Pseudo-código de un algoritmo de Optimización por Cúmulo de Partículas

Inicialización aleatoria de todo el enjambre;

repeat

 Evaluar $f(x_i)$;

repeat

 Actualizar velocidades:

$$v_i(t) = v_i(t - 1) + \rho_1 \times (p_i - x_i(t - 1)) + \rho_2 \times (p_g - x_i(t - 1));$$

 Mover a nueva posición: $x_i(t) = x_i(t - 1) + v_i(t)$;

if $f(x_i) < f(pbest_i)$ **then**

 └ $pbest_i = x_i$;

if $f(x_i) < f(gbest)$ **then**

 └ $gbest = x_i$;

 Actualizar (x_i, v_i) ;

until hasta evaluar todas las partículas i ;

until Criterio de parada;

Output: Mejor población o solución encontrada.

En el mismo se puede observar que, en primer lugar, la población es inicializada de forma aleatoria, luego, mediante una función de evaluación (que no necesariamente debe coincidir con la función objetivo), se determina para cada partícula su nivel de aptitud. Este proceso iterativo mejora gradualmente la dirección que las partículas deben tomar, con el objetivo de encontrar la región del espacio de búsqueda que contiene las buenas soluciones.

Optimización por Colonia de Hormigas (Ant Colony Optimization)

Como bien lo indica su nombre, el algoritmo de *Optimización por Colonia de Hormigas* (Ant Colony Optimization, ACO) [92] es una metaheurística inspirada en el comportamiento de colonias de hormigas. La metáfora se centra en el principio de comunicación entre hormigas, lo cual ocurre mediante el rastro de feromonas que cada hormiga va dejando a lo largo de su camino recorrido. De esta manera, mientras n hormigas aisladas se mueven con un comportamiento aleatorio, los agentes de una colonia de hormigas detectan dicho rastro y tienden a seguirlo. Éstas, a su vez, van dejando su propia feromona a lo largo del camino recorrido y, por tanto, lo hacen más atractivo, puesto que se ha reforzado el rastro de feromona. Sin embargo, la feromona también se va evaporando con el paso del tiempo provocando que el rastro de feromona sufra cierto debilitamiento. En definitiva, puede decirse que el proceso se caracteriza por una retroalimentación positiva, en la que la probabilidad con la que una hormiga escoge un camino aumenta con el número de hormigas que previamente hayan elegido el mismo camino (Algoritmo 4).

Algoritmo 4: Pseudo-código de un Algoritmo de Optimización por Colonia de Hormigas.

Inicializa los rastros de feromonas;

repeat

for *Cada hormiga* **do**

 Construcción de solución usando el rastro de feromonas;

 Actualiza los rastros de feromonas;

 Evaporación;

 Reforzamiento;

until *Criterio de parada*;

Output: Mejor solución encontrada o conjunto de soluciones.

2.4. Metaheurísticas Paralelas

Desde una mirada general, se puede decir que el desarrollo de metaheurísticas paralelas surge con el propósito de atacar dos objetivos principales: resolver instancias de problemas cada vez más grandes de forma más rápida (abordar problemas que no pueden ser resueltos de forma secuencial), y brindar mayor robustez a los métodos de búsqueda. Según Talbi [66], el paralelismo y la computación distribuida se pueden utilizar en el diseño e implementación de metaheurísticas, persiguiendo al menos alguno/s de los siguientes beneficios (motivos):

- *Acelerar el proceso de búsqueda*: Este es uno de los principales objetivos que se persiguen al momento de paralelizar metaheurísticas. La reducción del tiempo de búsqueda permite su aplicación en problemas dinámicos en tiempo real.
- *Mejorar la calidad de las soluciones*: Evidentemente, una mera implementación paralela de una metaheurística secuencial no permitirá mejorar la calidad de las soluciones encontradas. Sin embargo, este propósito sí será alcanzado si el paralelismo es utilizado para realizar una cooperación entre diferentes metaheurísticas o metaheurísticas iguales parametrizadas de forma distinta.
- *Robustez*: La robustez de una metaheurística se refiere al grado de sensibilidad que dicha metaheurística tiene en sus parámetros. En otras palabras, se puede decir que una metaheurística paralela será más o menos robusta si logra resolver, o no, de manera efectiva, diferentes problemas de optimización, y diferentes instancias de un mismo problema.
- *Resolver grandes y complejos problemas*: Este puede ser uno de los factores más importantes, junto con el primero, ya que, para cierto tipo de aplicaciones o problemas, puede ser completamente inviable intentar resolverlo de forma secuencial.

En concordancia con la clasificación presentada en la sección 2.3.1, las estrategias de paralelización de metaheurísticas difieren entre sí, ya sea para las metaheurísticas de trayectoria o para las poblacionales.

Para las *metaheurísticas de trayectoria* se han propuesto diferentes modelos paralelos, entre los cuales se utilizan con mayor frecuencia los siguientes: a) exploración y evaluación paralela del vecindario (o modelo de movimientos paralelos), b) modelo multi-inicio paralelo, y c) evaluación paralela de única solución (o modelo de aceleración de movimiento) [93].

- Modelo de movimientos paralelos: se define como un modelo master-worker. Ofrece el mismo resultado que una implementación secuencial, sólo que logra reducir el tiempo de procesamiento. El funcionamiento es sencillo, el maestro duplica la solución actual y, por cada una de ellas, se realiza la búsqueda de una solución candidata, finalmente los resultados se entregan al master.
- Modelo de multi-inicio paralelo: esta metodología consiste en lanzar simultáneamente diferentes instancias de la metaheurística de trayectoria secuencial.
- Modelo de aceleración de movimiento: en este modelo la calidad de cada movimiento se evalúa de forma paralela y centralizada.

Algoritmo 5: Esquema general del algoritmo evolutivo paralelo.

```

inicio
g = 0 /*número de generación*/
generar p poblaciones con n individuos
calcular la función de evaluación de cada individuo
while no se cumpla condicion de fin do
    /*Producir nueva generación*/
    seleccionar progenitores
    recombinar pares de progenitores
    mutar los descendientes con cierta probabilidad
    calcular la función de evaluación de cada individuo
    reemplazar los descendientes en la nueva generación
    if g % frecuencia == 0 then
        /* realizar migración de individuos*/
        enviar m mejores individuos a población vecina
        recibir m individuos de población vecina
        reemplazar m individuos en población actual
        evaluar cada individuo con la función objetivo
    g++
fin

```

Para las *metaheurísticas poblacionales* se puede decir que el paralelismo se encuentra implícito o que surge de manera natural, ya que cada solución candidata o individuo puede ser tratado de forma independiente al resto. Desde el punto de vista de paralelización de las metaheurísticas poblacionales, y tomando el enfoque propuesto por [94] se puede decir que existen tres diferentes aproximaciones: Única

Población y Evaluación en Paralelo, Única Población y Vecindarios Solapados, y Múltiples Poblaciones y Migración. En el esquema de Única Población y Evaluación en Paralelo, se trabaja con un paradigma del tipo master worker [61], con una población panmíctica donde la única carga de procesamiento que es distribuida a los workers es la evaluación de aptitud. La segunda aproximación consiste en trabajar con una población espacialmente estructurada, donde las operaciones de búsqueda se limitan a un pequeño vecindario, pero los mismos se superponen permitiendo una interacción entre todos los individuos. Y, por último, el Modelo de Islas opera con múltiples poblaciones donde los operadores se aplican entre individuos de la misma isla, y cada cierto tiempo o generaciones se intercambian individuos entre las diferentes islas.

Debido a que la presente tesis se centra en la utilización de metaheurísticas paralelas con características evolutivas de tipo poblacional, a continuación, se realiza una aproximación más detallada de las estrategias de paralelización utilizadas en las implementaciones vinculadas con este modelo.

2.4.1. Paralelización de Metaheurísticas Poblacionales Evolutivas

Las metaheurísticas poblacionales evolutivas son métodos de optimización, eficaces y eficientes, que pueden ser aplicados en la resolución de problemas de alta complejidad, dimensionalidad y dinamicidad. De forma general, se puede decir que estas metaheurísticas producen un conjunto de posibles soluciones que son sometidas a un proceso de iteraciones. En cada iteración, se aplican ciertas operaciones (*selección, cruzamiento, mutación y reemplazo*, varían según el tipo de metaheurística) que permiten realizar búsquedas de mejores soluciones al problema tratado. Cada solución candidata (individuo, partícula, vector, etc.) posee un valor que de termina la calidad de la solución. Éste valor suele conocerse como *aptitud* o *fitness*, y se calcula utilizando la función matemática que describe el problema. El valor de aptitud juega un rol fundamental, ya que permite caracterizar a las posibles soluciones según sus capacidades. De esta manera, aquellas con mejor aptitud deben tener mayor probabilidad de permanecer, participar en la generación de nuevas soluciones, de no ser reemplazados y, de esta forma, dejar sus rasgos en las poblaciones futuras [95].

Cuando estas metaheurísticas son utilizadas para resolver problemas complejos y/o de grandes dimensiones, suelen ser paralelizadas, dividiendo la carga de trabajo en N unidades de cómputo. Dentro de este tipo de metaheurísticas se encuentran los conocidos Algoritmos Evolutivos Paralelos, AEP (Parallel Evolutionary Algorithms, PEA). Un esquema general de un AEP se presenta en el Algoritmo 6. Tal como se puede observar, el proceso comienza con la generación de p poblaciones de n soluciones candidatas, las cuales suelen ser creadas mediante un proceso estocástico. Una vez que las poblaciones han sido generadas, se evalúa la aptitud inicial de cada solución, aplicando la función de fitness a cada una. Luego, comienza el ciclo evolutivo, el cual itera mientras no se cumpla la condición de finalización, generalmente hasta que se alcance un cierto umbral de aptitud o una máxima cantidad de iteraciones.

Posteriormente se realiza el proceso de perturbación, el cual consiste, según la metaheurística utilizada, en diferentes tipos de sub-procesos. Entre ellos se destaca la selección de soluciones candidatas a ser utilizadas para generar nuevas soluciones, proceso que suele brindar mayor probabilidad de elegir soluciones con mejores características.

Dichas soluciones son utilizadas para la generación de las nuevas soluciones; las operaciones que se involucran en esta etapa suelen consistir en la cruce de diferentes soluciones y la modificación o mutación. Las nuevas soluciones deben ser introducidas mediante algún tipo de proceso y criterio en la nueva población o conjunto de soluciones. Este procedimiento suele llamarse reemplazo, según la metaheurística involucrada.

Según la cantidad de poblaciones involucradas, el tratamiento de los operadores, la relación cómputo/comunicación y demás factores, los AE se pueden paralelizar siguiendo diferentes estrategias. Según [96] la clasificación de los AEP se puede resumir en cuatro grupos: paralelismo automático a nivel del compilador, master-worker con única población y evaluación en paralelo, paralelismo de grano grueso, distribuido o basado en islas (múltiples poblaciones) y de grano fino o masivamente paralelo. Las dos primeras estrategias sólo mejoran el tiempo de resolución del algoritmo, pero dicho paralelismo no influye en la calidad de los valores numéricos hallados.

En [97] se expone una taxonomía donde los modelos distribuidos de EAs se dividen en dos grandes grupos: *modelo de distribución por poblaciones* y *modelo de distribución por dimensiones*. La primera clasificación, consiste en distribuir individuos de una población o subpoblación en diferentes unidades de cómputo, y la segunda,

distribuye particiones de la dimensión del problema, en problemas más pequeños. A su vez, la clasificación de distribución basada en poblaciones, puede ser dividida en: modelo master-worker, modelo de islas, modelo celular o modelo jerárquico. Ambas clasificaciones coinciden en gran parte y mantienen diferencias menores en ciertos aspectos. Las estrategias de paralelización que han sido utilizadas en este trabajo, corresponden con las tres clasificaciones siguientes: modelo master-worker o master-worker, modelo de islas y modelo jerárquico.

El *modelo master-worker* (ver Fig. 2.3.a), se basa en el paradigma master-worker [95], donde una única población de individuos es administrada por un proceso central, usualmente denominado *master*. El master lleva a cabo gran parte de las funciones del AEP (selección, cruzamiento, mutación y reemplazo). En este esquema, sólo la evaluación de la aptitud de los individuos se realiza de forma paralela por los procesos *workers*. El *modelo de islas* (ver Fig. 2.3.b), utiliza múltiples poblaciones o islas, las cuales evolucionan de forma independiente. Cada población puede ser administrada por un worker o conjuntos de workers, dependiendo del nivel de paralelización con el cual el algoritmo sea implementado. Cada cierto número de generaciones, se produce un intercambio de individuos entre las diferentes islas o poblaciones, este proceso permite incrementar el grado de diversidad y minimizar la probabilidad de convergencia prematura o estancamiento en óptimos locales. El *modelo jerárquico* consiste en sumar las bondades de dos o más modelos diferentes en beneficio de la escalabilidad y/o la capacidad de búsqueda del método. Algunos de estos esquemas son: modelo de islas con master-worker, modelo de islas con modelo de islas y modelo de islas con doble jerarquía master-worker.

En la Fig. 2.3.c se observa un esquema de paralelismo basado en islas, tanto a nivel de jerarquía superior como inferior. Dentro de cada isla la paralelización también utiliza un esquema basado en islas. Esta estrategia de paralelización incluye un proceso de migración adicional dentro de cada isla de jerarquía mayor. En la Fig. 2.3.d, se puede observar un esquema del modelo jerárquico basado en islas con master-worker. En este esquema, el primer nivel de paralelismo corresponde a un modelo de islas con migración, y en cada isla se aplica un modelo master worker. Finalmente, en la Fig. 2.4, se observa un esquema basado en islas con doble jerarquía master-worker.

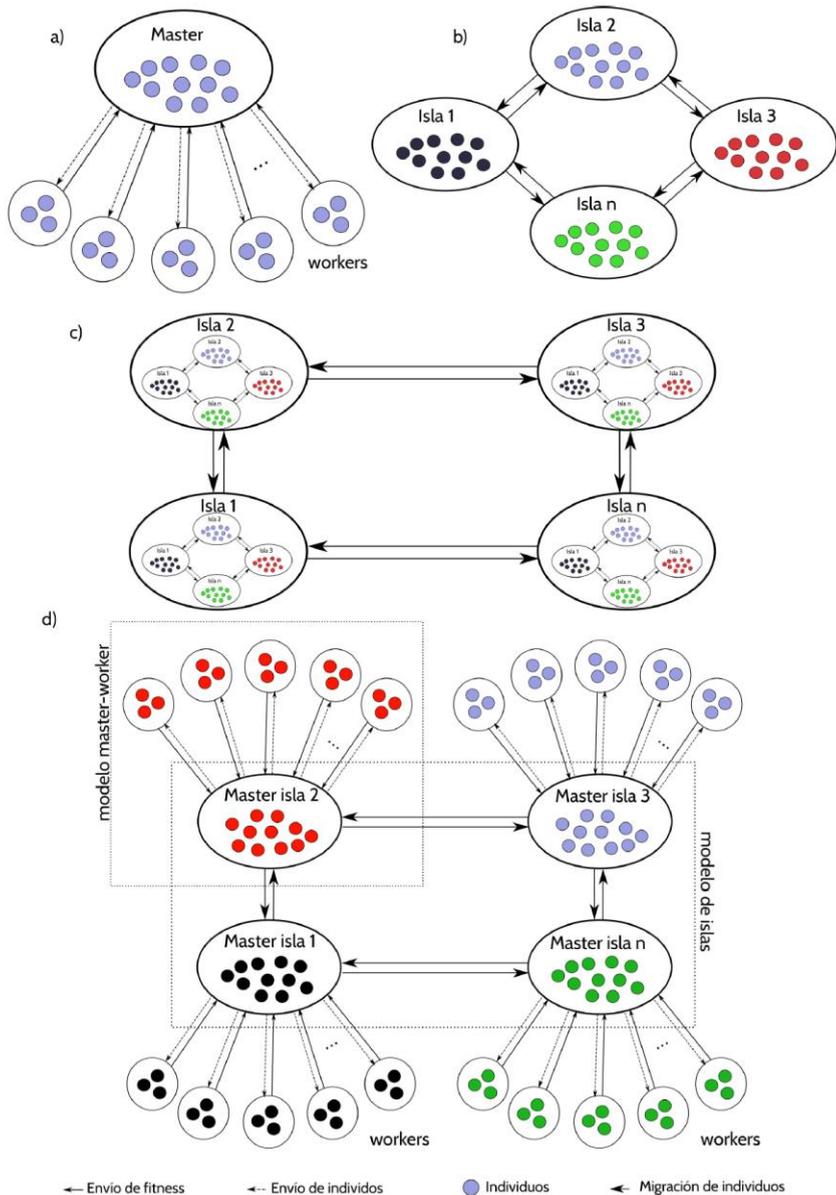


Figura 2.3: a) Esquema de paralelización master-worker. b) Esquema de paralelización basado en islas. c) Esquema de paralelización jerárquico basado en islas con modelo de islas. d) Esquema de paralelización jerárquico basado en islas con master-worker.

En el nivel superior, indicado como *modelo master-worker 1*, un proceso *master global* coordina el funcionamiento de la totalidad de las

islas del AEP. El master de cada isla cumple la función de worker respecto del master global. Posteriormente, en el segundo nivel de jerarquía master-worker (*modelo master worker 2*), cada master de isla controla a un conjunto de workers dentro de cada isla, quienes son los encargados de evaluar la función de aptitud. Es importante remarcar que los operadores de cruzamiento, mutación y reemplazo, son aplicados entre individuos de la misma isla. Evidentemente cada estrategia de paralelización ofrece diferentes ventajas en función de la implementación y el problema a abordar.

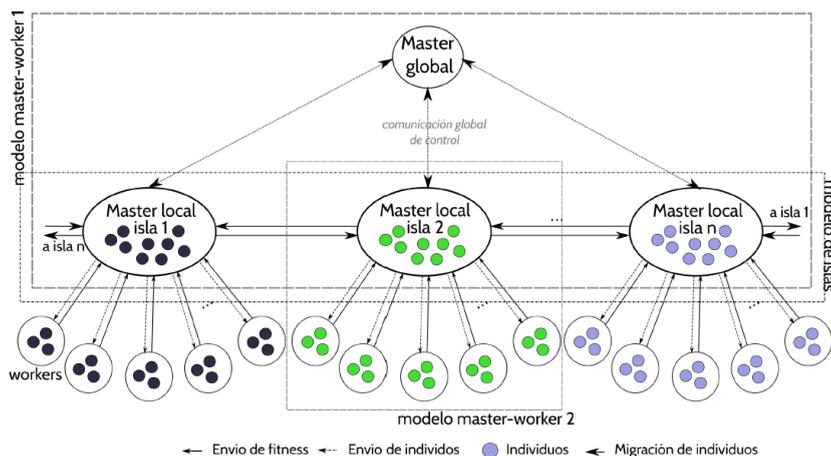


Figura 2.4: Esquema de paralelización jerárquico basado en islas con doble modelo master-worker.

2.5. Síntesis

Las técnicas de optimización han cobrado gran importancia en la resolución de problemas de alta complejidad computacional, entre ellos se destacan los problemas dinámicos, los cuales se relacionan íntimamente con los sistemas de predicción. Dentro de las herramientas disponibles para la resolución de este tipo de problemas, las metaheurísticas han demostrado ser muy robustas, eficientes y adaptables para resolver, de forma aproximada, gran cantidad de problemas de la vida real. En este capítulo se ha presentado una clasificación de estas técnicas como así también una breve descripción de las que posteriormente serán abordadas en mayor detalle en la presente tesis tras haber sido utilizadas en el diseño de los métodos presentados, estas son: Algoritmos Evolutivos, Evolución Diferencial y Optimización por Cúmulo de Partículas.

Predicción de Incendios mediante Métodos Guiados por Datos

El presente capítulo realiza una presentación de los métodos guiados por datos, sus variantes y las metodologías desarrolladas bajo estos conceptos.

3.1. Introducción

Tal como se ha formulado en el Capítulo 1, la predicción del comportamiento de un incendio forestal es una tarea compleja que no puede ser llevada a cabo, de forma aceptable, precisa y efectiva mediante la aplicación de procedimientos clásicos de predicción. Esto se debe principalmente a la presencia de tres factores que aportan incertidumbre al proceso de predicción: las dificultades para medir los parámetros de entrada en tiempo real, la dinamicidad del fenómeno y las limitaciones de los modelos de comportamiento de incendios forestales. Cada una de estos factores genera, en mayor o menor medida, diversos grados de incertidumbre, tanto interna como externa al modelo, lo que impide obtener predicciones cercanas a la realidad.

Tal como se observó en el esquema general de la predicción clásica presentado en el Capítulo 1 (Fig. 1.7), la predicción generada depende del resultado de una *única* simulación utilizando un *único* conjunto de parámetros de entrada. Es decir, los métodos de predicción clásica operan con una etapa que, a través de una única simulación, genera la predicción del frente de fuego para un instante de tiempo determinado. Este tipo de enfoque presenta grandes limitaciones, las cuales se hacen más notables cuando el comportamiento del fenómeno que se intenta predecir es fuertemente dependiente de factores dinámicos. En el caso de los incendios forestales, una de las principales fuentes de error se origina en la falta de precisión en los parámetros de entrada. Por ejemplo, existen ciertos parámetros que no pueden ser medidos de forma directa, por lo que sus valores deben ser establecidos a partir de mediciones indirectas, tal es el caso del contenido de humedad en la vegetación. También existen otros parámetros, como la velocidad y la dirección del viento, que suelen medirse en ciertos puntos del terreno,

por lo que los valores deben interpolarse a toda la superficie. Además, es importante tener en cuenta que estas variables no son independientes entre sí, es decir que el cambio en una de estas suele modificar el valor de las demás. Por lo tanto, tales aproximaciones evidentemente generan grandes imprecisiones que afectan en mayor o menor medida la calidad de salida del modelo. Además, la precisión de todo modelo no sólo depende de las incertidumbres externas, sino también de las internas, propias de la implementación del simulador. Debido a esto, ha sido necesario el desarrollo de nuevos enfoques que permitan sortear estas problemáticas y mejorar así la calidad de predicción.

Antes de continuar es importante destacar los conceptos fundamentales que terminan las diferencias esenciales entre simulación y predicción, para ello nos basaremos en el trabajo de Nelles [48].

3.1.1. Predicción

En la Fig. 3.1 se presenta un esquema del modelo de predicción de un paso. Como se puede observar, el modelo es alimentado con las entradas y salidas del proceso a modelar para un instante de tiempo anterior: $u(k - i)$ y $y(k - i)$, respectivamente. El modelo, con dicha información, predice la salida del proceso $\hat{y}(k)$ para el siguiente instante de muestreo. Generalmente, los procesos reales no tienen alimentación directa o en tiempo real, por lo tanto la salida $\hat{y}(k)$ no puede verse afectada inmediatamente por la entrada $u(k)$, siempre que el muestreo no se realice exactamente al mismo tiempo. En aquellos casos en que $u(k)$ no afecta a $\hat{y}(k)$ de forma instantánea la primera puede no ser incluida en el modelo. Sin embargo, los modelos que incluyen a $u(k)$ generalmente ofrecen mejor rendimiento dado que el efecto de u en y puede ser mucho más rápido que un intervalo de tiempo de muestreo [48]. Para este tipo de problemas de predicción, es necesario medir las salidas del proceso (o los estados).

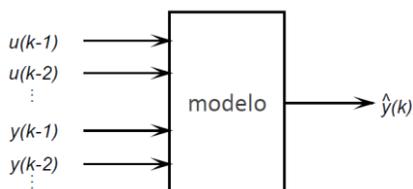


Figura 3.1: Representación gráfica de un modelo de predicción de un paso.

En el caso de predecir más de un paso en el futuro, esto se denomina predicción *multi-paso* o *l-step*. El número de pasos predichos en el futuro l se denomina horizonte de predicción.

3.1.2. Simulación

A diferencia de la predicción, en los procesos de simulación sólo se cuenta con las entradas del proceso ($u(k - 1), u(k - 2)$, ver Fig. 3.2), y la salida del modelo representa una estimación de su comportamiento. La simulación se utiliza con mucha frecuencia en diferentes áreas de la ciencia y la ingeniería. La diferencia fundamental entre la simulación y la predicción es que la simulación generalmente requiere componentes de retroalimentación dentro del modelo. Esto hace que la fase de identificación y modelado sea más difícil y requiere cuidados adicionales para garantizar la estabilidad del modelo.

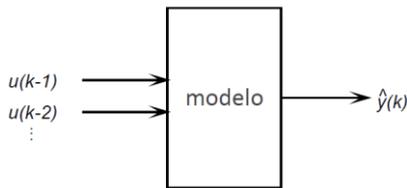


Figura 3.2: Representación gráfica de un modelo de simulación.

3.2. Métodos de Predicción Guiados por Datos

El concepto de *predicción guiada por datos* o *predicción de dos etapas* plantea la incorporación de una nueva fase o etapa en la cadena de predicción de los métodos clásicos (Fig. 3.3) [98], [63]. Dicha etapa se encarga de efectuar una calibración o ajuste de los parámetros de entrada del modelo antes de realizar la predicción. Este proceso consiste en efectuar una búsqueda iterativa que permita obtener parámetros mejorados (calibrados) que posibiliten realizar una predicción más cercana a la realidad, reduciendo de esta manera los efectos negativos causados por la incertidumbre. Esto implica que este tipo de métodos necesitan de un estado adicional de la evolución del proceso real del incendio para poder generar la predicción.

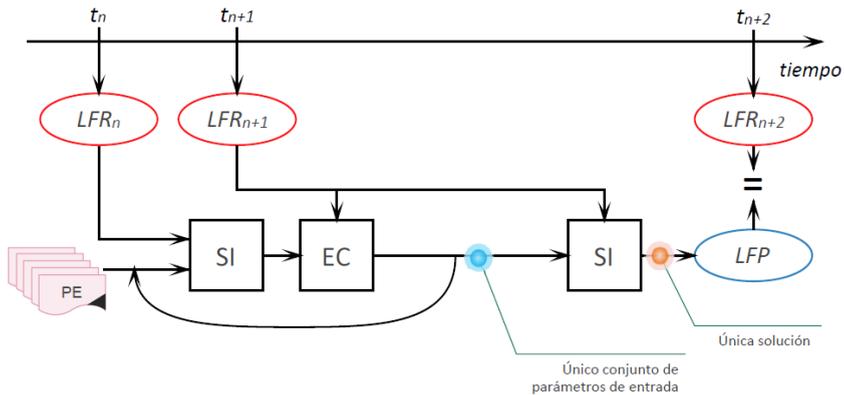


Figura 3.3: Esquema de Predicción Guiada por Datos. (LFR_x : estado real en tiempo x , SI: simulador de propagación de incendios forestales, LFP: estado de la línea de fuego propuesta por el modelo, EC: etapa de calibración, PE: grandes conjuntos de parámetros de entrada).

Este esquema se muestra en la Fig. 3.3. Como se puede observar, el simulador de incendios (SI) es alimentado por un gran conjunto de parámetros de entrada, los cuales son calibrados en una nueva etapa denominada Etapa de Calibración (EC). Para evaluar la calidad de los resultados obtenidos por el simulador y la etapa de calibración, se debe realizar una comparación entre la línea de fuego predicha y la línea de fuego real. Producto de la calibración se obtiene un único conjunto de parámetros de entrada (parámetros calibrados), los cuales son utilizados para realizar la predicción del instante de tiempo siguiente (t_{n+2}). De esta manera se puede decir que el enfoque de la predicción guiada por datos consiste en el desarrollo de métodos que utilizan técnicas de optimización con el fin de calibrar los parámetros de entrada. De esta manera, el proceso de optimización, se centra en encontrar un conjunto de valores de entrada que, si alimentan el simulador, permitan describir (con cierto nivel de confianza) el comportamiento previo del fenómeno, esperando que el mismo conjunto de valores sirva para describir o aproximar un instante futuro cercano.

Obviamente, los resultados generados por el método poseen validez temporal, es decir, que este proceso no será útil si el tiempo de calibración es superior al tiempo transcurrido entre (t_{n+1}) y (t_{n+2}) (Fig. 3.3). Por este motivo suele ser necesario, además de incrementar la calidad de predicción, acelerar el proceso de calibración. Este concepto se relaciona con el paradigma de computación urgente (UC, Urgent Computing) [58], en donde los sistemas poseen requerimientos de procesamiento y generación de resultados dentro de un lapso

determinado de tiempo, fuera de este intervalo los datos ya carecen de valor por lo que deben ser descartados.

3.2.1. Urgent Computing

UC requiere que los cálculos comiencen lo antes posible y se completen dentro de un plazo estipulado de tiempo a fin de brindar apoyo en la mitigación, respuesta y recuperación ante un evento que requiere atención inmediata. El resultado de los cálculos permite tomar decisiones rápidas y oportunas para mitigar las diferentes pérdidas. UC se define bajo las siguientes condiciones [58]:

- el cálculo opera bajo un plazo estricto, después del cual los resultados carecen de valor práctico.
- el inicio del evento que requiere el cálculo es impredecible.
- el cálculo requiere un uso de recursos significativo.

Algunas terminologías de UC:

- **caso o evento urgente**, es un evento que requiere atención inmediata, por ejemplo: una inundación, o un problema de alto impacto, como una fusión nuclear en un reactor. Los efectos pueden ser mitigados con el apoyo de sistemas u operaciones de cómputo, las cuales deben ser realizadas dentro de un estricto lapso temporal.
- **servicio urgente**, se refiere al grupo de actividades que brindan un servicio de carácter urgente, es decir, recolección de datos, comunicación, computación, toma de decisiones y trabajo de coordinación, para cumplir con lo requerido.
- **cálculo urgente**, es una actividad que debe comenzar en poco tiempo, es decir inmediatamente o tan pronto como sea posible, para simular o predecir un evento urgente. Un cálculo urgente se desencadena por un evento urgente.
- **producto urgente**, es el resultado procesado de un cálculo urgente que puede ser directamente utilizado para la toma de decisiones y actividades de coordinación para apoyar la mitigación de pérdidas.
- **sistema urgente**, es un sistema que permite cálculos urgentes para la toma de decisiones para la coordinación de acciones urgentes que permitan la mitigación de pérdidas y daños.

3.2.2. Métodos Guiados por Datos

Los métodos guiados por datos (Data Driven Methods, DDM) surgen con el propósito de subsanar las limitaciones impuestas por las técnicas clásicas, principal mente debido a la incertidumbre o falta de precisión en los valores de entrada de los modelos. De esta manera, los métodos realizan un ajuste para obtener una combinación *óptima* de parámetros de entrada. Sin embargo, como los DDM obtienen un único conjunto de valores de los parámetros de entrada, para aquellos parámetros de comportamiento dinámico, el valor encontrado no suele ser de gran utilidad para describir correctamente el comportamiento del modelo. Esta categoría de métodos se denomina Métodos Guiados por Datos de Única Solución (DDM-US, Data Driven Methods with Unique Solution) [99], [98], ver Fig. 3.4.

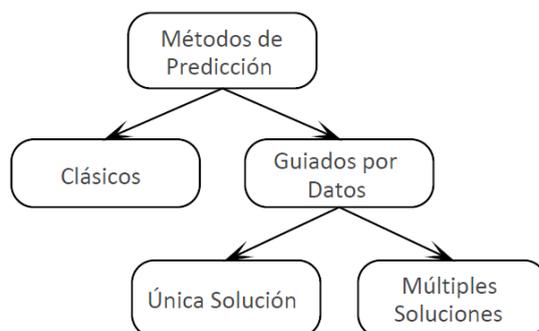


Figura 3.4: Clasificación de métodos de predicción.

Otra clasificación de los DDM trabaja con solapamiento de diversos casos o combinaciones de parámetros para efectuar las predicciones: Métodos Guiados por Datos con Solución Múltiple Solapada (DDM-MOS, Data Driven Methods with Multiple Overlapping Solutions) [63]. Una representación gráfica de este tipo de métodos se puede observar en la Fig. 3.5. Esta clasificación considera un gran número de combinaciones de parámetros de entrada, con los cuales se alimenta las etapas siguientes efectuando las predicciones en base a la totalidad de combinaciones de parámetros previamente calibrados.

En esta clasificación se encuentran los siguientes métodos desarrollados:

- Sistema Estadístico para la Gestión de Incendios Forestales (Statistical System for Forest Fire Management, S²F²M) [100].

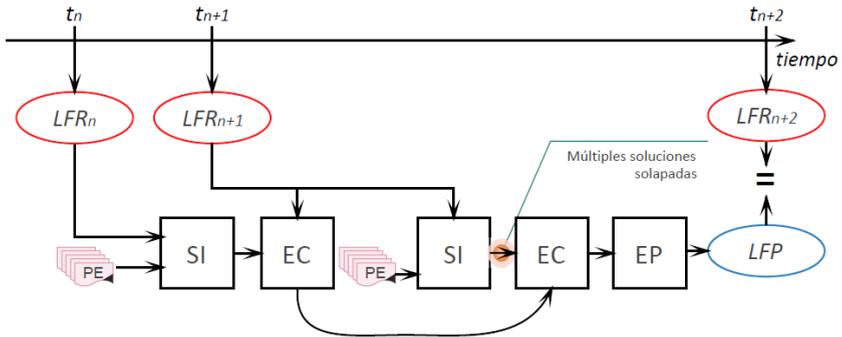


Figura 3.5: Esquema de Predicción Guiada por Datos con Múltiples Soluciones Solapadas. (LFR_x : estado real en tiempo x , SI: simulador de propagación de incendios forestales, LFP: estado de la línea de fuego propuesta por el modelo, EC: etapa de calibración, PE: grandes conjuntos de parámetros de entrada, EP: etapa de predicción).

- Sistema Estadístico Evolutivo (Evolutionary Statistical System, ESS) [101].
- Sistema Estadístico Evolutivo con Modelo de Islas (Evolutionary Statistical System with Island Model, ESS-IM) [102].
- Sistema Estadístico Evolutivo con Modelo de Islas y Evolución Diferencial (Evolutionary Statistical System with Island Model - Differential Evolution, ESS-IM-DE) [103].
- Sistema Estadístico Híbrido Evolutivo con Modelo de Islas (Hybrid Evolutionary Statistical System with Island Model, HESS-IM) [104].

Los DDM están fuertemente relacionados con los Métodos de Asimilación de Datos (Data Assimilation Methods, DAM), que se caracterizan por la incorporación de datos en un modelo en ejecución [105]. Desde este punto de vista, se puede decir que los métodos listados anteriormente corresponden a una conjunción entre DDM-MOS y DAM, esto se debe a que los mismos no tienen como objetivo encontrar un solo conjunto de datos con valores específicos. En cambio, el objetivo es encontrar un conjunto de valores que permita mejorar el resultado final y también incluir la incorporación de datos en el momento de la ejecución. Otro enfoque que guarda especial relación con los DDM-MOS es el de los Sistemas de Aplicación Guiados por Datos Dinámicos (DDDAS, Dynamic Data Driven Application Systems). El concepto de los DDDAS consiste en incorporar datos de forma dinámica y dotar a las aplicaciones con la capacidad de dirigir dinámicamente los procesos de medición. Algunas entradas de datos

pueden ser adquiridas en tiempo real (on-line) o pueden ser obtenidos a través del almacenamiento local. DDDAS ofrece la promesa de una mejora en los métodos de modelado, aumentando las capacidades de predicción, análisis y mejorando la eficiencia de las simulaciones [106].

El análisis preciso y la predicción del comportamiento de un sistema complejo siguen siendo tareas muy difíciles en la actualidad. Hoy en día, los sistemas de simulación son bastante complejos y además carecen de la capacidad de describir con precisión los fenómenos que se pretenden representar. Incluso los modelos complejos de sistemas producen simulaciones que difieren, o no predicen la realidad de forma aceptable, situación que se acentúa en los casos en los que existen condiciones que varían de forma dinámica en el tiempo. Las simulaciones de aplicaciones que incorporan dinámicamente nuevos datos, archivados o a partir de mediciones en línea de los sistemas actuales, aseguran análisis, predicciones y controles más precisos, así como de resultados más fiables [107].

Un DDDAS puede ser visto como una metodología para mejorar los resultados de los modelos mediante la utilización de datos adicionales incorporados en tiempo de ejecución para mejorar o refinar selectivamente el modelo original. Además, el paradigma DDDAS fomenta la capacidad de una aplicación de simulación que permita controlar y guiar las mediciones, creando de esta manera una herramienta más eficaz para dichos procesos. La capacidad de guiar el proceso de medición y centrarse selectivamente en un subconjunto del espacio de medición puede resultar en mediciones más eficientes, lo que puede ser deseable en reducir el costo, el tiempo de recolección o mejorar la calidad de los datos recolectados.

3.3. Síntesis

La predicción de fenómenos dinámicos requiere de metodologías y/o sistemas que permitan mejorar los resultados de las simulaciones sin modificar los modelos que definen el comportamiento del fenómeno bajo estudio. En este sentido, los métodos guiados por datos y las diferentes aproximaciones descritas en este capítulo son herramientas de vital importancia para lograr tales objetivos.

Evidentemente, el enfoque más prometedor requerirá de la incorporación de datos en tiempo real que brinden al modelo una mejor representación de la realidad. Esto sin duda tampoco deja de ser una

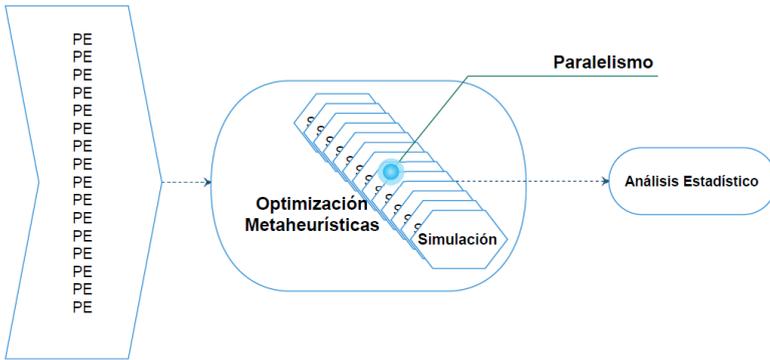
tarea sencilla ya que requiere determinar qué variables medir, con qué frecuencia y cómo tratarlas para que el modelo pueda adaptarse al flujo de esta información. Sin embargo, es importante tener en cuenta que los métodos abordados en la presente tesis consisten en métodos de predicción bajo condiciones extremas de incertidumbre, es decir, sin la incorporación de datos en tiempo real, ni con el conocimiento inicial de los parámetros de entrada.

Métodos Propuestos

El presente capítulo aborda la descripción de los métodos de reducción de incertidumbre propuestos: a) Sistema Estadístico Evolutivo con Modelo de Islas (ESS-IM, por sus siglas en inglés) y b) Sistema Estadístico Híbrido Evolutivo con Modelo de Islas (HESS-IM, por sus siglas en inglés). Se presentan sus características, como así también los detalles de implementación.

4.1. Introducción

Tal como se ha presentado en el capítulo anterior, las metodologías guiadas por datos brindan un enfoque acertado cuando se trata de desarrollar sistemas de predicción de fenómenos bajo condiciones de incertidumbre y alta dinamicidad. En el contexto de este trabajo, la metodología desarrollada y detallada en el presente capítulo consiste en un método que centra sus esfuerzos en minimizar el problema de la incertidumbre en los parámetros de entrada utilizando técnicas de: simulación [15], análisis estadístico [108], computación de alto rendimiento [55] y optimización metaheurística [66]. De forma sintética se puede visualizar la relación de éstos componentes en la Fig. 4.1. Tal como se puede observar, la simulación se utiliza en el centro de nuestro sistema en donde se llevan a cabo múltiples ejecuciones de un simulador de comportamiento de incendios forestales, cada una con diferentes configuraciones de parámetros de entrada (PE). Tales configuraciones de parámetros representan un valor de aptitud en cada paso de predicción, los cuales se optimizan mediante diferentes técnicas metaheurísticas, que se ejecutan de forma concurrente. El paralelismo se vuelve necesario debido al costo computacional de estos procesos y, también, debido a que los resultados deben ser obtenidos en el menor tiempo posible. Finalmente, la estadística permite estimar un patrón de comportamiento del incendio, basado en un análisis de probabilidad considerando la totalidad de las simulaciones llevadas a cabo, esperando que dicho patrón permita realizar una la predicción inmediata futura. Estas técnicas, integradas en los métodos propuestos, permiten incrementar la calidad de predicción frente a las limitaciones ya conocidas de la predicción clásica.



4.1: Relación entre componentes principales del método. **PE**: parámetros de entrada.

El principio de funcionamiento de ESS-IM y HESS-IM se ilustra en la Fig. 4.2. En dicha gráfica se representa al instante actual como t_2 , teniendo como datos de entrada el estado previo del incendio t_1 y la predicción a realizar se refleja en el instante t_3 . De esta manera, el sistema tiene como objetivo determinar aquellas configuraciones de parámetros de entrada del modelo que generan mejor aptitud, y por lo tanto posibiliten describir “con la mayor fidelidad posible” el comportamiento del fuego en un instante pasado. Esperando que dichos parámetros, sigan teniendo validez durante un período de tiempo posterior llamado t_{lp} (tiempo límite de predicción o ventana temporal) donde $t_{lp} \geq 2t_1$, pudiendo ser usados para generar la predicción en un futuro cercano (t_3). Como se puede observar, este enfoque coincide con la arquitectura DDM-MOS, y al igual que se describió en el capítulo anterior, el mismo requiere de un estado adicional de avance del incendio para poder iniciar la cadena de predicción (en este ejemplo puntual la primera predicción sería generada para el instante t_3).

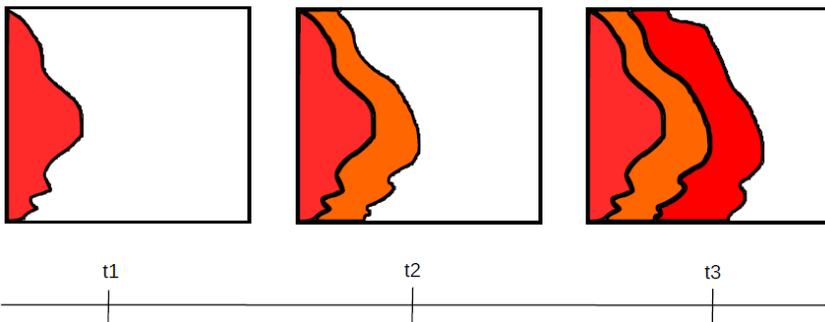


Figura 4.2: Representación del límite de previsibilidad.

En comparación con la predicción clásica se puede observar la Fig. 4.3, donde el enfoque clásico genera la primera predicción para el instante t_{n+1} , mientras el enfoque basado en DDM-MOS la genera para el instante t_{n+2} .

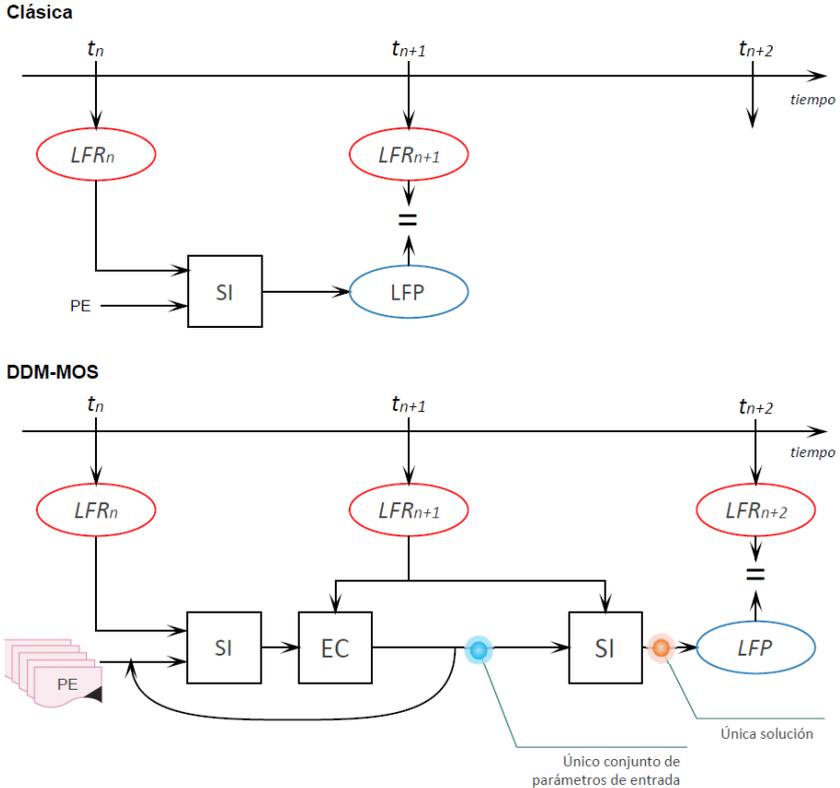


Figura 4.3: Estados de evento real necesarios en enfoque Clásico y DDM-MOS. LFR_x : Línea de Fuego Real en tiempo x , SI: Simulador de propagación de Incendios forestales, EC: Etapa de Calibración, LFP: Línea de Fuego Propuesta por el simulador, PE: Parámetros de Entrada.

4.2. Descripción de los Métodos

ESS-IM y HESS-IM son métodos de reducción de incertidumbre desarrollados para su aplicación en la predicción de desastres naturales. En este trabajo ambos métodos han sido aplicados a la predicción del comportamiento de incendios forestales. Es decir que, según los conceptos mencionados en el Capítulo 1, la *caja negra* del modelo de simulación, en este caso se corresponde con el modelo de

comportamiento de incendios forestales de Rothermel [47]. Puntualmente se trata del simulador de incendios forestales implementado en ANSI C mediante la librería *fireLib* [109]. A continuación, se describen cada uno de los cuatros componentes de los métodos mencionados en la introducción del presente capítulo.

4.2.1. Componente Simulación

Según [110], la modelización de los incendios de superficie ha sido objeto de investigación durante las últimas cinco décadas en los centros de investigación de incendios forestales de todo el mundo.

Producto de estas investigaciones han surgido diferentes modelos basados en diversos factores, tales como: el tipo de combustible, las características del terreno, las condiciones meteorológicas, el calor desprendido en el frente de fuego, la cantidad de combustible consumido, entre otros. Además, los métodos desarrollados incluyen perspectivas teóricas, empíricas y/o semiempíricas, de las cuales no todas se han implementado de forma operativa. En la Tabla 4.1 se presenta un resumen de los más citados en la literatura. En este contexto el modelo de Rothermel es uno de los más utilizados.

El modelo de Rothermel define el comportamiento del fuego asumiendo que éste avanza sobre superficies donde prevalecen combustibles finos y continuos. El modelo considera la energía generada por el fuego, la transmisión de calor al combustible cercano al fuego y también la energía absorbida por el combustible. Se toma en cuenta tanto combustibles vivos como muertos (combustibles que poseen tejido vivo, cuya humedad depende exclusivamente de la humedad atmosférica [111]), y sus contenidos de humedad. También se consideran los efectos del viento y de la pendiente en la transferencia del calor, la carga y grado de compactación del combustible, el tamaño de las partículas que lo componen, entre otros. Todos estos factores determinan la forma y la velocidad en que el fuego se propaga.

A pesar de que el modelo de Rothermel no considera ni modela fuego de copas ni torbellinos de fuego, entre otros, sigue siendo el modelo más utilizado hasta la fecha. El mismo se originó con once modelos de combustible que posteriormente fue ampliado a trece [112]. Cada modelo contiene información sobre las propiedades de la vegetación en una región particular con cierta granularidad. Es decir, la superficie se divide en celdas, donde cada celda tiene propiedades que modelan el tipo de combustible que prevalece. La mayoría de los simuladores de comportamiento del fuego basan sus cálculos en este modelo [63]. Desde que se publicó el documento de Rothermel en 1972, se han

desarrollado varios simuladores de propagación de incendios. A la fecha, la mayor cantidad de implementaciones actuales de software de simulación de incendios forestales usan este modelo. Algunas de las alternativas evaluadas son descriptas brevemente en las subsecciones siguientes.

Tabla 4.1: Clasificación de los modelos de propagación de incendios de superficie (1946-1999) [46], [110].

Autor y fecha	Tipología	Operatividad
Fons (1946) [113]	Teórica	Nula
Emmons (1964) [114]	Teórica	Baja
McArthur (1966) [115]	Empírica	Alta
Van Wagner (1967) [116]	Teórica	Nula
Thomas (1967) [117]	Teórica	Nula
Anderson (1969) [118]	Teórica	Nula
Frandsen (1971) [119]	Semiempírico	Nula
Rothermel (1972) [120]	Semiempírico	Alta
Pagni y Peterson (1973) [121]	Teórica	Nula
Telisin (1974) [122]	Teórica	Nula
Steward (1974) [123]	Teórica	Nula
Konev y Sukhinin (1977)	Teórica	Nula
Cekirge (1978) [124]	Teórica	Nula
Trabaud (1979) [125]	Semiempírico	Nula
Fuji (1980) [126]	Teórica	Alta
Noble et al. (1980) [127]	Empírica	Nula
Grishin et al. (1983) [128]	Teórica	Alta
Hwang y Xie (1984) [129]	Teórica	Nula
Albini (1985, 1986)	Teórica	Nula
De Mestre et al. (1989) [130]	Teórica	Nula
Weber (1989) [131]	Teórica	Nula
Grishin (1997) [132]	Teórica	En desarrollo
Dupuy (1997) [133]	Teórica	En desarrollo
Linn (1997) [134]	Teórica	En desarrollo
Santoni (1998) [135]	Teórica	En desarrollo
Viegas et al. (1998) [136]	Empírica	En desarrollo
Cheney et al. (1998) [137]	Empírica	En desarrollo
Larini et al. (1998) [138]	Teórica	En desarrollo

4.2.1.1. BEHAVE (1986)

El sistema BEHAVE se desarrolló durante varios años de investigación y fue publicado entre 1984 y 1986 [139]. BEHAVE está formado por dos subsistemas: el subsistema de modelado de combustible, Burgan y Rothermel 1984 (programas NEWMDL y TSTMDL) [140], y el subsistema de predicción del comportamiento de incendios forestales, Andrews 1986 (programas FIRE1 y FIRE2) [139]. Estos subsistemas le permitían a BEHAVE ofrecer dos utilidades principales. Por un lado, un módulo para usar los modelos de combustible de Rothermel y desarrollar nuevos modelos de combustible que podían integrarse en el simulador. Por el otro lado, un módulo que permitía ejecutar simulaciones y estimar la predicción de quema de un incendio determinado. La salida del simulador se presentaba en una tabla que representaba los tiempos de llegada de cada celda en la simulación. No ofrecía herramientas de visualización para las simulaciones, por lo cual no era válido para situaciones de tiempo real. Sigue siendo una herramienta útil para el desarrollo del modelo de combustible. BEHAVE, compuesto por cinco programas realizados en FORTRAN que corrían sobre el sistema operativo MS-DOS [141], es considerado el primer simulador basado en el modelo de Rothermel.

4.2.1.2. *fireLib* (1996)

fireLib es una librería escrita en C optimizada para aplicaciones altamente iterativas, la cual encapsula el algoritmo de comportamiento de fuego BEHAVE [51, 52] para predecir la propagación de incendios en dos dimensiones mediante simulación de crecimientos basado en celdas o autómatas celulares (ver Fig. 4.4). Al trabajar con autómatas celulares se simula la tasa de tiempo para la propagación del fuego, indicando el tiempo del frente de fuego para alcanzar cada celda. De esta manera cada celda registra el tiempo en que es alcanzada por el fuego.

Los algoritmos son mucho más simples, pero la geometría de la rejilla puede influenciar en la forma del fuego. Tal como se puede observar en la Fig. 4.5, el fuego comienza en una celda (celda de color rojo) y se propaga a las celdas vecinas según el modelo de comportamiento y los parámetros de entrada así lo determinen. El mismo caso puede observarse en la Fig. 4.5, donde cada celda posee un valor numérico del momento en que dicha celda fue alcanzada por el fuego, donde se utiliza el valor 0 (cero) para las celdas que no se han quemado.

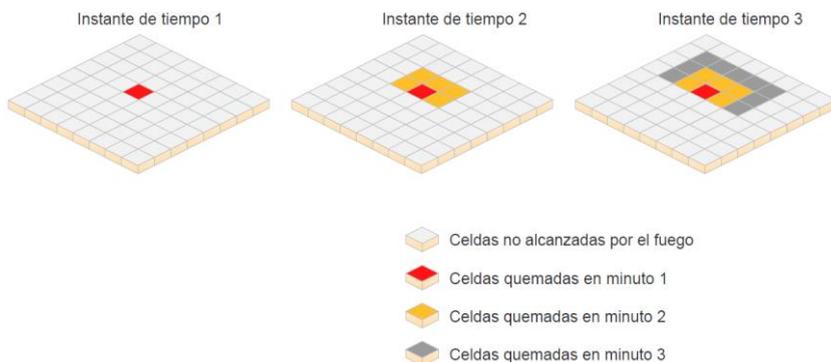


Figura 4.4: Representación gráfica de progresión del fuego mediante un autómata celular.

0	0	0	0	0	0	0	0
0	3	3	3	3	3	0	0
0	3	2	2	2	3	0	0
0	3	2	1	2	3	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Figura 4.5: Representación temporal de progresión del fuego mediante un autómata celular.

El simulador utilizado se denomina fireSim, y corresponde a un simulador de comportamiento de incendios forestales implementado mediante la librería *fireLib*. Éste hace uso de un autómata celular en donde el terreno es dividido en celdas y por medio de la relación entre los vecinos colindantes se evalúa si la celda ha sido alcanzada o no por el fuego, y en caso de serlo registra en qué instante. Este tipo de simulador es paramétrico, determinístico y de eventos discretos (para más información ver [109]).

4.2.1.3. FIRETEC (1997)

FIRETEC es un simulador de incendios que no se basa en las ecuaciones de propagación elípticas desarrolladas por Rothermel. En

lugar de usar datos semi empíricos para calcular sus métodos de propagación, usa modelos teóricos de reacciones químicas y transferencias de calor para descifrar dónde se va a propagar el fuego [142].

El modelo considera distintas escalas: los combustibles más pequeños (hoja o rama), elementos de una escala mediana (rama o árbol) y la macroescala. Los promedios de las micro y mesoescala se insertan cuidadosamente en la macroescala a fin de poder resolver un cálculo de diferencia finita. Por otro lado, si bien se prevé dos tipos de quema (local y no local), FIRETEC representa localmente las interacciones combinadas de todos los tipos de quemado, ya que utiliza una velocidad de reacción efectiva única. Es decir, todo el calor se genera en el lugar donde se encuentra el combustible, y posteriormente es transportado. FIRETEC permite representar una gran cantidad de entornos, incluidas regiones de transición, como aquellas con vegetación y terreno heterogéneos. Simula la aceleración y desaceleración de los frentes de fuego al variar la pendiente, fuego de copa en presencia de fuertes vientos, entre otros. Actualmente, FIRETEC está combinado con un modelo computacional de dinámica de fluidos, HIGRAD [143], que representa el flujo del aire y sus ajustes al terreno, vegetación y al fuego mismo. HIGRAD/FIRETEC permite simular los procesos dinámicos que ocurren dentro de un incendio y la forma en que estos procesos se alimentan y se alteran entre sí.

4.2.1.4. FARSITE (1998)

FARSITE (Fire Area Simulator) es un software diseñado para simular el crecimiento/avance de incendios forestales utilizando modelos existentes de comportamiento de BEHAVE y del Sistema Canadiense de Predicción de Comportamiento de Incendios Forestales (Forestry Canada Fire Danger Group 1992) [144].

Produce perímetros de fuego vectoriales (polígonos) a intervalos de tiempo específicos. Los vértices de los polígonos contienen información sobre la velocidad y la intensidad de la propagación del fuego, que se interpolan para producir imágenes tipo ráster (ASCII). Debido a que FARSITE puede generar mapas espaciales, es útil para producir análisis detallados del comportamiento y los efectos del fuego usando sistemas de información geográfica (SIG). Esto es particularmente útil para explorar las conexiones entre diferentes modelos de comportamiento del fuego, revelar las implicaciones de sus

suposiciones para impulsar el crecimiento e identificar componentes faltantes entre los distintos modelos.

FARSITE funciona como un simulador de incendios forestales a gran escala, incorpora modelos existentes de fuego de superficie, fuego de copa, entre otros, permitiendo salidas gráficas del comportamiento del fuego (longitud de la llama, intensidad de la línea del frente, velocidad de propagación y calor por unidad de área, entre otros).

Los cálculos unidimensionales para cada modelo se aplican directamente a los vértices del frente del fuego. Sin embargo, esto supone que el comportamiento del fuego en cada vértice es independiente de cualquier interacción espacial o temporal con su entorno, por lo que FARSITE no permite simular de manera confiable incendios masivos. Por otro lado, FARSITE supone que las formas de los incendios son elípticas en condiciones uniformes e independientes del tipo de combustible. Además, considera que al variar las condiciones de combustión la forma elíptica es perada es instantáneamente alcanzada. Otro supuesto que asume este modelo es que la aceleración del fuego depende del tipo de combustible, pero no de su comportamiento. Actualmente FARSITE ya no posee soporte y es incluido en la versión 6 de FlamMap [145].

4.2.1.5. BehavePlus (2001)

BehavePlus fue el sucesor del antiguo BEHAVE. Desde la versión 1 hasta la 6- beta (última versión publicada hasta la fecha) ha sido diseñado para ejecutarse sobre diferentes versiones de Microsoft Windows. A diferencia de BEHAVE, las versiones de BehavePlus están compuestas por un único paquete de software desde donde se puede acceder a todas las funcionalidades.

En la primera versión de BehavePlus [146], ésta se definía como una aplicación para predecir el comportamiento de los incendios forestales para fines de manejo. Diseñado para ser utilizado por operadores que están familiarizados con combustibles, el clima, la topografía, las situaciones de incendios forestales y la terminología asociada. BehavePlus utiliza los mismos modelos matemáticos que el original BEHAVE. Los modelos en BehavePlus están agrupados en 9 módulos: probabilidad de ignición, comportamiento del fuego de superficie, comportamiento del fuego de corona, tamaño de la zona de seguridad, tamaño y forma del fuego de la fuente puntual, contención del fuego debido a la acción de supresión, distancia máxima de detección, altura

de la copa quemada y mortalidad de árboles. Las versiones futuras de BehavePlus incorporarán modelos de fuego adicionales [147].

BehavePlus es parte de un conjunto de sistemas de comportamiento de incendios que incluye FlamMap [145], FARSITE [50], entre otros.

4.2.1.6. WRF-Fire (2009)

WRF-Fire [148] es un módulo de física en WRF (Weather Research and Forecasting Model) [149], se encuentra disponible desde 2010, permite simular el crecimiento y la propagación de incendios forestales. Los procesos físicos en WRF-Fire son un subconjunto de los algoritmos implementados por [150, 151], pero en WRF-Fire se limita actualmente a incendios de superficie, es decir, un incendio que se propaga a través de combustibles que se encuentran sobre la superficie del terreno, pasto o arbustos [152], sin considerar otros tipos de incendios.

WRF-Fire combina el modelo meteorológico WRF con SFIRE [153], un modelo de comportamiento de fuego superficial basado en fórmulas semi-empíricas, que calculan la tasa de propagación de la línea de fuego según las propiedades del combustible, las velocidades del viento de WRF y la pendiente del terreno. La simulación de la propagación del fuego está basada en el método de ajuste de nivel (level-set-method, LSM). Esto permite considerar tanto la influencia del clima (vientos, humedad, entre otros) en el comportamiento de los incendios forestales, como la del fuego en la atmósfera, a través de los flujos de calor y vapor de la combustión, y la evaporación de la humedad de los combustibles [154].

4.2.2. Componente Estadístico

La estadística es una ciencia formal que se ocupa de la recopilación, presentación, análisis y uso de datos para tomar decisiones, resolver problemas y diseñar productos y procesos [108]. Por lo tanto, las estadísticas permiten buscar correlaciones y dependencias entre las variables que afectan un fenómeno particular físico o natural y el fenómeno en sí. En el desafío tratado en nuestro trabajo, la predicción del comportamiento de incendios forestales necesita de un conjunto de parámetros (también considerados como factores o variables) que determinan y afectan el comportamiento del fenómeno, por ejemplo, la velocidad y dirección del viento, la humedad en la vegetación, la pendiente, entre otros. Estos factores alimentan el modelo de predicción, cuyo objetivo es determinar el comportamiento del fuego

de acuerdo con los valores de entrada de cada variable. Debido a que cada parámetro puede tener valores diferentes (independientemente del resto), es necesario evaluar los efectos de cada parámetro junto con los valores de los otros. Por lo tanto, una forma eficiente de resolver este problema, desde el punto de vista estadístico, es el diseño de un experimento factorial.

Según [108], un experimento factorial es uno en el que los factores varían entre sí. En otras palabras, es un tipo de diseño experimental en el que cada valor de un factor se prueba en combinación con cada valor de otro factor (en general, en un experimento factorial, se prueban todas las combinaciones posibles de valores [108]). En nuestro contexto, cada combinación posible de valores de parámetros se denomina escenario (este concepto se tratará en profundidad en las siguientes secciones). Por lo tanto, si queremos determinar qué porción de una superficie de terreno (dividida en m celdas) será alcanzada o no por el fuego, se pueden definir n diferentes escenarios y calcular la probabilidad de ignición (P_{ign}) para cada celda de la siguiente manera [63]:

$$P_{ign}(C) = \frac{n_C}{n} \quad (4.1)$$

donde $P_{ign}(C)$ es la probabilidad de ignición para la celda C y n_C es el número de escenarios donde se quemó la celda C .

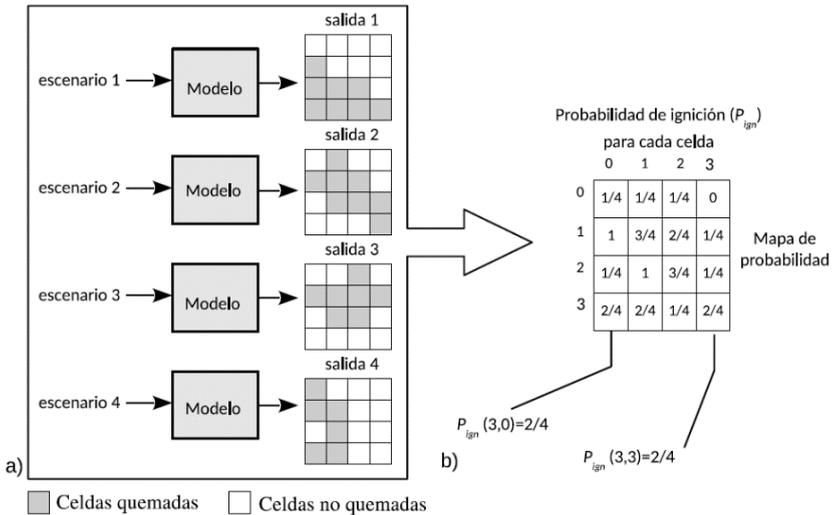


Figura 4.6: a) Ejemplo de cálculo para construir el mapa de probabilidad, b) mapa de probabilidad.

Para entender mejor el concepto, puede considerarse el ejemplo de la Fig. 4.6.a. En tal ejemplo, el número de escenarios es $n = 4$, por lo tanto, se obtienen 4 salidas intermedias (*salida 1, 2, 3 y 4*) y considerando la ecuación (4.1) se puede calcular la *probabilidad de ignición* que tiene en cuenta estos cuatro escenarios. De esta manera, es posible generar el mapa de probabilidad de la Fig. 4.6.b. Este mapa, como puede verse, es una matriz con un valor asociado a cada celda que representa la probabilidad de ignición de cada celda. El conjunto de celdas, cuyo valor P_{ign} es mayor o igual que un determinado valor particular P_K , donde $0 \leq P_K \leq 1$, constituye lo que llamamos mapa de probabilidad con probabilidad P_K . La Fig. 4.7 muestra la representación gráfica de este concepto.

4.2.2.1. Patrón de Comportamiento y Evaluación

Tal como hemos mencionado anteriormente, HESS-IM es un método general de reducción de incertidumbre aplicado a la predicción del comportamiento de incendios forestales. Sin embargo, es necesario aclarar que cuando se hace mención a “reducir la incertidumbre en los parámetros de entrada” no hacemos referencia a obtener los valores reales de las variables que alimentan al modelo en un instante determinado. En cambio, se busca obtener n configuraciones de parámetros de entrada que, en su conjunto, permitan describir el comportamiento del fenómeno en un instante de tiempo conocido (pasado cercano), esperando que las condiciones varíen lo menos posible, permitiendo así efectuar la predicción con mayor exactitud y calidad en un instante futuro cercano. Este concepto se denomina, en el contexto del método propuesto, *patrón de comportamiento del modelo*.

Hallar un patrón de comportamiento del modelo es el objetivo central de ESS IM y HESS-IM, el cual, como se acaba de mencionar, será independiente a los valores particulares de los parámetros de entrada del modelo. Ya que, por más que contásemos con un simulador perfecto, igualmente tendríamos errores de predicción por la falta de precisión de los parámetros de entrada, dado que existen ciertos parámetros que no pueden medirse directamente y, por lo tanto, deben estimarse a partir de medidas indirectas, incluso sin considerar que algunos también varían con mucha frecuencia durante el mismo incendio.

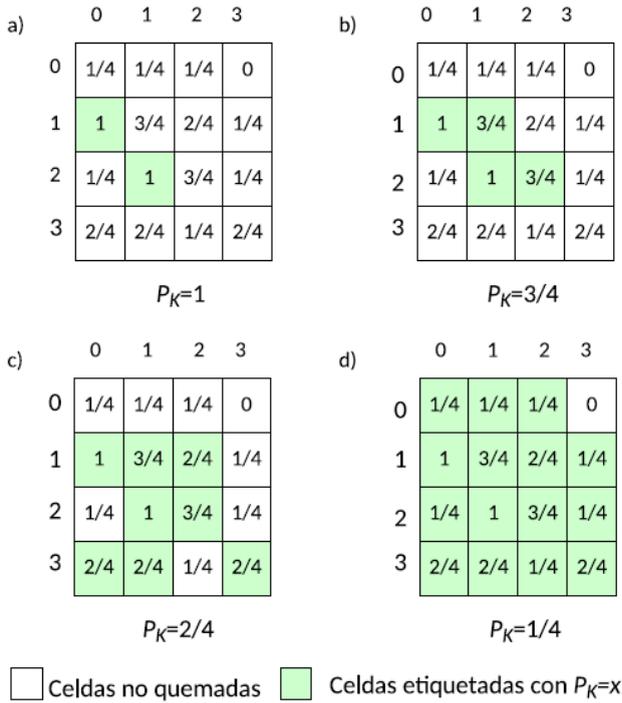


Figura 4.7: Cuatro ejemplos de mapas de probabilidad con probabilidades: a) $P_K = 1$, b) $P_K = 3/4$, c) $P_K = 2/4$ y d) $P_K = 1/4$.

El patrón de comportamiento se calcula comparando el mapa de probabilidad (generado a partir de las simulaciones) con el mapa real en el instante de tiempo t_i . Dicha comparación se realiza mediante una función basada en el índice de Jaccard [155] la cual se describe en la ecuación 4.2 [63].

$$Fitness = \frac{|A \cap B|}{|A \cup B|} \quad (4.2)$$

En donde A representa el conjunto de celdas en el mapa real sin el subconjunto de celdas quemadas antes de comenzar la simulación, y B representa el conjunto de celdas en el mapa de simulación sin el subconjunto de celdas quemadas antes de comenzar la simulación. La representación gráfica de este proceso puede observarse en la Fig. 4.8. En la misma, el patrón de comportamiento se representa con la abreviatura K_{ign} (o valor clave de ignición) para un intervalo de tiempo específico. En otras palabras, el valor de K_{ign} consiste en un valor

particular de P_{ign} , cuyo mapa de probabilidad asociado proporciona la mejor coincidencia con la propagación del fuego real.

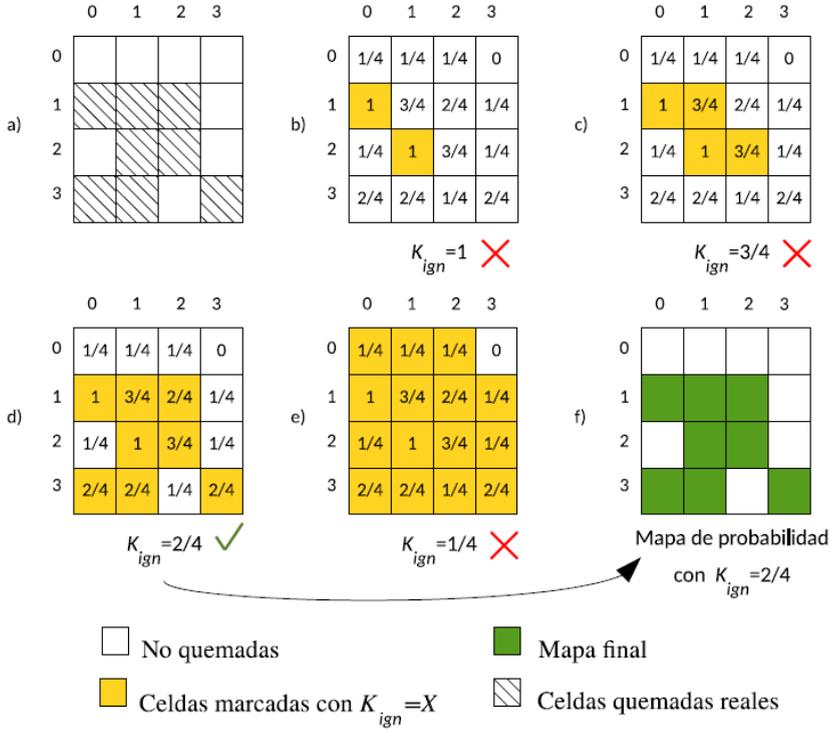


Figura 4.8: a) representación del mapa real quemado, b) a f) ejemplo de una búsqueda del valor clave de ignición K_{ign} .

4.2.3. Componente de Optimización Metaheurística

En la sección 2 se presentaron los beneficios de utilizar metaheurísticas en la resolución de problemas complejos de alta dinamicidad. Sin embargo, el potencial de éstas puede verse incrementado si diferentes metaheurísticas son utilizadas de forma colaborativa, tratando de maximizar las bondades de cada una y reduciendo sus puntos débiles.

4.2.3.1. Metaheurística híbrida evolutiva

Un esquema general de una metaheurística híbrida evolutiva bajo un esquema de colaboración basada en migración entre islas se presenta en el Algoritmo 6. Tal como se puede observar, el proceso comienza con

la generación de p conjuntos de n soluciones candidatas, las cuales suelen ser creadas mediante un proceso estocástico. Una vez que los conjuntos de soluciones han sido generados, se evalúa la aptitud inicial de las soluciones, aplicando la función de *fitness* a cada solución. Luego, comienza el ciclo evolutivo instanciando diferentes metaheurísticas. Cada metaheurística itera mientras no se cumpla la condición de finalización, es decir, hasta que se alcance un cierto umbral de aptitud o una máxima cantidad de iteraciones.

Algoritmo 6: Esquema general de una metaheurística híbrida evolutiva.

```

1: inicio
2:  $g=0$  /*número de iteración*/
3: generar  $p$  conjuntos con  $n$  soluciones candidatas
4: calcular la función de evaluación de cada solución
5: while no se cumpla condición de finalización do
6:   /* Producir nueva iteración */
7:   operadores ejecución de operadores que modifican las soluciones (según metaheurística implementada)
8:   calcular la función de evaluación de cada solución
9:   if  $g \% \text{frecuencia} == 0$  then
10:    /* realizar migración de soluciones */
11:    enviar  $m$  mejores soluciones a conjunto vecino
12:    recibir  $m$  soluciones de conjunto vecino
13:    reemplazar  $m$  soluciones en conjunto actual
14:    evaluar cada solución con la función objetivo
15:     $g++$ 
16: fin

```

Posteriormente se realiza el proceso que permite modificar a las soluciones con el objetivo de que las mismas se acerquen más a una solución óptima del problema (alteración de las soluciones candidatas). Ese proceso implementa las funciones que cada metaheurística necesita para modificar las soluciones candidatas. A continuación, se realiza el proceso de migración, el cual se efectúa cada cierta cantidad de iteraciones (frecuencia de migración), donde m individuos son seleccionados y enviados por cada isla a sus vecinos (diferentes metaheurísticas). La migración, selección y el reemplazo (dentro del proceso de migración) suelen ser elitistas (es decir, se seleccionan los mejores m individuos de la isla origen, que reemplazan a los peores m individuos en la isla destino) aunque también son válidas otras estrategias. A medida que el conjunto de soluciones evoluciona, las mismas deben aproximarse a una solución óptima para el problema tratado.

4.2.3.2. Metaheurística híbrida evolutiva paralela

Las metaheurísticas híbridas evolutivas pueden implementarse de forma secuencial o paralela. Cuando las metaheurísticas son utilizadas para resolver problemas complejos y/o de grandes dimensiones, usualmente éstas son paralelizadas, dividiendo la carga de trabajo en N unidades de cómputo. Las metaheurísticas pueden paralelizarse siguiendo diferentes criterios y estrategias, algunas de ellas sólo mejoran el tiempo de respuesta del algoritmo sin influir en los resultados obtenidos [97].

HESS-IM utiliza un esquema de paralelización basado en islas con doble jerarquía master-worker (ver sección 2.4.1). Un esquema general puede observarse en la Fig. 4.9. Como puede apreciarse, en el nivel superior se encuentra el modelo master worker L1, donde un proceso *master global* o *monitor* coordina el funcionamiento de la totalidad de las islas (workers L1). El master local de cada isla cumple la función de worker respecto del master global. Posteriormente, en el segundo nivel de jerarquía master-worker, cada master de isla controla a un conjunto de workers dentro de cada isla (workers L2). Es importante remarcar que cada isla inicializa una población diferente, las cuales son perturbadas mediante diferentes metaheurísticas. La colaboración entre cada metaheurística se lleva a cabo por el operador de migración, el cual intercambia individuos entre las diferentes islas y/o metaheurísticas cada ciertos eventos (es decir, cantidad de generaciones realizadas, valor de aptitud de la población, detección de estancamientos en óptimos locales, etc.).

HESS-IM utiliza de forma colaborativa tres metaheurísticas evolutivas presentadas en la sección 2: Algoritmos Evolutivos [86], Evolución Diferencial [87] y Optimización por Cúmulo de Partículas [88]. La elección de estas metaheurísticas ha sido sustentada en función de experimentaciones previas realizadas, y en base a la literatura, es importante tener en cuenta que el diseño de HESS-IM permite ser fácilmente adaptado a otras metaheurísticas evolutivas paralelas.

4.2.4. Componente de Computación de Alto Rendimiento

La predicción 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 sentido, la computación de alto rendimiento ha permitido abordar problemas cada vez más grandes y complejos, beneficiado por el avance tecnológico de los últimos años.

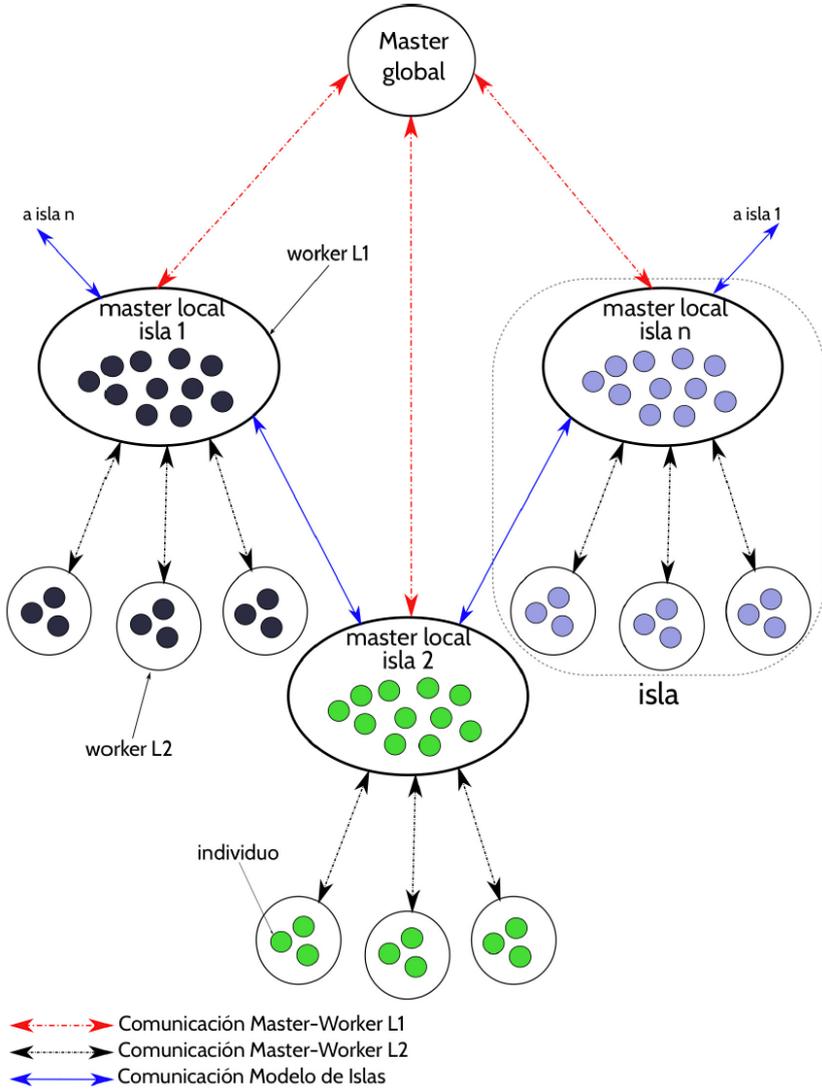


Figura 4.9: Esquema de paralelización jerárquico basado en islas con doble modelo master-worker.

Tal como se acaba de mencionar en la sección 4.2.3.2, HESS-IM utiliza una arquitectura de doble jerarquía master/worker la cual en su implementación original incorporó únicamente paralelismo de memoria

distribuida bajo La librería de programación denominada Interfaz de Paso de Mensajes (MPI, Message Passing Interface) [156]. En la Fig. 4.9 se puede observar la existencia de tres tipos de procesos: monitor o master global, worker de nivel 1 o master de isla, y worker de isla. Es decir, en el nivel superior (L1) el proceso monitor coordina a sus respectivos procesos workers. En el nivel inferior de jerarquía (L2) cada uno de estos procesos opera como master de isla. Finalmente, cada isla contiene su conjunto de workers de isla. La comunicación entre los procesos se ve reflejada en las conexiones entre cada entidad.

Debido a que las metaheurísticas involucradas en el proceso de optimización necesitan someter cada solución candidata al simulador de incendios, resulta indispensable reducir la cantidad de evaluaciones de la función objetivo, o bien el tiempo de procesamiento asociado a este proceso. Persiguiendo este objetivo, se implementó HESS-IM bajo un esquema de paralelismo heterogéneo de memoria distribuida/compartida bajo MPI/CUDA (Compute Unified Device Architecture) [157]. De esta manera, y dado que el componente de simulación de HESS-IM, *fireSim*, se encuentra basado en la librería *fireLib*, el cual originalmente fue implementado para CPU, y dado que la simulación es la operación de mayor costo computacional, se incorporó una implementación de *fireLib* para GPU basada en el trabajo de [158]. En este sentido las herramientas utilizadas para la paralelización de los métodos desarrollados se presentan en las siguientes subsecciones.

4.2.4.1. MPI

La Interfaz de Paso de Mensajes (MPI, por sus siglas en inglés) puede decirse que es la librería de programación más utilizada para programar en arquitecturas de memoria distribuida. MPI es un estándar para facilitar la escritura de programas mediante el paso de mensajes, el estándar MPI define la sintaxis y la semántica de las rutinas útiles de la biblioteca. MPI es la librería de programación dominante en informática de alto rendimiento. Al utilizar MPI, múltiples procesos MPI computan en paralelo y se comunican llamando a las rutinas de la biblioteca MPI. En la inicialización del programa, se crea un conjunto fijo de procesos. Por lo general, el rendimiento óptimo se logra cuando cada proceso MPI se mapea en un núcleo separado. Para facilitar la coordinación entre procesos, la interfaz MPI proporciona funcionalidad para comunicación, sincronización y topología virtual. En la

arquitectura presentada en la Fig. 4.9 MPI es utilizado para comunicar los procesos monitor, master y workers.

4.2.4.2. OpenMP

Dentro del esquema de paralelismo de memoria compartida, los procesos se comunican a través de la escritura y lectura la memoria compartida. Bajo esta clasificación podemos encontrar dos implementaciones más frecuentes, OpenMP y POSIX Threads. OpenMP [159] es una interfaz de programación de aplicaciones (API) para la programación multiproceso de memoria compartida en múltiples plataformas. Provee un conjunto de directivas, rutinas y variables de entorno. A partir del conjunto de directivas se puede crear threads, realizar operaciones de sincronización y distribuir la carga de trabajo entre threads. OpenMP impone una estructura jerárquica entre los threads, ya que sigue el modelo fork/join. En este modelo la ejecución del programa comienza como un único hilo *master*, que se ejecuta secuencialmente hasta encontrar una región paralela (directiva *parallel*). En ese momento, el master crea un conjunto de *threads* (*fork*). A partir de ese instante, cada *thread* ejecuta las sentencias encerradas en la región paralela en forma concurrente. Cuando todos los *threads* finalizan, se sincronizan y terminan (*join*), quedando solo el *thread master*, el cual continúa la ejecución.

Debido a la creciente masificación de los procesadores multinúcleo, la tendencia actual propone explotar al máximo los recursos computacionales siendo necesario hacer uso de implementaciones híbridas basadas en MPI y OpenMP. Las diferentes combinaciones de implementación pueden clasificarse en tres tipos de implementaciones: a) MPI-puro: donde el paralelismo está expresado a partir del modelo de paso de mensajes entre los diferentes procesos involucrados, b) OpenMP puras: en donde el nivel de paralelismo se realiza dentro de cada nodo de cómputo. En este caso, no se realiza paso de mensaje entre los hilos involucrados ya que los mismos comparten una memoria común. La creación, destrucción y sincronización de estos *threads* es realizada por la librería en el inicio y fin de las secciones paralelas o son especificadas explícitamente por el programador, y c) MPI + OpenMP (Híbrido): estas implementaciones son una alternativa para aprovechar el paralelismo de ambos niveles. De esta forma, un proceso, estaría compuesto por: 1) una región serie, 2) una región de comunicación entre procesos y 3) una región paralela definida por las secciones de OpenMP. OpenMP ha sido utilizado en evaluaciones

experimentales, y no se encuentra en la implementación final de ESS-IM ni HESS-IM, sin embargo, ha sido utilizado para paralelizar tareas dentro de cada nodo representado en la Fig. 4.9.

4.2.4.3. CUDA

Hoy en día las unidades de procesamiento gráfico (GPU) forman parte de los aceleradores de aplicaciones más populares, y están ampliamente integradas en clústeres HPC. Por su parte, CUDA es el estándar de facto para programar las GPU de NVIDIA. En este contexto, el programador tiene que escribir los llamados *kernels* CUDA (código fuente para GPUs) que se ejecutan simultáneamente por muchos hilos en la GPU. Un programa CUDA incluye código de *host* (código para CPU) y *device* (dispositivo, código para GPU). El código de *host* es puro código ANSI C y el código del dispositivo (GPU) es la extensión de ANSI C que proporciona palabras clave para etiquetar funciones paralelas o núcleos de datos. Ambos códigos se compilan por separado, siendo el primero por el compilador estándar de C y el compilador NVIDIA o un compilador de código abierto para el código del dispositivo. El código del host se ejecuta en el host por el CPU y el código del dispositivo realiza lo propio en la GPU. Es importante tener en cuenta que debido a que CUDA es propiedad de NVIDIA, el rendimiento de las aplicaciones realizadas con CUDA está optimizado para las GPU de NVIDIA. Es importante mencionar que ESS-IM no fue implementado de forma heterogénea (CPU-GPU). HESS-IM en su versión heterogénea utiliza MPI para comunicar los procesos entre nodos (ver Fig. 4.9) y CUDA en algunos nodos para instanciar la etapa de simulación de incendios (SI).

4.3. Integración

La evolución de los métodos contempla varias fases desarrolladas cronológicamente tal como se muestra en la Fig. 4.10 y la Tabla 4.2, y son descritas a continuación.

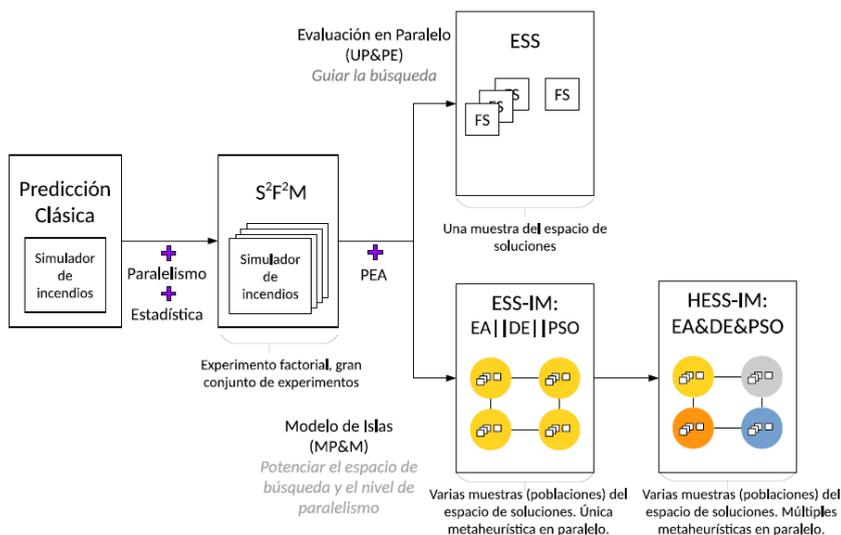


Figura 4.10: Evolución de los métodos a partir de la predicción clásica.

Tabla 4.2: Arquitectura, nivel de paralelismo por métodos desarrollados. EA: Algoritmos Evolutivos, PSO: Optimización por Cúmulo de Partículas, DE: Evolución Diferencial, UP&EP: Única Población y Evaluación en Paralelo, MP&M: Múltiples Poblaciones y Migración.

Versión	Arquitectura			Metaheurística			Paralelismo	
	MPI	OpenMP	CUDA	EA	PSO	DE	UP&EP	MP&M
ESS	x			x			x	
ESS-IM	x	x		x				x
HESS-IM	x		x	x	x	x		x

4.3.1. ESS

Como puede observarse en la Fig. 4.10, el primer método de la taxonomía DDM MOS consistió en la implementación de un método de reducción de incertidumbre basado en análisis estadístico y cómputo distribuido. Este método, conocido como S²F²M - Sistema Estadístico para la Gestión de Incendios Forestales (Statistical System for Forest Fire Management) [100], [63], consiste en la obtención de un patrón de comportamiento del modelo (sección 4.2.2.1), independientemente de los valores particulares de los parámetros de entrada del mismo. Es decir, todos los escenarios posibles (diferentes combinaciones de parámetros de entrada) son generados mediante un experimento factorial (sección 4.2.2). El sistema evalúa (simula) para cada escenario el comportamiento del modelo. Los resultados obtenidos son

combinados para determinar la tendencia del comportamiento del modelo, ajustándolo con la observación del mismo en un instante previo, ya que dicho patrón es utilizado para realizar la predicción del próximo paso.

Debido a la naturaleza de los experimentos factoriales y a la cantidad de parámetros y rangos válidos para cada uno de ellos, este método requiere de un gran número de operaciones, siendo muy demandante de tiempo de ejecución, por lo que es implementado bajo un esquema paralelo bajo MPI. Realizar el cálculo del patrón de comportamiento en base a todas las posibles combinaciones de parámetros genera que se incorporen, de forma inevitable, escenarios que no aportan un valor significativo al resultado, ya sea por ser redundantes o por tratarse de casos negativos que degradan los resultados ofrecidos por el método. Con el propósito de mitigar tal comportamiento, esta metodología fue enriquecida con una nueva etapa de optimización de los escenarios, con el propósito que sólo fuesen considerados (para el cálculo del patrón de comportamiento) aquellos escenarios previamente *optimizados*.

Este renovado enfoque da origen al método denominado Sistema Estadístico Evolutivo (ESS, Evolutionary Statistical System) [101]. Es importante remarcar que ESS opera con una única población. Luego que toda la población ha sido evaluada, el proceso master evoluciona la población, incluye los resultados parciales y realiza la predicción para cada instante de tiempo actual. El proceso master incorpora una etapa de optimización la que implementa un algoritmo evolutivo que permite buscar escenarios (en este contexto llamados individuos) que se adapten de la mejor manera al problema, para posteriormente introducirlos al método estadístico. Es importante aclarar que, de esta manera se considera sólo una muestra de dicho experimento factorial para constituir una población de individuos, la cual es explotada por el componente evolutivo. La incorporación del componente evolutivo permite guiar la búsqueda de las soluciones y reducir el número de casos en estudio en la etapa estadística. Dicho de otra manera, el método genera predicciones basadas en el análisis estadístico de las soluciones de una población de escenarios.

Para comprender mejor el funcionamiento de esta metodología se presenta en la Fig. 4.11 un esquema general de su funcionamiento.

Dicho esquema representa la arquitectura a nivel de paralelismo como así también de funcionamiento interno del método.

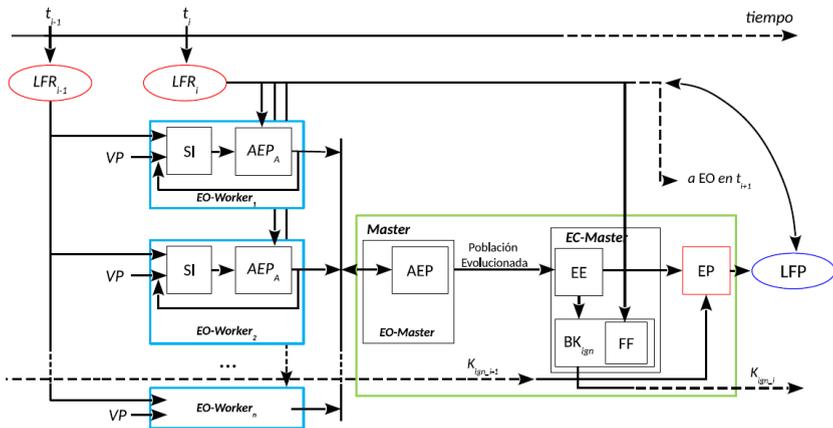


Figura 4.11: ESS: Diagrama esquemático del Sistema Estadístico Evolutivo. LFR_x : estado real en tiempo x , SI: simulador de propagación de incendios forestales, EO: Etapa de Optimización, EC: Etapa de Calibración, AEP: Algoritmo Evolutivo Paralelo, EE: Etapa Estadística, BK: Búsqueda de K_{ign} , K_{ign} : valor clave de ignición; FF: Función de Fitness, EP: Etapa de Predicción, LFP: Línea de Fuego Propuesta por el modelo, VP: Vector de Parámetros de entrada.

A grandes rasgos se puede observar que el sistema se encuentra dividido en dos grandes etapas: por un lado la Etapa de Optimización (EO), la cual es llevada a cabo tanto en el proceso master² (EO-Master), como así también en cada uno de los procesos workers (EO-Worker n). La primera de ellas se encarga de coordinar la evolución de la población de individuos mediante el algoritmo evolutivo paralelo (cuadro AEP), iterando hasta que la población alcance cierto nivel de aptitud. La segunda realiza la evaluación de la aptitud de los individuos en paralelo (cuadro AEP_A). Por el otro lado, la Etapa de Calibración (EC), la cual consta de una Etapa Estadística (EE), es alimentada con la población evolucionada proporcionada por la etapa de optimización en su conjunto (EO-Worker n y EO-Master). La salida de la EE consiste en un mapa de probabilidad (ver Fig. 4.8), que es utilizado para dos propósitos. Por un lado, en la etapa de búsqueda del valor clave de ignición K_{ign} (BK), para calcular el patrón de comportamiento del incendio K_{ign} , el cual será utilizado para realizar la predicción en el próximo instante de tiempo. Y por el otro lado, la salida de BK en tiempo t_i se combina con el mapa de probabilidad proporcionado por

² ESS está basado en un modelo master-worker [95]. Procesos dentro de la ejecución: 1 proceso máster, n procesos workers.

EE en tiempo t_{i+1} para generar la predicción en t_{i+1} en la etapa de predicción (**EP**). La evaluación de la aptitud de los resultados se lleva a cabo en la etapa que implementa la función de fitness (**FF**). Es importante notar que la calibración para el instante de tiempo t_{i+1} se solapa en t_i con los cálculos de predicción.

Cada método posee una serie de parámetros de configuración que condicionan fuertemente su desempeño, tanto en términos de calidad de predicción, como así también en relación a la cantidad de tiempo de procesamiento. En el caso de ESS los parámetros que deben inicializarse antes de ejecutar una instancia del sistema se presentan en la Tabla 4.3. Como se puede observar, los siete primeros corresponden a características específicas que hacen al funcionamiento interno de la metaheurística, y el octavo es un parámetro vinculado a la cantidad de recursos computacionales que relaciona la cantidad de procesos a utilizar en la ejecución (unidades de procesamiento, memoria y red).

Tabla 4.3: ESS: detalle de parámetros de configuración inicial.

Parámetro	Descripción	Rango
Umbral de fitness	valor de aptitud alcanzado por al menos un individuo de la población	[0,1]
Iteraciones	cantidad máxima de iteraciones a llevarse a cabo en una ejecución	Entero positivo
Individuos	número de individuos que componen la población de soluciones candidatas	Entero positivo
Workers	número de procesos que se instanciarán y participarán como procesos workers en toda la ejecución	Entero positivo (dependiente de la capacidad del cluster)
Criadores	porcentaje de la población que será utilizado para generar nuevos individuos	[0,1]
Mín. cruzamiento	mínima probabilidad de cruzamiento	[0,1]
Máx. cruzamiento	máxima probabilidad de cruzamiento	[0,1]
Mutación	valor que determina la probabilidad de mutación	[0,1]

4.3.2. ESS-IM

Como se describió en la sección 2.4.1, los algoritmos evolutivos paralelos se clasifican en tres grupos principales: Única Población y Evaluación Paralela (UP&PE), Única Población y Vecindarios Solapados (UP&ON) y Múltiples Poblaciones y Migración (MP&M). El esquema de paralelismo implementado en ESS se corresponde con el primero de éstos, en ESS-IM (Sistema Estadístico Evolutivo con

Modelo de Islas) se incrementó el nivel de paralelismo del AEP pasando a un esquema de Múltiples Poblaciones y Migración (también conocido como Modelo de Islas). Este esquema permite aumentar la diversidad en el AEP al operar con múltiples poblaciones gestionadas en diferentes islas. ESS-IM, al igual que ESS, está basado en el paradigma master-worker. Sin embargo, el primero utiliza dicho paradigma en dos niveles: master-worker L1 y master-worker L2, como puede observarse en la Fig. 4.13. En el *nivel 1*, el proceso monitor se encarga de inicializar cada una de las islas al comienzo de la ejecución, luego recibe los resultados tras el procesamiento en cada una de las islas, y por último lleva a cabo la predicción del comportamiento del incendio para el paso de simulación actual.

Es importante destacar la diferencia entre un paso de simulación y uno de predicción: para ello observemos la Fig. 4.12.

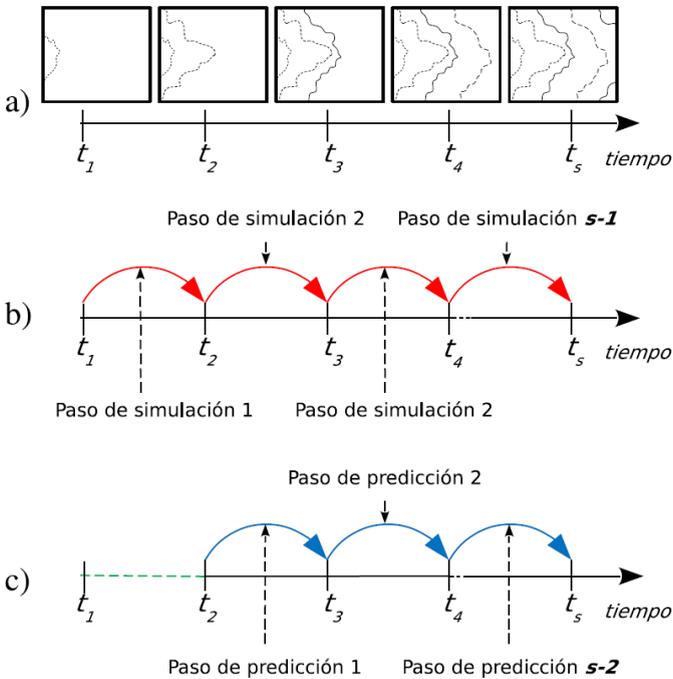


Figura 4.12: a) Representación del avance del incendio, b) pasos de simulación y c) predicción.

Aquí se puede ver que para un incendio de duración determinada, existen t_s muestras representativas del avance del fuego, por lo tanto $s-1$ pasos de simulación y $s-2$ pasos de predicción. Esto es debido a

que los métodos descritos precisan de una etapa de calibración de parámetros de entrada, por lo que el primer paso de simulación es utilizado para realizar la calibración inicial de parámetros de entrada. Luego de ésta, el sistema se encuentra en condiciones para efectuar la primera predicción correspondiente al segundo paso de simulación, que equivale al primer paso de predicción (es decir, entre t_2 y t_3). Por lo tanto, para todo caso de prueba se obtienen $s-2$ valores de calidad de predicción, uno para cada paso de predicción.

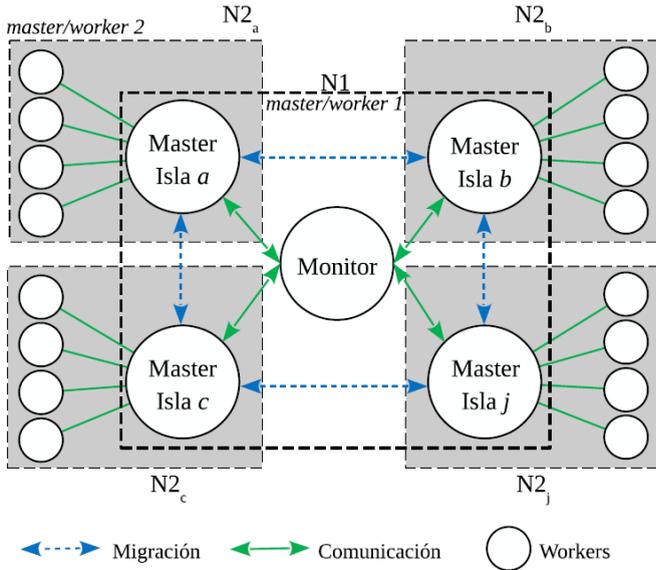


Figura 4.13: Modelo de comunicación y representación del tratamiento de la población (master-worker L1: nivel 1 de master-worker, master-worker L2: nivel 2 de master-worker).

El *nivel 2*, está compuesto por diferentes instancias de este esquema (master-worker), donde cada isla contiene un proceso master coordinador y su respectivo grupo de procesos workers. El master de cada isla distribuye los individuos a los workers, un individuo en ESS-IM representa un conjunto particular de parámetros de entrada que alimentan al modelo, del mismo modo que en ESS.

4.3.2.1. Individuos

Cada parámetro de entrada del simulador fireSim posee un rango de valores válidos. En el contexto de las metaheurísticas, éstos rangos son

la base para crear valores aleatorios en dichos intervalos. Estos valores se utilizan para definir los diferentes componentes de cada individuo en la población (ver Fig. 4.14).

Estos parámetros son: modelo de combustible (Model), velocidad del viento (WinSpd), dirección del viento (WinDir), pendiente (Slope), aspecto u orientación de la pendiente (Aspect), contenido de humedad en el combustible muerto pasada una hora (M1), 10 horas (M10) y 100 horas (M100) y humedad del combustible herbáceo vivo (Mherb). Cada individuo está representado por un vector de números con doble precisión, ya que para representar estas variables no es posible utilizar una representación binaria. Es importante señalar que cada individuo constituye la entrada de una simulación particular, y dichos valores de entrada son globales para todo el mapa simulado, es decir, para cada simulación se usan los mismos valores en cada celda del mapa.

Población	individuo 1	Model	WindSpd	WinDir	Slope	M1	M10	M100	Mherb
	individuo 2	Model	WindSpd	WinDir	Slope	M1	M10	M100	Mherb
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	individuo n	Model	WindSpd	WinDir	Slope	M1	M10	M100	Mherb

Individuo = Posible solución

Figura 4.14: Representación de una población en ESS-IM y HESS-IM.

4.3.2.2. Funcionamiento

Los workers, llevan a cabo la simulación y la evaluación de la aptitud de cada uno de los individuos en paralelo, retornando los resultados a su respectivo master. Cada master, además, es responsable de llevar a cabo la evolución de la población y de efectuar las migraciones de los individuos entre las distintas islas en función de la topología de comunicación y los parámetros de migración configurados. Es importante remarcar que la evolución de las islas se efectúa en paralelo.

Un esquema detallado del funcionamiento de ESS-IM se presenta en la Fig. 4.15. Puede observarse que, al igual que en ESS, el sistema cuenta con dos Etapas de Optimización (**EO-Worker** y **EO-Master**) las que son instanciadas en cada isla. Por su parte se diferencia de ESS por contar con dos Etapas de Calibración (**EC-Master** y **EC-Monitor**) y una Etapa de Predicción (**EP**) vinculada de forma directa al proceso monitor.

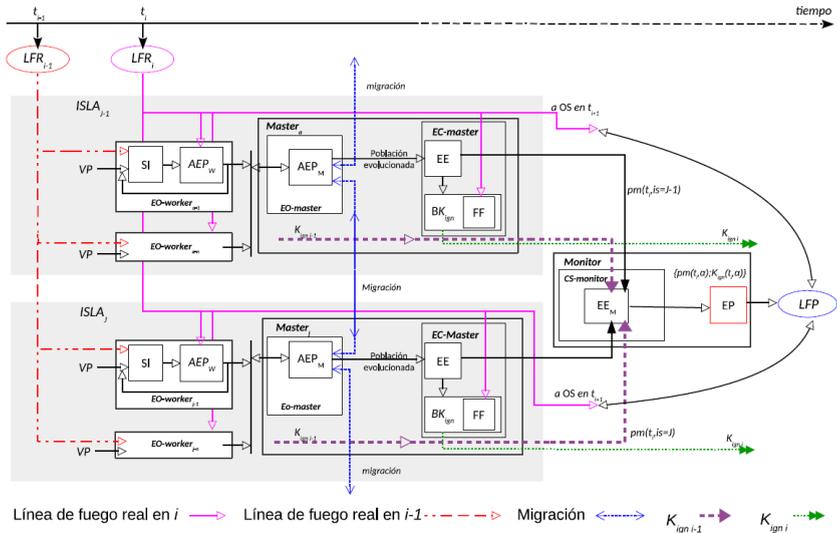


Figura 4.15: ESS-IM: Diagrama General del Sistema Estadístico Evolutivo con Modelo de Islas: **SI**: Simulador de Incendios; **AEP_M**: Algoritmo Evolutivo Paralelo en master; **AEP_W**: Algoritmo Evolutivo Paralelo en worker (evaluación de aptitud); **EO**: Etapa de Optimización; **EE**: Etapa Estadística; **EE_M**: Etapa Estadística monitor, **BK_{ign}**: Búsqueda K_{ign} ; **K_{ign}**: valor clave de ignición; **FF**: Función de Fitness; **EC**: Etapa de Calibración; **EP**: Etapa de Predicción; **LFP**: Línea de Fuego Propuesta; **LFR_x**: Línea de Fuego Real en el instante x ; **VP**: Vector de Parámetros de entrada; **pm**($t_i, is = j$): mapas de probabilidad generados por cada isla.

La etapa **EO-Worker** lleva a cabo la porción del Algoritmo Evolutivo Paralelo (**AEP_W**) que efectúa la evaluación de la población mediante la función de fitness. La etapa **EO-Master** opera en cada una de las islas, siendo responsable de llevar a cabo la evolución de la población, operación que se realiza iterando hasta que la población alcanza cierto nivel de calidad. Esta etapa también es encargada de efectuar la migración de individuos.

La etapa **EC-Master** se encarga de implementar el método estadístico [100]. El mapa de probabilidad generado por la **EE-Master** se utiliza como entrada por la etapa **BK** (Búsqueda de K_{ign}) para realizar la búsqueda del valor clave de ignición que será utilizado en el próximo instante de predicción.

La etapa **EC-Master** realiza las mismas funciones que en ESS, con la diferencia de que la salida de **EE** se envía a **EC-Monitor** junto con los $j - 1$ valores K_{ign} calculados por las restantes islas. Esta etapa ofrece tres modos de operación: **EC Monitor-1**, **EC-Monitor-2** y **EC-**

Monitor-3. Finalmente, la etapa **EP** realiza la predicción de la línea de fuego (**LFP**) para cada paso de simulación en función del mapa de probabilidad y del valor K_{ign} calculado por **EC-Monitor**. Cada uno de los modos de trabajo ofrece distintas ventajas y desventajas tanto en calidad de predicción como en rendimiento:

EC-Monitor-1 (mejor): En esta modalidad se selecciona el mejor valor K_{ign} entre los valores obtenidos por cada una de las islas. Esto implica que cada isla que haya concluido con el paso actual de simulación no podrá avanzar al paso siguiente hasta que todas las islas hayan concluido con el paso actual. En este esquema se prioriza la calidad de predicción penalizando el rendimiento del método, ya que pueden existir momentos en que los nodos de una isla completa no efectúen procesamiento alguno. Es importante resaltar aquí que una isla estará compuesta por al menos dos unidades de procesamiento. Sin embargo, éstas usualmente se conforman con mayor cantidad de nodos, por lo que la cantidad de procesadores o núcleos del clúster que estarán ociosos puede ser considerable.

EC-Monitor-2 (global): En esta modalidad, el valor de K_{ign} se calcula en función de un mapa estadístico generado a partir de los mapas estadísticos resultantes de cada isla. Esta configuración también implica una sincronización global debido a que es necesario esperar a que todas las islas envíen los resultados al nodo monitor.

EC-Monitor-3 (primero): En esta configuración se intenta brindar el mejor rendimiento posible, ya que la predicción se realiza utilizando el valor K_{ign} de la isla que finalice primero (es decir, el hallado en el menor tiempo posible). En términos de calidad de predicción, esta modalidad ofrece resultados variados ya que el mismo depende de las características de la población de la isla que finalice en menor tiempo.

4.3.2.3. Parámetros de inicialización

Al igual que en ESS, ESS-IM posee parámetros de inicialización, los cuales afectan directamente el rendimiento del método tanto en calidad de predicción como en tiempo de ejecución. Cada uno de estos parámetros son listados en la Tabla 4.4, junto con los parámetros involucrados en ESS. En dicha tabla “X” o “-” significa que un parámetro está o no presente en dicho método. En la Tabla 4.5 se puede observar una descripción detallada de los parámetros de ESS-IM, donde la mayoría de éstos se relacionan con el AEP.

Tabla 4.4: ESS y ESS-IM: Parámetros de configuración de cada método.

Parámetro	ESS	ESS-IM
Umbral de fitness	X	X
Cantidad de iteraciones	X	X
Cantidad de individuos	X	X
Cantidad de individuos por Isla	-	X
Cantidad de islas	-	X
Cantidad de individuos a migrar	-	X
Criterio de migración	-	X
Frecuencia de migración	-	X
Cantidad de workers	X	X
Cantidad de workers por Isla	-	X
Topología de comunicación	-	X
Criadores	X	X
Mínimo cruzamiento	X	X
Máximo cruzamiento	X	X
Porcentaje de mutación	X	X

4.3.2.4. ESS-IM-DE

La implementación basada en islas de ESS-IM sentó las bases del diseño de un framework que permitió continuar experimentando con otras metaheurísticas evolutivas paralelas. En este sentido, y siguiendo con la arquitectura de ESS-IM, se desarrolla ESS-IM con Evolución Diferencial (ESS-IM-DE) [103]. Esta implementación modifica la etapa de optimización del proceso *master* de ESS-IM (**EO-Master**), utilizando en lugar del AEP la metaheurística Evolución Diferencial. El funcionamiento es idéntico a ESS-IM, el proceso *master* de cada isla es el encargado de inicializar la población y aplicar los operadores de mutación y cruzamiento para generar nuevos individuos en la etapa

Tabla 4.5: ESS-IM: detalle de parámetros de configuración inicial.

Parámetro	Descripción	Rango
Umbral de fitness	valor de aptitud alcanzado por al menos un individuo de la población	[0,1]
Iteraciones	cantidad máxima de iteraciones a llevarse a cabo en una ejecución	Entero positivo

Individuos	número de individuos que componen la población de soluciones candidatas	Entero positivo
Individuos por islas	número de individuos que componen la población de soluciones candidatas dentro de una isla	Entero positivo
Islas	Número de instancias de islas a crearse que juntas componen la población de soluciones dentro del sistema	Entero positivo
Individuos a migrar	Número de individuos que serán migrados de una isla a otra en cada instancia de migración llevada a cabo entre las islas participantes	Entero positivo
Migración	Tipo de migración, elitista, semi-elitista, aleatoria, mixto	Entero positivo
Frecuencia	Número de iteraciones entre las cuales se produce una migración de individuos	Entero positivo
Workers	número de procesos que se instanciarán y participarán como procesos workers en toda la ejecución	Entero positivo (dependiente de la capacidad del cluster)
Workers por isla	número de procesos workers que conformarán una isla determinada	Entero positivo
Criadores	porcentaje de la población que será utilizado para generar nuevos individuos	[0,1]
Mín. cruzamiento	mínima probabilidad de cruzamiento	[0,1]
Máx. cruzamiento	máxima probabilidad de cruzamiento	[0,1]
Mutación	valor que determina la probabilidad de mutación	[0,1]

EO-Master. 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 aplicando la función de aptitud y retornando al **master** la aptitud obtenida.

El proceso *master*, determina si los individuos generados en el proceso de mutación y cruzamiento sobrevivirán a la siguiente generación. Para ello, verifica si la aptitud obtenida es mayor a la aptitud del individuo

corriente, si es así, el nuevo individuo reemplaza al individuo corriente. Este proceso se realiza para cada individuo de la población. Es importante remarcar que el resto de las etapas se mantienen sin modificaciones.

4.3.3. HESS-IM

Tal como se describió en el Capítulo 2, las metaheurísticas pueden paralelizarse siguiendo diferentes criterios y estrategias, algunas de ellas sólo mejoran el tiempo de respuesta del algoritmo sin influir en los resultados obtenidos [97]. Aprovechando el esquema de paralelización basado en islas con doble jerarquía master-worker implementado en ESS-IM, y basado en trabajos como [160, 161], se decidió potenciar la capacidad de búsqueda del algoritmo mediante la hibridación de metaheurísticas. Siguiendo con el esquema de ESS-IM, la hibridación contempló sólo metaheurísticas poblacionales, pero también se evaluaron alternativas que incluían metaheurísticas de trayectoria.

Tal como se ha mencionado en secciones previas, HESS-IM utiliza una técnica de optimización híbrida y colaborativa entre diferentes metaheurísticas poblacionales bajo un esquema de paralelización basado en islas con doble jerarquía master-worker. Este esquema permite a cada isla operar con una metaheurística diferente sobre el mismo espacio de búsqueda, intercambiando entre sí soluciones candidatas cada ciertos intervalos de tiempo, o a demanda de determinados eventos. El esquema de comunicación entre los diferentes procesos se mantiene similar al de ESS-IM. En el nivel superior se encuentra el *modelo master-worker 1*, donde un proceso master global coordina el funcionamiento de la totalidad de las islas. El master local de cada isla cumple la función de worker respecto del master global. Posteriormente, en el segundo nivel de jerarquía master-worker (*modelo master-worker 2*) cada master de isla controla a un conjunto de workers dentro de cada isla. Es importante remarcar aquí una diferencia fundamental entre ESS-IM y HESS-IM. Si bien ambos inicializan una población diferente en cada isla, HESS-IM da un tratamiento particular a cada una, ya que son perturbadas o modificadas mediante diferentes metaheurísticas. La colaboración entre cada metaheurísticas se lleva a cabo por el operador de migración, el cual intercambia individuos entre las diferentes islas y/o metaheurísticas cada ciertos eventos (por ejemplo, cantidad de generaciones realizadas, valor de aptitud de la población, detección de estancamientos en óptimos locales, etc.).

HESS-IM es un método general de reducción de incertidumbre que ha sido plateado como un framework adaptativo para la implementación de técnicas de optimización colaborativa basada en diferentes metaheurísticas evolutivas/poblacionales. Asumimos que un framework es un conjunto de clases que incorpora un diseño abstracto para soluciones a una familia de problemas relacionados [162]. De esta manera, HESS-IM permite vincular n metaheurísticas evolutivas para resolver de forma colaborativa la optimización de soluciones del método de predicción. Este trabajo presenta la implementación de HESS-IM utilizando tres metaheurísticas poblacionales evolutivas: a) Algoritmos Evolutivos, b) Evolución Diferencial y c) Optimización por Cúmulo de Partículas utilizando MPI conjuntamente con CUDA. Se realizaron experimentaciones contrastando resultados entre las diferentes metaheurísticas EA, EA/DE y EA/DE/PSO.

Algoritmo 7: Pseudo-código del funcionamiento de HESS-IM.

Input:

-mapa de entrada (LFR_n), /* monitor */
 -parámetros de entrada del escenario,
 -parámetros de entrada de metaheurísticas,
 $i = 1$; /* Definición del paso inicial */ /* monitor */
 $k = p_predicción()$; /* Calcula cantidad de pasos de predicción*/ /* monitor */

while $i \leq k$ **do**

```

  /* Comienza bucle por cada paso de simulación */ /* monitor */
  k++;
  while iteraciones < max_iter & fitness < umbral do
    /* Comienza bucle de iteración evolutiva */ /* master */
    instanciación de metaheurísticas según parámetros de entrada; /* master */
    evaluación de soluciones candidatas; /* worker */
    operadores de cada metaheurística; /* master */
    migración de soluciones; /* master */
    evolución de la población; /* master */
  calcular  $Kign_{i+1}$ ; /* master */ /* monitor */
  if  $i \geq 2$  then
    _ predicción ( $PFL_i + 1$ ); /* monitor */

```

Output:

- un total de $k - 1$ valores de calidad de predicción por paso
 - mapa de evolución de incendio para cada paso de predicción mejor solución encontrada

Un esquema general del funcionamiento de HESS-IM puede observarse en la Fig. 4.16 y en el Algoritmo 7. La cadena de predicción comienza en el proceso *monitor*, éste envía a cada una de las n islas dos conjuntos de datos iniciales:

a) Conjunto de datos del incendio: un mapa con el frente de fuego real, intervalos de tiempo a considerar para cada paso de predicción, valores y rangos para cada parámetro de entrada del modelo.

b) Conjunto de datos de las metaheurísticas: cantidad de islas, metaheurística que cada isla deberá utilizar, cantidad de workers por isla, parámetros de migración, parámetros propios de cada metaheurística. Un detalle de estos parámetros se puede observar en la Tabla 4.6.

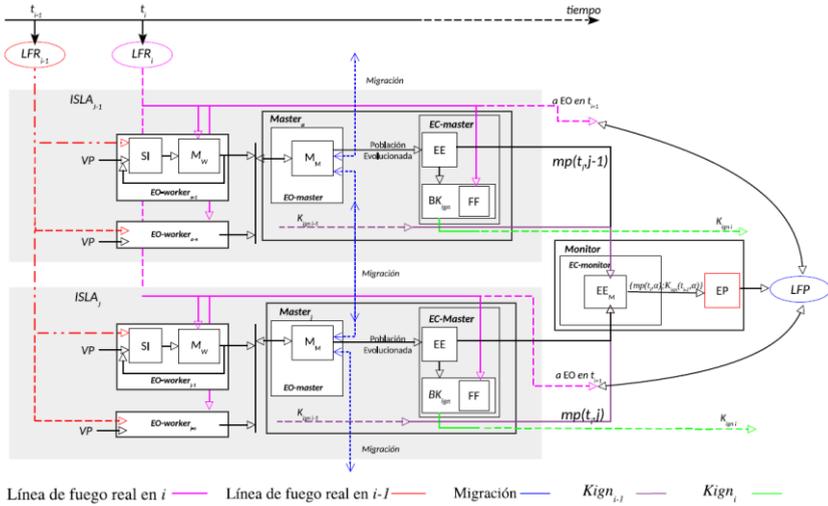


Figura 4.16: HESS-IM: Sistema Estadístico Híbrido Evolutivo con Modelo de Islas. **SI**: simulador de incendios; **M_M** : etapa metaheurística en master; **M_W** : etapa metaheurística en worker (evaluación de fitness); **EO**: etapa de optimización; **EE**: etapa estadística; **BK_{ign}** : búsqueda de K_{ign} ; **K_{ign}** : valor clave de ignición; **FF**: función de fitness; **EC**: etapa de calibración; **EP**: etapa de predicción; **LFP**: línea de fuego predicha; **LFR_x** : línea de fuego real en tiempo x ; **VP**: vector de parámetros de entrada; **EE_M** : etapa estadística en proceso monitor; **mp** : mapa de probabilidad.

Dichos conjuntos de datos son recibidos por el proceso *master* de cada isla, donde se instancia la Etapa de Optimización (EO_{master}), que da inicio a la metaheurística requerida por el proceso *monitor*, ya sea: EA, DE o PSO. Dicho procesamiento se realiza en la Etapa Metaheurística, la cual se encuentra dividida en dos subetapas: por un lado, la Etapa Metaheurística, del proceso master (M_M), y por el otro, la Etapa Metaheurística de los procesos workers (M_W).

Tabla 4.6: HESS-IM: detalle de parámetros de configuración inicial.

Parámetro	Descripción	Rango
Umbral de fitness	valor de aptitud alcanzado por al menos un individuo de la población	[0,1]
Iteraciones	cantidad máxima de iteraciones a llevarse a cabo en una ejecución	Entero positivo
Individuos	número de individuos que componen la población de soluciones candidatas	Entero positivo
Individuos por islas	número de individuos que componen la población de soluciones candidatas dentro de una isla	Entero positivo
Islas	Número de instancias de islas a crearse que juntas componen la población de soluciones dentro del sistema	Entero positivo
Individuos a migrar	Número de individuos que serán migrados de una isla a otra en cada instancia de migración llevada a cabo entre las islas participantes	Entero positivo
Criterio	Tipo de migración, elitista, semi-elitista, aleatoria, mixto	Entero positivo
Frecuencia	Número de iteraciones entre las cuales se produce una migración de individuos	Entero positivo
Workers	número de procesos que se instanciarán y participarán como procesos workers en toda la ejecución en toda la ejecución	Entero positivo (dependiente de la capacidad del cluster)
Workers por isla	número de procesos que serán designados a la función de worker en cada isla	Entero positivo
(EA) Criadores	porcentaje de la población que será utilizado para generar nuevos individuos	[0,1]
(EA) Mín. cruzamiento	mínima probabilidad de cruzamiento	[0,1]
(EA) Máx. cruzamiento	máxima probabilidad de cruzamiento	[0,1]

(EA) Mutación	valor que determina la probabilidad de mutación	[0,1]
(DE) f, Factor de mutación	Constante F utilizada en el operador de mutación de Evolución Diferencial para generar individuos candidatos	[0,1]
(DE) Cr, Cruzamiento	Constante Cr utilizada en el operador de cruzamiento de Evolución Diferencial para generar individuos candidatos	[0,1]
(PSO) c1 y c2, Coeficientes de aceleración	Representan la atracción que tiene una partícula hacia su propio éxito o hacia el éxito de sus vecinos.	$\leq 2,0$
(PSO) w, Factor de inercia	Valor utilizado para determinar el equilibrio entre la búsqueda global y local.	[0,4; 0,9]
(PSO) Tamaño del vecindario	Valor que determina diferentes vías de comunicación entre sus miembros	$[2, n * \frac{n-1}{2}]$

La etapa M_M realiza diferentes operaciones dependiendo del momento en que se encuentre operando el método y en función de la metaheurística configurada en dicha isla. Sin embargo, existen ciertas funciones que todas las islas realizan en común:

- 1) Inicialización de la población: al inicio del proceso de predicción, esta etapa se encarga de inicializar la población de individuos y de enviar los mismos a los workers para que éstos sean tratados.
- 2) Migración de individuos: mientras los workers procesan a los individuos, la etapa M_M lleva a cabo la migración de individuos hacia las islas vecinas. La migración se realiza en función de la topología de comunicación y tipo de migración configurada. Esta última puede ser: a) con mayor frecuencia entre islas que operen con la misma metaheurística, o b) con un único valor de frecuencia para todas las islas.
- 3) Evolución de la población: cuando se alcanza la cantidad máxima de iteraciones o el umbral de aptitud, dentro de cada iteración, se realiza la evolución de la población hacia la próxima generación evolutiva.

4) Finalización y envío: cuando la población ha evolucionado, cumpliendo la condición de finalización, se envía, desde la etapa M_M la población final de individuos hacia la Etapa de Calibración (EC_{master}).

Por su lado, los workers tienen como función principal evaluar la aptitud de cada individuo mediante la utilización del simulador de comportamiento de incendios forestales (SI). Dicho simulador está basado en el modelo definido por Rothermel [47] e implementado mediante la librería *fireLib* [109] para CPU y CUDA-FGM [158] para GPU, la cual corresponde a una implementación de *firelib* para GPU. El cálculo de la evaluación de aptitud en el instante t_i necesita contar con la línea de fuego real del incendio (LFR) en t_{i-1} (es decir, LFR_{i-1}) y los valores de los parámetros de entrada, los cuales se almacenan en los vectores de parámetros (VP) (los individuos de la población).

La función de aptitud utilizada para determinar la calidad de cada individuo está basada en el índice de Jaccard [155], ver Ec. (4.2). Esta función permite comparar el mapa simulado con el mapa real en t_i (LFR_i) y, en función de la diferencia entre los dos mapas, se determina el valor de aptitud para cada individuo. La Etapa de Calibración del proceso master (EC_{master}) calcula un mapa de probabilidad en función de todos los individuos de la población final.

Dicho mapa se utiliza para obtener el valor clave de ignición (Key Ignition Value, K_{ign}), el cual representa el patrón de comportamiento del incendio y es obtenido en la etapa Búsqueda del K_{ign} (BK_{ign}).

Para cada instante de tiempo i , cada isla j genera un $K_{ign}(t_i, j)$ y un mapa de probabilidad $m_p(t_i, j)$, los cuales son enviados a la etapa de calibración del proceso monitor ($EC_{Monitor}$). El primero es enviado a $EC_{Monitor}$ en t_{i+1} y el segundo a $EC_{Monitor}$ en t_i , de esta manera se conforma en t_i un par $\{m_p(t_i, \alpha); K_{ign}(t_i - 1, \alpha)\}$, donde α indica el número de isla que obtuvo mejor desempeño en términos de aptitud. Éstos valores finalmente son ingresados en la etapa de predicción (EP) la cual realiza el cálculo de predicción (línea de fuego predicha, LFP) para el próximo instante de tiempo LFR_{i+1} . Debido a esto HESS-IM no genera predicción en el primer instante de tiempo, ya que en dicho momento se calcula el primer K_{ign} .

Es importante tener en cuenta que la hibridación tiene en consideración las diferentes velocidades de los algoritmos, con el propósito de que todas las metaheurísticas avancen al mismo tiempo. Se estudiaron configuraciones de cada metaheurística con el objetivo de garantizar dicho fin, aunque de igual manera los intercambios de información

entre las islas se realizan de manera asincrónica, evitando de este modo cualquier problema por desacoplamiento de velocidades.

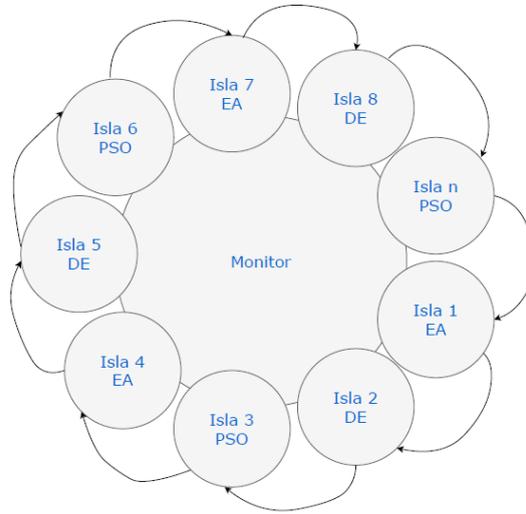


Figura 4.17: Ejemplo de generación de topología de comunicación entre islas según metaheurísticas instanciadas.

Con el propósito de mantener la diversidad en la búsqueda, se considera una topología de anillo donde se repiten las tres metaheurísticas por igual y en diferentes órdenes de secuencia, haciendo de esta manera que la búsqueda sea más completa (por ejemplo, ver Fig. 4.17), manteniendo un esquema de migración unidireccional de un solo sentido. Dentro de la experimentación llevada a cabo también se testeó el resultado de HESS-IM utilizando una topología anillo aleatoria, donde las islas pueden compartir información con cualquier isla (de cualquier metaheurística instanciada).

4.4. Síntesis

Este capítulo ha presentado la arquitectura y los detalles de la implementación de los métodos de reducción de incertidumbre ESS-IM y HESS-IM, desde el simulador utilizado hasta la integración de sus diferentes componentes: estadística, optimización metaheurística y computación de alto rendimiento. Además, se ha proporcionado detalles del funcionamiento interno que implica dicha integración.

Evaluación Experimental y Análisis de Resultados

El presente capítulo describe las experimentaciones realizadas a fin de evaluar y comparar los resultados obtenidos tras la aplicación de los métodos presentados en el capítulo anterior en casos de estudio correspondientes a quemas controladas e incendios forestales. A lo largo del mismo, se han evaluado distintos aspectos y configuraciones de los métodos desarrollados. En cada experimento se describe la configuración de los algoritmos, los objetivos a evaluar, y se realiza un análisis de los resultados obtenidos.

5.1. Introducción

Los métodos presentados han sido evaluados en dos tipos de casos de estudio o escenarios. Por un lado, los casos de quemas prescritas o quemas controladas, los cuales se caracterizan por desarrollarse mediante el uso controlado del fuego para disminuir la vegetación bajo ciertas condiciones específicas según un objetivo propuesto. Y por el otro lado, los casos denominados incendios forestales, los cuales corresponden a incendios ocurridos en condiciones naturales, es decir, fuegos no programados o programados y que han salido de control, que afectan de diversas formas a las superficies forestales.

5.1.1. Casos de quemas controladas

Estos casos se tratan de quemas controladas intencionales llevadas a cabo en Portugal (precisamente en Serra de Lousã, Gestosa) en el marco del proyecto SPREAD [163]. Se trata precisamente de 5 terrenos, cada uno posee características diferentes respecto a dimensiones, pendiente, tiempo de inicio de ignición inicial, entre otros. Para generar cada caso de estudio fue necesario realizar en cada terreno una quema intencional, registrando la evolución del avance del fuego durante cierta cantidad de minutos. El período de duración total de cada incendio ha sido dividido

en lapsos o intervalos de tiempo denominados pasos de simulación. Estos pasos de simulación corresponden a instantes en donde se realizan las comparaciones de calidad de predicción entre el resultado de la simulación y el estado real del incendio.

En la Tabla 5.1 puede observarse la información detallada para cada uno de los cinco casos de estudio, y en las Figs. 5.1, 5.2, 5.3, 5.4 y 5.5 la representación de la evolución del fuego en cada terreno. En cada figura la coloración negra representa el área que no ha sido alcanzada por el fuego (celdas no quemadas), mientras que el resto de las coloraciones indican cómo el fuego ha ido avanzando sobre el terreno. Además, en cada figura se observa el tamaño del terreno en cada uno de sus lados, como así también el tiempo en minutos en que el fuego ha ido alcanzando cada porción del terreno.

Tabla 5.1: Casos de quemas controladas: ancho y largo en celdas, superficie en (m²), pendiente del terreno (en grados) y detalle de cada paso de simulación (en minutos).

Ca- so	Anch o (m)	Largo (m)	Sup. (m ²)	Pendiente (grados)	Inicio (min)	Incremento (min)	Fin (min)
1	20	30	600	6	2,0	2,0	10,0
2	89	109	9.701	21	2,0	2,0	14,0
3	95	123	11.68 5	21	2,0	2,0	12,0
4	89	91	8.099	21	2,5	2,5	12,5
5	75	126	9.450	19	3,0	1,0	9,0

A continuación, se presenta un detalle de cada uno de los casos de estudio:

- **Caso 1:** Este terreno posee una superficie de 600 (m²), con celdas de 0,3333m x 0,3333m de tamaño. El fuego ha sido iniciado de forma lineal sobre un borde del terreno. El incendio es registrado desde el minuto 2 hasta el minuto 10, y las muestras son generadas cada 2 minutos, por lo tanto, el caso posee tres pasos de predicción: minuto 6, 8 y 10, ver Fig. 5.1.
- **Caso 2:** Este terreno posee una superficie de 9.701 (m²), donde el fuego ha sido iniciado de forma lineal sobre un borde del terreno en el minuto 2 finalizando en el minuto 14, este experimento posee cinco pasos predicción: 6, 8, 10, 12 y 14, ver Fig. 5.2. En este caso, al igual que en los restantes, el tamaño de celda es de 1m x 1m.

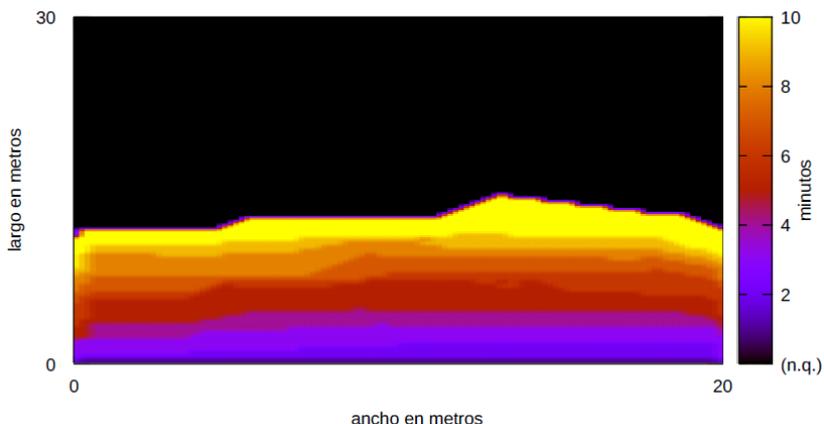


Figura 5.1: Caso 1: Representación gráfica del terreno y la evolución del incendio sobre el mismo.

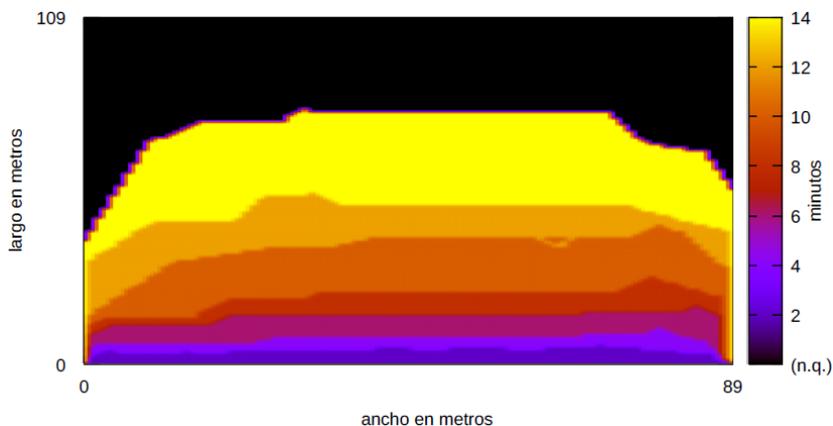


Figura 5.2: Caso 2: Representación gráfica del terreno y la evolución del incendio sobre el mismo.

- **Caso 3:** Terreno con mayor superficie $11.685 \text{ (m}^2\text{)}$, donde el fuego ha sido iniciado en un punto dentro del terreno. La quema ha sido registrada con inicio en el minuto 2 finalizando en el minuto 12 (con intervalos de 2 minutos). Este experimento posee cuatro pasos de predicción: minutos 6, 8, 10 y 12, ver Fig. 5.3.

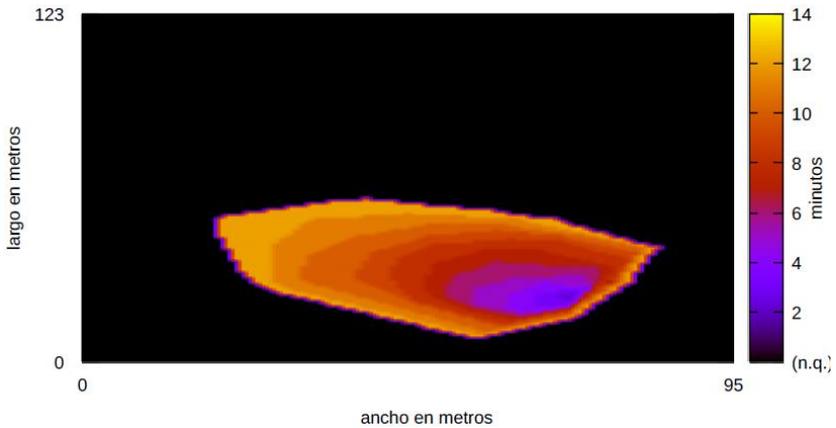


Figura 5.3: Caso 3: Representación gráfica del terreno y la evolución del incendio sobre el mismo.

- **Caso 4:** Este terreno posee una superficie de $8.099 \text{ (m}^2\text{)}$, fuego iniciado de forma lineal sobre un borde del terreno en el minuto 2,5 finalizando en el minuto 12,5. Experimento con tres pasos de predicción: minutos 7,5, 10 y 12,5, ver Fig. 5.4.

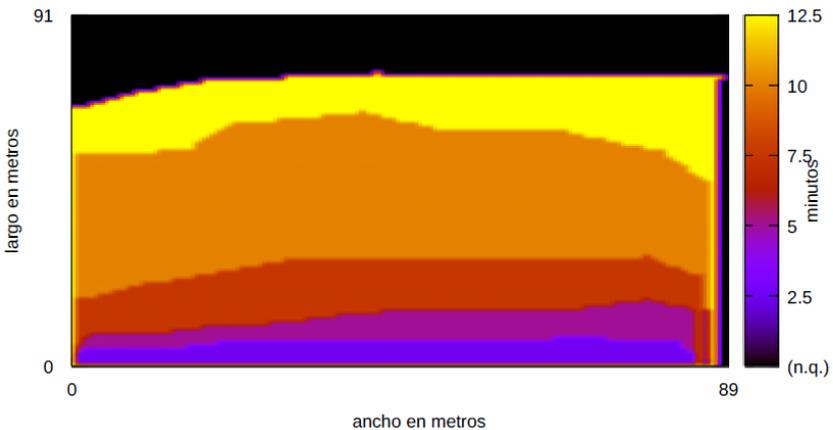


Figura 5.4: Caso 4: Representación gráfica del terreno y la evolución del incendio sobre el mismo.

- **Caso 5:** Este terreno posee una superficie de $9.450 \text{ (m}^2\text{)}$, con fuego iniciado de forma central en el interior del terreno en el minuto 3 finalizando en el minuto 9, pero con incrementos de 1 minuto. Por lo tanto, este experimento cuenta con cinco pasos de predicción: minutos 5, 6, 7, 8 y 9, ver Fig. 5.5.

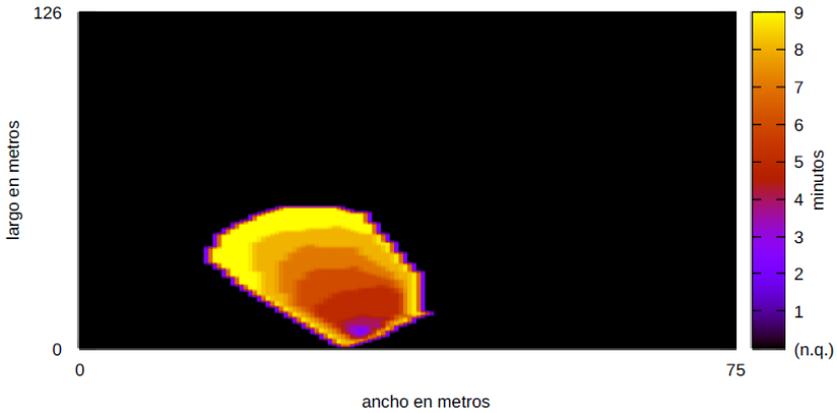


Figura 5.5: Caso 5: Representación gráfica del terreno y la evolución del incendio sobre el mismo.

5.1.2. Casos de incendios forestales

Los llamados casos de incendios forestales, corresponden a incendios ocurridos en condiciones naturales, por consiguiente, se trata de terrenos de grandes dimensiones. En la Tabla 5.2 puede observarse la información detallada para ambos incendios, como el tamaño de la superficie quemada (medida en hectáreas), duración, ubicación y el año de producido el incendio. Es importante mencionar que estos incendios no han sido seleccionados al azar, es decir para su elección se han considerado que estos cumplan con dos requisitos fundamentales: a) que la duración del incendio nos permita contar con al menos 4 instantes de simulación, y b) que las imágenes satelitales disponibles del evento no contaran con cobertura de nubes que impidieran determinar con precisión el área quemada por los incendios. Finalmente es importante destacar que, en este tipo de casos de estudio, las dificultades para lograr una predicción aceptable aumentan considerablemente, debido a la complejidad de los escenarios. Los incendios forestales seleccionados corresponden a dos fenómenos ocurridos en diferentes años, en 2008 y 2018, en Argentina y Estados Unidos, respectivamente.

Tabla 5.2: Casos de incendios forestales: ubicación, superficie total del mapa tratado, superficie total quemada al finalizar el incendio, duración del mismo, año de ocurrencia.

Caso	Ubicación	Sup. Q. (ha)	Duración (días)	Año
1	Córdoba (AR)	51.852	6	2008
2	California (EU)	19.797	7	2018

El primero de ellos, denominado en este trabajo como caso de incendio forestal 1, corresponde a un incendio ocurrido en el año 2008 en la localidad de Capilla del Monte, Provincia de Córdoba, Argentina. Cuya duración se presenta desde el 29 de agosto al 3 de septiembre de dicho año [164]. El área se ubica en el cordón oriental de las Sierras Chicas, que abarcó los departamentos de Ischilín, Punilla, Totoral y Colón, con una superficie aproximada de 51.852 hectáreas (Fig. 5.6).

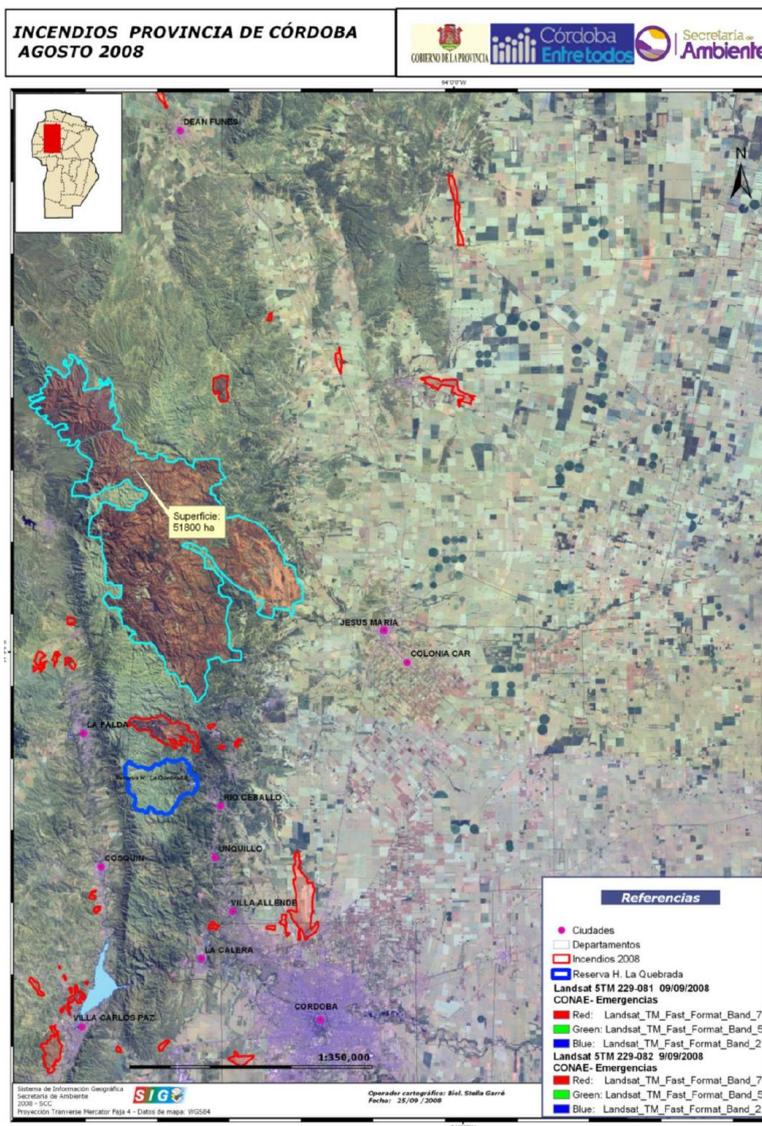


Figura 5.6: Caso de incendio forestal 1: Incendios Provincia de Córdoba, 2008. Fuente Secretaría de Ambiente, Gobierno de Córdoba.

Según la interpretación realizada por la Secretaría de Ambiente de la Provincia de Córdoba, a partir de imágenes Landsat 5 TM, del día 9 de septiembre de dicho año (provista por CONAE-Emergencias), ocurrieron 49 focos de incendios en la zona centro-norte de la provincia de Córdoba. Estimando una superficie total afectada de 66.000 hectáreas aproximadamente. Cabe aclarar que producto de este incendio poblaciones aledañas debieron ser evacuadas. En la Fig. 5.7 puede observarse el área quemada en el último día de incendio.

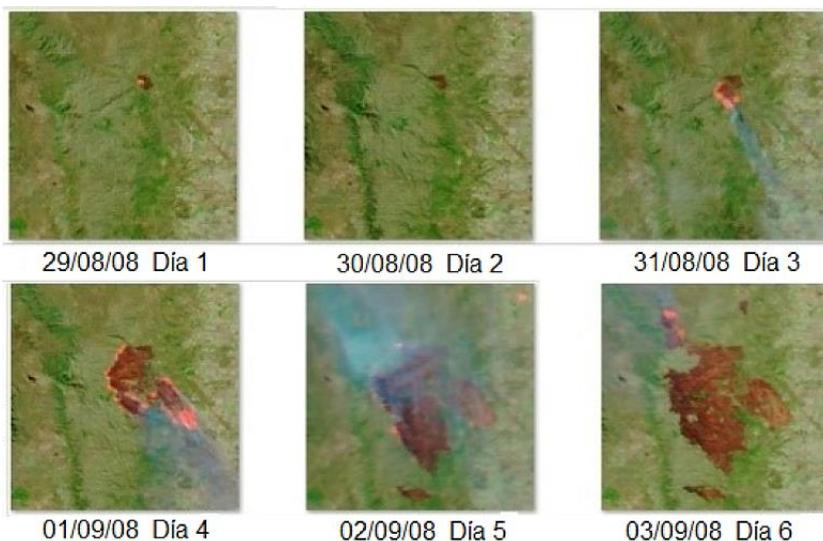


Figura 5.7: Caso de incendio forestal 1: Secuencia de imágenes satelitales día a día donde puede apreciarse el área quemada y no quemada (evolución del incendio), como así también la presencia de humo desde el día 3 al día 5. Fuente MODIS.

Respecto al segundo caso (caso de incendio forestal 2), el mismo corresponde a un incendio ocurrido en el año 2018, en California en el Condado de Mendocino, entre los meses de julio y agosto. Este incendio ha sido de los mayores registrados en la historia de California [165]. El evento estuvo compuesto por dos incendios separados denominados: River Fire y Ranch Fire (ver Fig. 5.9), que ardieron en los condados de Mendocino, Lake, Colusa y Glenn consumiendo más de 180.000 hectáreas. Este caso de estudio corresponde al incendio del sur (River Fire), ocurrido en el noreste de Hopland, California, el cual consumió 19.797 hectáreas.

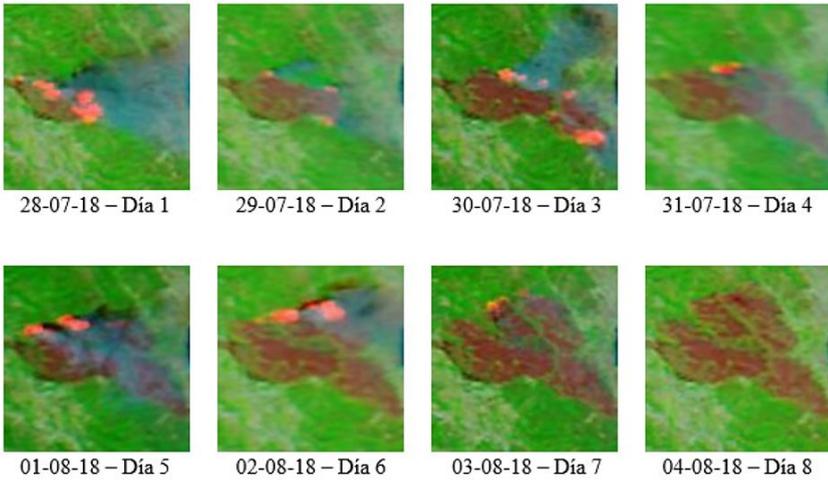


Figura 5.8: Caso de incendio forestal 2: Incendio en California, 2018. Secuencia de imágenes satelitales día a día, donde puede apreciarse el área quemada y no quemada (evolución del incendio). Fuente MODIS.

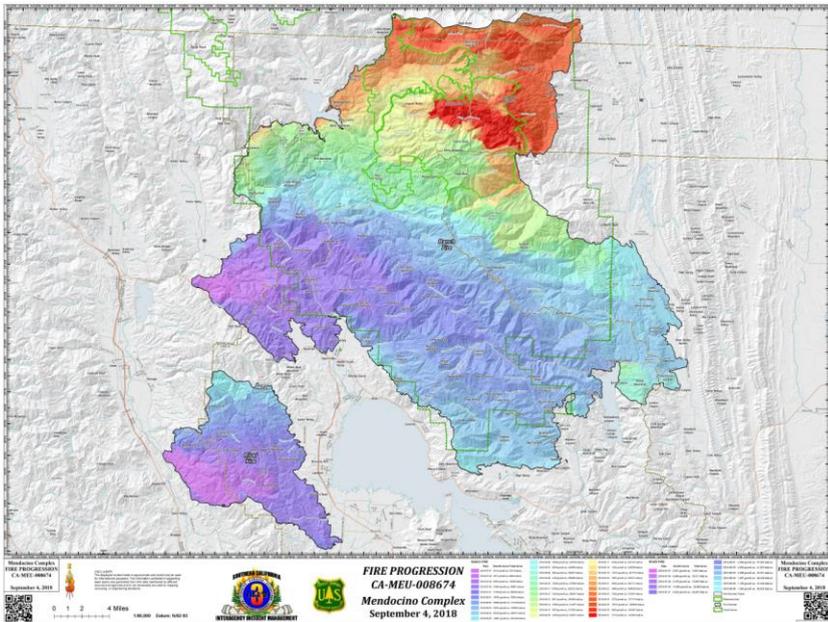


Figura 5.9: Mapa de progresión y perímetros de área quemada de incendios Ranch y River fires [166], 4 de septiembre de 2018.

Si bien el incendio tuvo una duración mayor al presentado en las imágenes de la Fig. 5.8, el mapa utilizado en el sistema de predicción

únicamente registra los días: 27, 28, 29, 30, 31 de julio, y para agosto los días 1 y 2. Los detalles de cómo estos casos fueron generados a partir de los registros de imágenes satelitales se describen en el Capítulo 7.

5.2. Experimentos

A continuación, se presentan los resultados obtenidos en las diferentes experimentaciones llevadas a cabo durante el desarrollo de la presente tesis. Se presenta un análisis comparativo entre diferentes métodos, desde el punto de vista de calidad de resultados (calidad de predicción) y en términos de tiempo de ejecución.

La experimentación realizada se encuentra dividida en tres secciones:

- Comparación A): Esta sección presenta una comparación entre los métodos: S^2F^2M , ESS y ESS-IM. La experimentación utiliza implementaciones basadas en MPI.
- Comparación B): Esta sección presenta una comparación entre los métodos: ESS y ESS-IM. La experimentación utiliza implementaciones basadas en MPI.
- Comparación C): Esta sección presenta una comparación entre los métodos: ESS-IM y HESS-IM (EA y DE) y ESS-IM y HESS-IM (EA, DE y PSO). La experimentación utiliza implementaciones basadas en MPI y GPU.

5.2.1. Comparación A): S^2F^2M , ESS y ESS-IM

En este apartado se presentan los resultados comparativos obtenidos tras aplicar los métodos S^2F^2M , ESS y ESS-IM a los 5 experimentos correspondientes a incendios controlados presentados en la sección 5.1.1. En cada experimento se han definido pasos de tiempo discretos para representar el avance del frente de fuego. En la Tabla 5.1, junto con la información de las dimensiones y pendientes para cada uno de los casos de estudio utilizados, se puede observar también los valores de tiempo inicial y final de cada simulación y su respectivo valor de incremento.

Tal como se mencionó en la sección 3.2, es importante tener en cuenta que, para fines de calibración y predicción, los tres métodos aquí comparados necesitan de una línea de fuego real adicional a la que precisa la predicción clásica, por lo que no pueden proporcionar predicciones al finalizar el primer paso de simulación. En dicho

instante, estos métodos aplican la etapa de calibración. Por esta razón, las gráficas presentan los resultados a partir de la predicción generada en el segundo paso de simulación (predicción realizada entre t_2 y t_3 , ver Fig. 4.12).

Es importante tener en cuenta que cada método opera con diferentes conjuntos de parámetros de entrada, por su lado, S²F²M utiliza como entrada los rangos de valores para cada parámetro de fireLib, como así también parámetros que indican los recursos asignados, en cambio ESS y ESS-IM necesitan además de éstos, aquellos que determinan el entorno de ejecución del algoritmo evolutivo (ver sección 4.3). Tales parámetros pueden condicionar fuertemente el desempeño de cada método, ante el mismo caso de análisis. Por este motivo, teniendo en cuenta la información acerca de los experimentos y del modelo de Rothermel [120], para algunos de los parámetros se han especificado rangos (en particular en aquellos que presentan incertidumbre), y, además, parte de esta información se ha medido durante los experimentos, y el resto se ha tomado de los valores estándar usados por BehavePlus [51].

En la Tabla 5.3 se presentan cada uno de los parámetros que alimentan al modelo de predicción de incendios forestales.

Tabla 5.3: Detalle de parámetros utilizados por fireLib.

Parámetro	Descripción	Rango	Unidad/medición
Model	Modelo de combustible de Rothermel	1-13	-
WindSpd	Velocidad del viento	0-80	millas/horas
WindDir	Dirección del viento	0-360	grados en sentido horario desde el norte
M1	Humedad del combustible muerto en 1 hora	1-60	porcentaje
M10	Humedad del combustible muerto en 10 horas	1-60	porcentaje
M100	Humedad del combustible muerto en 100 horas	1-60	porcentaje
Mherb	Humedad del combustible herbáceo vivo	30-300	porcentaje
Slope	Pendiente del terreno	0-81	grados
Aspect	Dirección de la pendiente	0-360	grados en sentido horario desde el norte

Es importante tener en cuenta que estos parámetros están presentes en los tres métodos, y los valores de cada uno de ellos son particulares para cada experimento. Dado que S²F²M utiliza el concepto de experimento

factorial, aquellos parámetros con valores dinámicos, además de contar con el rango determinado, poseen un valor de incremento utilizado para recorrer así cada rango de valores. En los otros dos métodos, al reemplazar el experimento factorial por un método de optimización metaheurístico, sólo se utilizan los rangos de cada parámetro para la generación de los individuos, siendo el valor de incremento innecesario. De este modo, S²F²M no requiere de parámetros adicionales a los descritos en la Tabla 5.3. En las Tablas 5.4, 5.5 se presentan los parámetros utilizados por ESS y ESS-IM, respectivamente, los cuales han sido seleccionados a base de un estudio de la literatura y de un análisis de calibración de parámetros.

Tabla 5.4: ESS: parámetros utilizados en metaheurística.

Parámetro	Valor
Umbral de fitness	0,7
Iteraciones	200
Individuos	200
Workers	2, 4, 8, 16, 32
Criadores	0,8
Mínima prob. cruzamiento	0,2
Máxima prob. cruzamiento	0,6
Mutación	0,5

5.2.1.1. Evaluación de Calidad de Predicción

En la Fig. 5.10, se presentan los resultados de calidad de predicción (aptitud o fitness) obtenida por cada método para cada caso de estudio. Las gráficas presentan en el eje de abscisas el valor en minutos correspondiente a cada instante de predicción, y en el eje de ordenadas, el valor de aptitud (dentro del rango [0, 1]), donde un valor igual a 1 indica una predicción perfecta, y por el contrario un valor igual a 0 una predicción completamente errónea.

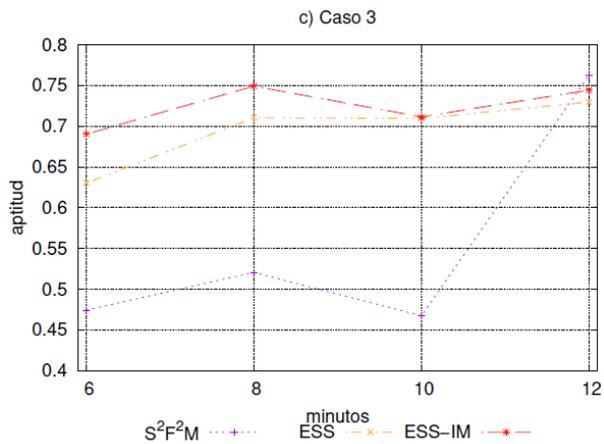
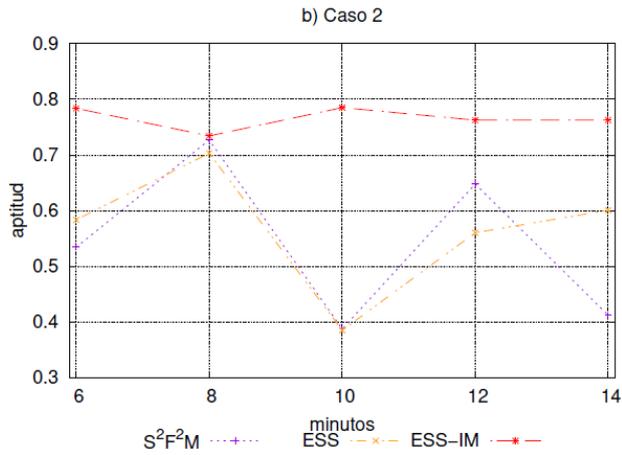
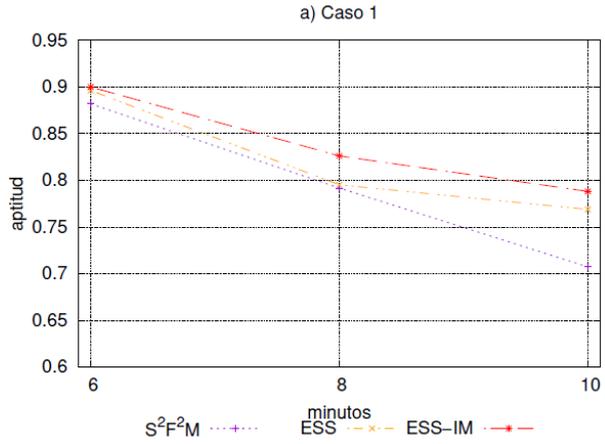
Tal como se puede observar, en la totalidad de los casos presentados, la respuesta de ESS-IM supera la de los métodos: S²F²M y ESS. Sin embargo, en momentos puntuales, los valores obtenidos por ESS-IM son levemente inferiores a los resultados de alguno de los otros métodos (como sucede en el caso 4 en el minuto 12,5, en el caso 3 en el minuto 12 y en el caso 5 en el minuto 7). Los resultados demuestran que el

incremento en paralelismo del AEP permite optimizar el funcionamiento del mecanismo de búsqueda interno, esto es debido a que en ESS la metaheurística opera con una única población de individuos, y en ESS-IM el método opera múltiples poblaciones (distribuidas en islas) que realizan una búsqueda más eficiente sobre el espacio de búsqueda. Con el propósito de que cada isla no converja a óptimos locales es introducido el operador de migración, el cual intercambia individuos entre las diferentes islas/poblaciones. Este cambio de arquitectura (entre ESS y ESS-IM) logra de esta manera mejorar la calidad de predicción. Otra observación a tener en cuenta es que debido al comportamiento determinístico de ESS y ESS-IM, la experimentación de estos métodos debe contemplar un valor promedio de una cierta cantidad de ejecuciones.

Tabla 5.5: ESS-IM: parámetros utilizados en metaheurística.

Parámetro	Valor
Umbral de fitness	0,7
Iteraciones	200
Individuos por isla	200
Islas	6
Individuos a migrar	6%
Migración	Elitista
Frecuencia	cada 2 generaciones
Dirección de migración	sentido horario
Topología de comunicación	anillo
Workers por isla	2-6
Criadores	0,8
Mínima prob. cruzamiento	0,2
Máxima prob. cruzamiento	0,6
Mutación	0,5

En este sentido, los valores indicados tanto para ESS como para ESS-IM corresponden al promedio de 20 ejecuciones, no así en cambio para S²F²M, ya que este brinda una salida determinista.



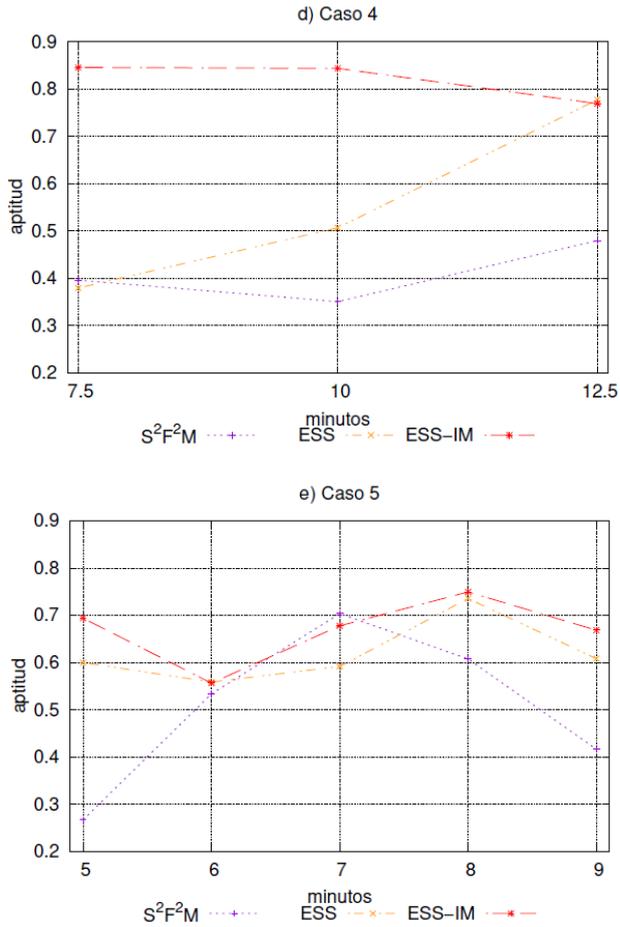


Figura 5.10: Comparación de calidad de predicción entre métodos: S²F²M, ESS y ESS-IM. Casos de quemas controladas 1, 2, 3, 4 y 5.

5.2.1.2. Evaluación de Rendimiento

Respecto a la evaluación en términos de tiempo de ejecución, se realizó una evaluación de rendimiento únicamente entre los métodos evolutivos (ESS y ESS IM), ya que S²F²M ya había sido contrastado con ESS en estudios previos [101]. Con el propósito diseñar la experimentación de la forma más equitativa posible, y considerando la arquitectura de cada método (ESS opera con una única población y ESS-IM opera con múltiples poblaciones) se evaluaron diferentes alternativas. Finalmente se determinó que la carga de individuos entre cada experimentación debía ser balanceada. De esta manera se

establecieron diferentes experimentos en donde ambos métodos utilizaban la misma cantidad de individuos. Por lo tanto, ESS operó en las experimentaciones con 204 individuos, mientras que ESS-IM utilizó tres configuraciones diferentes: a) 2 islas de 102 individuos, b) 4 islas de 51 individuos y c) 6 islas de 34 individuos.

Los resultados obtenidos se presentan en la Tabla 5.6, el tiempo se encuentra representado en horas, minutos y segundos. En ella se puede observar un rendimiento prácticamente equivalente entre ambos métodos al considerar ESS y ESS-IM con dos islas, dado que la diferencia, a favor de ESS, es menor al 1%. Es importante tener en cuenta, que todos los experimentos utilizan la misma cantidad de recursos computacionales. Debido a esto, al incrementar la cantidad de islas y reducir la cantidad de individuos por islas, necesariamente también se disminuye la cantidad de workers por isla. Esto ocasiona que aumente la cantidad de operaciones que debe llevar a cabo el método (al existir más islas), pero a su vez cada isla cuenta con menos recursos disponibles para desarrollar las tareas, por lo tanto, esto produce demoras en la resolución de las mismas.

Tabla 5.6: Tiempo de ejecución promedio obtenido por cada método (ESS y ESS IM).

Experimento	ESS	ESS-IM (2 islas)	ESS-IM (4 islas)	ESS-IM (6 islas)
Caso 1	00:48:45	00:50:10	00:51:37	00:54:31
Caso 2	00:56:40	00:57:20	00:59:07	01:02:41
Caso 3	00:59:45	01:01:15	01:03:37	01:08:21
Caso 4	00:56:55	00:58:35	01:00:36	01:04:38
Caso 5	01:16:04	01:18:06	01:19:30	01:22:18

Es importante tener en cuenta que la ventaja de operar con islas radica en poder contar con múltiples procesos de búsqueda que operan en simultáneo en regiones diferentes del espacio de soluciones. Para ello es necesario realizar un balance acorde a la cantidad de recursos computacionales y los parámetros de configuración del método, para cumplir de forma adecuada con la carga de trabajo requerida. Finalmente, la experimentación se llevó a cabo en un cluster de computadoras Linux (22 nodos con procesadores Intel-Q9550, 4G de memoria RAM y Gigabit Ethernet), bajo un entorno MPI [156].

5.2.2. Comparación B): ESS y ESS-IM

En esta sección se presentan los resultados obtenidos tras efectuar la comparación de calidad de predicción y rendimiento computacional entre ESS y ESS-IM, luego de incluirse, en este último, diferentes alternativas de operación en la etapa **EC Monitor** del método (ver sección 4.3.2.2).

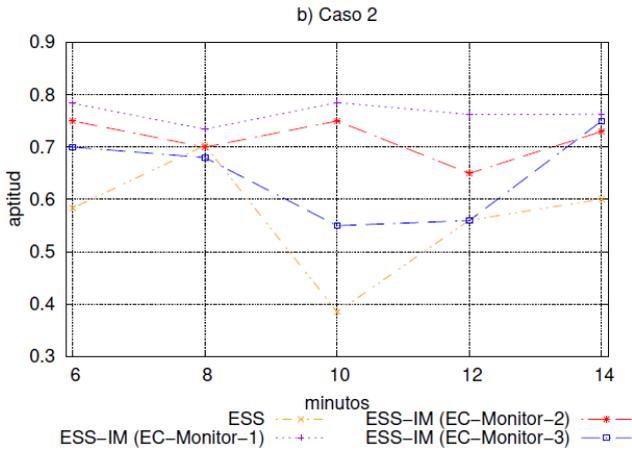
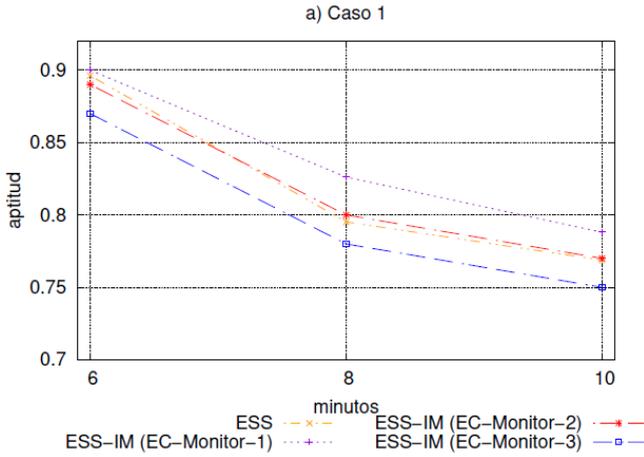
5.2.2.1. Evaluación de Calidad de Predicción

La experimentación realizada se llevó a cabo aplicando los métodos en los casos de estudio 1, 2, 3 y 4 correspondientes a la Tabla 5.1. Como mencionamos anteriormente ESS y ESS-IM centran su funcionamiento en algoritmos evolutivos paralelos, lo que implica que ambos métodos posean comportamiento no determinístico debido a que los individuos de cada población son generados mediante la utilización de semillas calculadas de forma estocástica. Entonces, para efectuar una comparación equitativa entre ambos métodos, se utilizaron 30 semillas diferentes entre sí, pero en conjunto idénticas para cada experimento. Es decir, los cuatro casos de quemas controladas fueron ejecutados en 30 instancias cada uno por cada método, haciendo uso del mismo conjunto de 30 semillas. Los experimentos fueron lanzados de esta manera con el objetivo de poder determinar una media de los resultados para cada experimento y cada método.

Los resultados obtenidos, en términos de calidad de predicción, para cada caso de estudio se observan en las Figs. 5.11 para los experimentos 1, 2, 3 y 4 respectivamente. En términos generales se puede observar que ESS-IM, principalmente en la modalidad EC-Monitor-1, supera a los resultados proporcionados por ESS, salvo algunas excepciones: en el caso 1 (Fig. 5.11.a), donde ESS supera en el primer paso de predicción a ESS-IM tanto en EC-Monitor-2 como en EC-Monitor-3. En el segundo y último paso de predicción sólo supera a EC-Monitor-3.

En el segundo experimento (Fig. 5.11.b), ESS sólo obtiene mejor calidad de predicción que ESS-IM en la modalidad EC-Monitor-3 (en el segundo paso de predicción). En el tercer experimento (Fig. 5.11.c), ESS supera a ESS-IM (en modo EC-Monitor-3) durante todo el experimento. En el minuto 10 obtiene el mismo desempeño que EC-Monitor-1 y el segundo mejor desempeño en el último paso de predicción.

Y finalmente en el cuarto experimento (Fig. 5.11.d), en el minuto 12,5, ESS obtiene mejor desempeño que todas las versiones de ESS-IM. Pero si consideramos la comparación general para el mismo experimento, se puede observar que en los primeros dos pasos de predicción (minuto 7,5 y minuto 10), ESS-IM (EC-Monitor 1 y EC-Monitor-2) han obtenido mejores resultados que ESS. Sin embargo, estos valores disminuyen en el último paso (12,5 min). Por su parte, ESS-IM (EC-Monitor 3) obtiene un muy buen resultado en el primer paso. En el segundo paso, su calidad de predicción disminuye tomando un valor cercano al de ESS, y en el último paso alcanza el menor desempeño de todo el experimento.



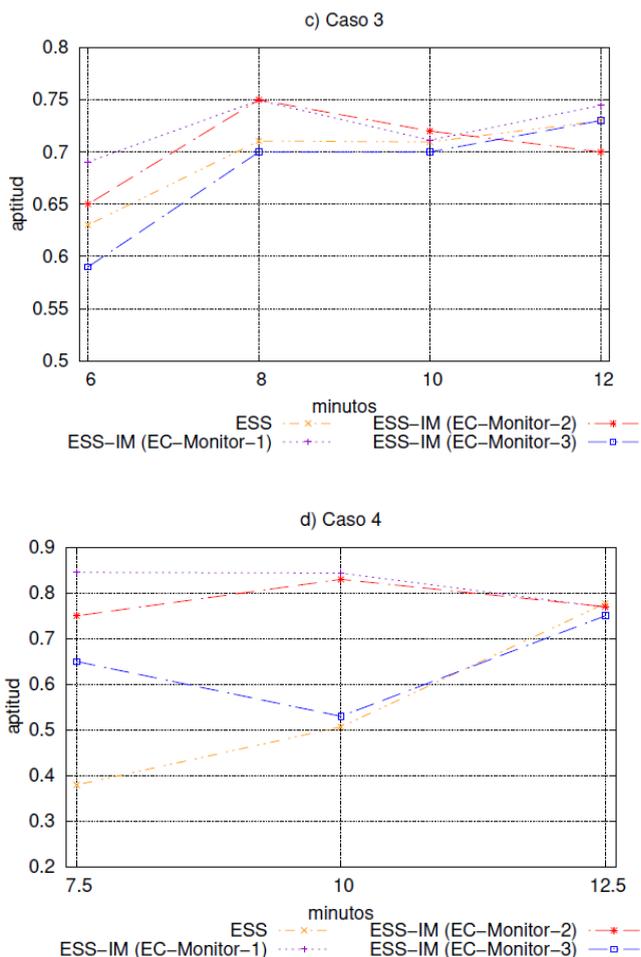


Figura 5.11: Comparación de calidad de predicción entre métodos: ESS y ESS-IM. Casos de quemas controladas 1, 2, 3 y 4.

Tal como se puede observar, en términos generales, el modo de operación EC Monitor-1 de ESS-IM permite generar los mejores niveles de calidad de predicción del método en los casos evaluados. Es importante recordar que en esta modalidad ESS-IM utiliza el mejor valor clave de ignición hallado por todas las islas (K_{ign}). También se aprecia en esta experimentación, que la modalidad EC-Monitor-2 la cual calcula un K_{ign} a diferencia de utilizar un valor previamente obtenido por alguna isla en las etapas anteriores de ejecución, genera resultados próximos a EC-Monitor-1, e incluso en ocasiones puntuales supera a EC-Monitor-1. Y finalmente, EC-Monitor-3, modalidad que

prioriza el rendimiento al utilizar el valor K_{ign} generado, en cada paso de simulación, en el menor tiempo posible obtiene resultados inferiores (dentro de los generados por ESS-IM), pero competitivos en relación a ESS y más aún en tiempo de respuesta.

5.2.2.2. Evaluación de Rendimiento

Respecto a la evaluación de rendimiento de ambos métodos, éstos han sido analizados mediante la métrica conocida como *speed-up* [167], que se define como la relación entre el tiempo necesario para resolver un problema utilizando una sola unidad de procesamiento y el tiempo requerido para resolver el mismo problema en un entorno paralelo/distribuido con x elementos idénticos de procesamiento. En (5.1) se puede observar la manera de calcular el *speed-up*, donde $t(1)$ representa al tiempo de ejecución secuencial y $t(N)$ el tiempo de ejecución en paralelo con N unidades de procesamiento. Debido a la arquitectura de ESS y ESS-IM, cada uno de ellos ha sido diseñado para operar mínimamente con 2 y 3 procesos, respectivamente. En ESS se necesita de un proceso master y un worker, y en ESS-IM es necesario un monitor, un master y un worker. Debido a esto ha sido necesario adaptar la implementación de ambos métodos para poder operar con menor cantidad de procesos, es decir con 1 y 2 procesos.

$$S = \frac{t(1)}{t(N)} \quad (5.1)$$

En la Fig. 5.12 se puede observar el resultado del análisis de *speed-up* realizado tras comparar, en términos de rendimiento, cada una de las metodologías descritas anteriormente. La gráfica corresponde a la utilización de: 1, 2, 4, 8, 16 y 32 unidades de procesamiento en la simulación de un caso particular. Para confeccionar cada una de las curvas de *speed-up* se efectuó la misma cantidad de simulaciones considerando para cada una el valor promedio.

En términos generales se puede observar que todos los métodos evaluados presentan muy buen resultado hasta la utilización de 16 unidades de procesamiento. A partir de aquí, dos modalidades de ESS-IM (EC-Monitor-2 y EC-Monitor-3) continúan con la misma tendencia mostrando un muy buen desempeño, cercano al *speed-up* ideal, en contraste con ESS-IM en modo EC-Monitor-1 y ESS.

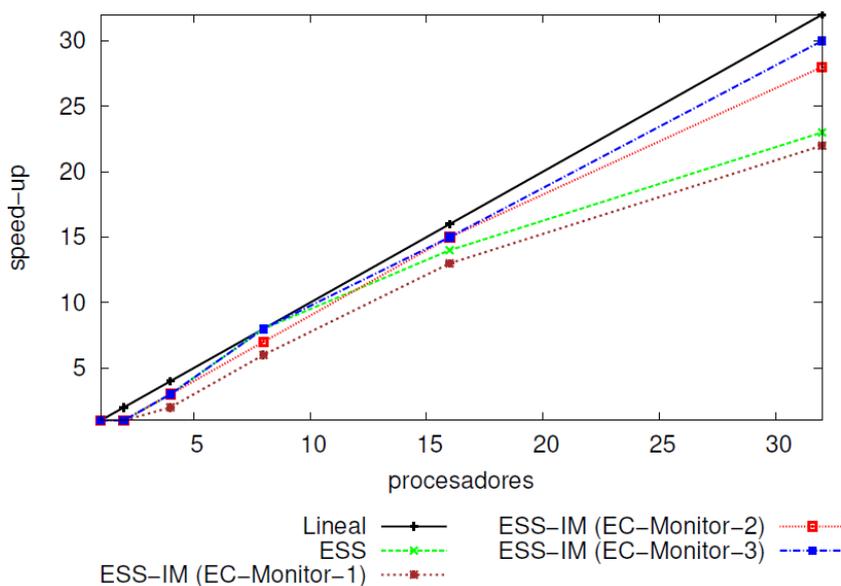


Figura 5.12: Análisis de *speed-up* entre ESS y ESS-IM en modalidades: EC-Monitor 1, EC-Monitor-2, y EC-Monitor-3.

Si bien la ganancia de rendimiento obtenida por estos últimos disminuye al trabajar con 32 procesos, lo cual es un factor importante a tener en cuenta, si se consideran los resultados de calidad de predicción obtenidos por estos métodos, se puede decir que ESS-IM en modo EC-Monitor-1 y ESS ofrecen un “muy buen” rendimiento, ya que los niveles de calidad de predicción son superiores a los demás y al utilizar mayor cantidad de unidades de procesamiento se ha logrado disminuir el tiempo de ejecución. Con el propósito de mejorar la escalabilidad de estos métodos, haciendo un uso más eficiente de los recursos, y con el objetivo de obtener mayor ganancia de rendimiento se llevó a cabo otra investigación correspondiente a la incorporación de arquitecturas GPU.

Tal como se ha podido observar en los resultados presentados en esta sección, la configuración de la etapa de calibración del proceso monitor es crucial ya que impacta directamente sobre los resultados. Tanto en nivel de calidad de predicción como en rendimiento, se esperaba que los mejores resultados fueran obtenidos usando las modalidades EC-Monitor-1 y EC-Monitor-3 respectivamente, lo cual ha podido verificarse en los resultados presentados. En el caso de la calidad de predicción, el mejor valor de K_{ign} garantiza aproximar de mejor manera el comportamiento de los resultados simulados con la predicción, y en términos de tiempo/rendimiento, al tomar en valor obtenido en menor

tiempo se garantiza concluir con la predicción en menor tiempo que el resto de las modalidades.

Debido a que el análisis de *speed-up* no se ha llevado a cabo en todas las experimentaciones de esta tesis, se ha incluido una tabla comparativa de consumo de tiempo, al igual que en la sección 5.2.1.2. Dicha tabla (Tabla 5.7) presenta el tiempo total incurrido por ESS y ESS-IM en sus tres modalidades: EC-Monitor-1, EC-Monitor-2 y EC-Monitor-3. Tal como se puede observar, el mejor rendimiento es ofrecido por ESS-IM en la modalidad EC-Monitor-3, como se había adelantado previamente. Esto se logra debido a que ESS-IM en esta modalidad no necesita esperar a que todas las islas finalicen cada paso de simulación, es decir, avanza al momento que se produzca la finalización de al menos una isla. Esto permite que el proceso de predicción, paso a paso, avance de manera más rápida. Posteriormente se encuentra la modalidad EC-Monitor-2, y finalmente se encuentra la modalidad EC-Monitor-1, la cual se encuentra por debajo de ESS. Es importante destacar que en esta experimentación se utilizó la misma cantidad de recursos en todos los experimentos, utilizando un balance en cantidad de individuos en la población entre los ESS y ESS-IM. Al igual que el caso anterior, la experimentación se llevó a cabo en un cluster de computadoras Linux (22 nodos con procesadores Intel-Q9550, 4G de memoria RAM y Gigabit Ethernet), bajo un entorno MPI [156].

Tabla 5.7: Tiempo de ejecución promedio obtenido por cada método (ESS y ESS-IM: ESS-IM(1) EC-Monitor-1, ESS-IM(2) EC-Monitor-2 y ESS-IM(3) EC-Monitor-3).

Experimento	ESS	ESS-IM (1)	ESS-IM (2)	ESS-IM (3)
Caso 1	00:48:45	00:52:12	00:46:01	00:44:12
Caso 2	00:56:40	00:58:23	00:51:04	00:44:22
Caso 3	00:59:45	01:02:49	00:56:04	00:52:00
Caso 4	00:56:55	00:59:33	00:53:19	00:50:53

5.2.3. Comparación C): ESS-IM y HESS-IM

En esta sección se presentan, en primer lugar, los resultados obtenidos tras evaluar la calidad de predicción de una versión preliminar de HESS-IM en contraste con la metodología previa ESS-IM. Esta implementación de HESS-IM utiliza una técnica de optimización híbrida y colaborativa entre las metaheurísticas: a) Algoritmos

Evolutivos y b) Evolución Diferencial. Para ello, ambos métodos fueron aplicados a tres quemas controladas, caso 1, 3 y 5 de la Tabla 5.1.

Para ambos métodos el criterio de finalización de cada metaheurística consistió en alcanzar un umbral de aptitud igual a 0,7, o llegar a un máximo de 200 iteraciones. Además, el tamaño de la población para cada metaheurística fue establecido en 200 soluciones. Para los parámetros particulares de cada metaheurística se utilizó una configuración clásica de valores, es decir, una configuración basada en valores de referencia de la literatura y de experimentaciones previamente realizadas. La información detallada de los parámetros de configuración se encuentra listada en la Tabla 5.8.

Tabla 5.8: Configuración de parámetros HESS-IM, EA y DE.

Parámetro	Valor
Umbral de fitness	0,7
Iteraciones	200
Individuos por isla	200
Islas	5
Individuos a migrar	20,00%
Criterio	semi-elitista
Frecuencia	cada iteración
Workers por isla	7
(EA) Criadores	0,8
(EA) Mínima prob. cruzamiento	0,2
(EA) Máxima prob. cruzamiento	0,6
(EA) Mutación	0,5
(DE) f, Factor de mutación	0,9
(DE) Cr, Cruzamiento	0,3

5.2.3.1. Evaluación de Calidad de Predicción

Experimentación ESS-IM y HESS-IM (EA y DE):

En la Fig. 5.13 y las Tablas 5.9, 5.10 y 5.11 se presentan los resultados de calidad de predicción tras aplicar cada método a los casos de estudio antes descritos. En términos generales se observa un mejor desempeño por parte del método HESS-IM, sin embargo, en algunos instantes puntuales de tiempo se ve superado por ESS-IM.

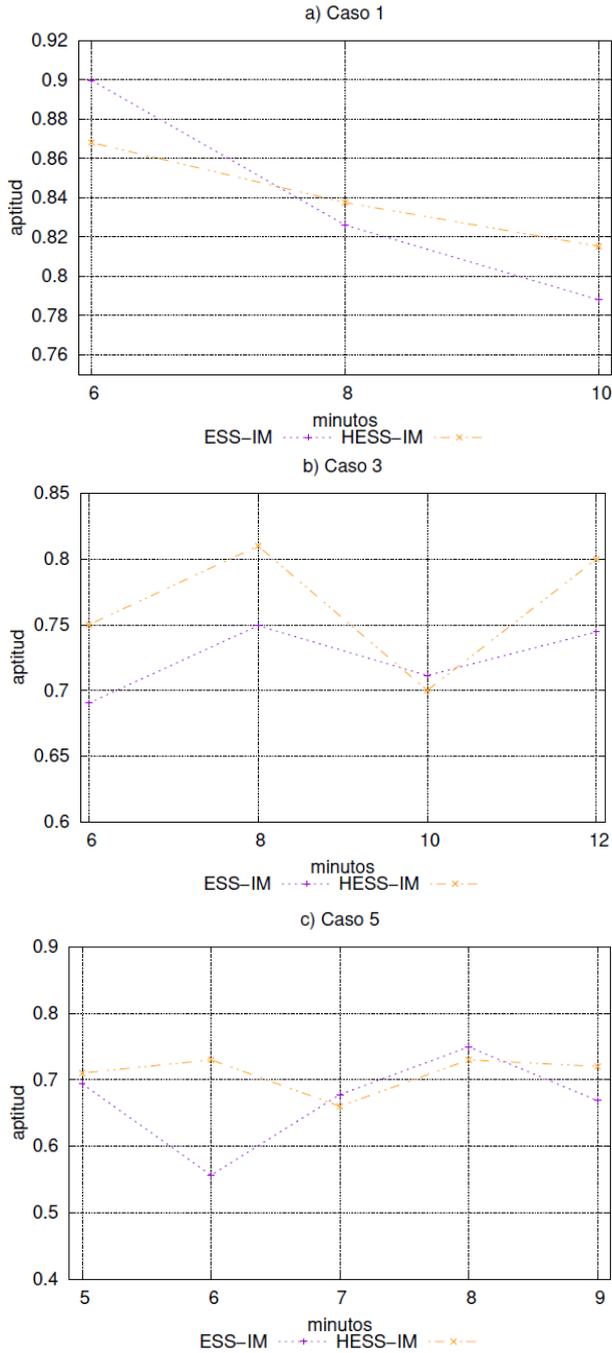


Figura 5.13: Comparación de calidad de predicción entre métodos: ESS-IM y HESS IM (EA y DE). Casos 1, 3, y 5.

En el caso 1, tal como se puede observar en la Fig. 5.13.a, ESS-IM obtiene la mejor calidad de predicción en el primer instante de tiempo, en los siguientes instantes (minutos 8 y 10), HESS-IM supera a ESS-IM. En las tablas se destaca en negrita los mejores valores de cada paso de predicción por cada instante de tiempo, además, se muestra el promedio general que cada método ha obtenido. También se ha marcado con un asterisco (*), el mejor registro de calidad de predicción de todo el experimento (ver para el caso 1 la Tabla 5.9). De esta manera, el mejor valor de calidad de predicción (en instantes de tiempo independientes) ha sido obtenido por ESS-IM en el minuto 6, y en promedio general se observa que HESS-IM ofrece mejor desempeño.

Tabla 5.9: Caso 1: Calidad de predicción y promedio general.

Intervalo (min)	ESS-IM	HESS-IM
2 a 4	-	-
4 a 6	0,899909*	0,8681990
6 a 8	0,8261334	0,8376037
8 a 10	0,7880610	0,8153010
Promedio	0,8380344	0,8403679

En el caso 3 (ver Fig. 5.13.b), HESS-IM obtiene un rendimiento bastante superior sobre ESS-IM, en contraste con el caso anterior. Como se puede observar, HESS-IM obtiene la mejor calidad de predicción en tres de los cuatro pasos de predicción (minuto 6, 8 y 12, Tabla 5.10). El mejor registro de calidad de predicción de todo el experimento ha sido obtenido por HESS-IM en el minuto 8 (en el segundo paso de predicción), y en promedio general también se observa que HESS-IM es el método con mejor calidad alcanzada. En el minuto 10, corresponde al instante donde HESS IM presenta la menor calidad de predicción, y justamente es donde ESS supera levemente a ESS-IM.

Tabla 5.10: Caso 3: Calidad de predicción y promedio general.

Intervalo (min)	ESS-IM	HESS-IM
2 a 4	-	-
4 a 6	0,69041180	0,75054000
6 a 8	0,74937500	0,8102500*
8 a 10	0,71142720	0,70540000
10 a 12	0,74456900	0,80584000
Promedio	0,72394575	0,76800750

En el caso 5 (Fig. 5.13.c y Tabla 5.11) ESS-IM obtiene el mejor registro de predicción de todo el experimento (minuto 8), también la mejor calidad, en dos de los cinco instantes de tiempo. HESS-IM obtiene mejor calidad de predicción en 3 de 5 pasos de predicción, con un valor de calidad promedio bastante superior al de ESS-IM.

Tabla 5.11: Caso 5: Calidad de predicción y promedio general.

Intervalo (min)	ESS-IM	HESS-IM
3 a 4	-	-
4 a 5	0,69380440	0,71058000
5 a 6	0,55625060	0,73046800
6 a 7	0,67776860	0,66025000
7 a 8	0,7494070*	0,73036000
8 a 9	0,66837460	0,71985000
Promedio	0,66912104	0,71030160

Tal como se ha podido observar, los resultados obtenidos indican que el enfoque colaborativo entre ambas metaheurísticas permite obtener mejoras en el mecanismo de optimización global del método, esto se ve reflejado en el incremento de la calidad de predicción en más del 50 % de los resultados. Al igual que en el resto de los experimentos (dado que ambos métodos poseen comportamiento estocástico), los resultados obtenidos fueron generados a partir del promedio de 30 ejecuciones (utilizando el mismo conjunto de semillas para ambos métodos). Además, todas las pruebas fueron desarrolladas en un clúster Linux, con 44 núcleos de cómputo distribuidos en CPUs quad-core, con 4 GB de RAM, bajo un entorno MPI y una red Gigabit Ethernet.

Experimentación ESS-IM y HESS-IM (EA, DE y PSO):

Al incorporarse la metaheurística Optimización por Cúmulo de Partículas (PSO) a HESS-IM se llevó a cabo una nueva experimentación, nuevamente entre ESS-IM y HESS-IM, es decir, utilizando una metaheurística híbrida basada en a) Algoritmo Evolutivo, b) Evolución Diferencial y c) Optimización por Cúmulo de Partículas. Con el propósito de determinar el impacto de incluir PSO a la hibridación, la comparación también incluye la implementación de HESS-IM con EA y DE. Por este motivo en las gráficas siguientes se observan tres curvas, una para ESS-IM, otra para HESS-IM con EA y DE, nombrada como HESSIM (a), y la tercera para HESS-IM con EA,

DE y PSO, denominada HESS-IM (b). La experimentación se llevó a cabo utilizando los casos de quemas controladas 1, 3 y 5 de la Tabla 5.1 y también se aplicó a los casos 1 y 2 de la Tabla 5.2 (es importante recordar que estos últimos corresponden a incendios forestales de gran escala, donde el nivel de complejidad crece de manera drástica). Para todos los experimentos se utilizó la misma configuración de parámetros de inicialización del método, tales parámetros se listan en la Tabla 5.12. En la Fig. 5.15 se presentan los resultados de calidad de predicción para los casos de quemas controladas. Donde se puede observar que la inclusión de PSO como componente adicional a la metaheurística híbrida le permite al método mejorar considerablemente la calidad de predicción, obteniendo el mejor desempeño en la totalidad de experimentos y pasos de predicción.

Tabla 5.12: Configuración de parámetros HESS-IM, EA, DE y PSO.

Parámetro	Valor
Umbral de fitness	0,7
Iteraciones	200
Individuos por isla	200
Islas	5
Individuos a migrar	20,00%
Criterio	semi-elitista
Frecuencia	1 iteración
Workers por isla	7
(EA) Criadores	0,8
(EA) Mín. cruzamiento	0,2
(EA) Máx. cruzamiento	0,6
(EA) Mutación	0,5
(DE) f, Factor de mutación	0,9
(DE) Cr, Cruzamiento	0,3
(PSO) c1 y c2, Coeficientes de aceleración	0,5
(PSO) w, Factor inercial	0,6
(PSO) Tamaño del vecindario	incremental

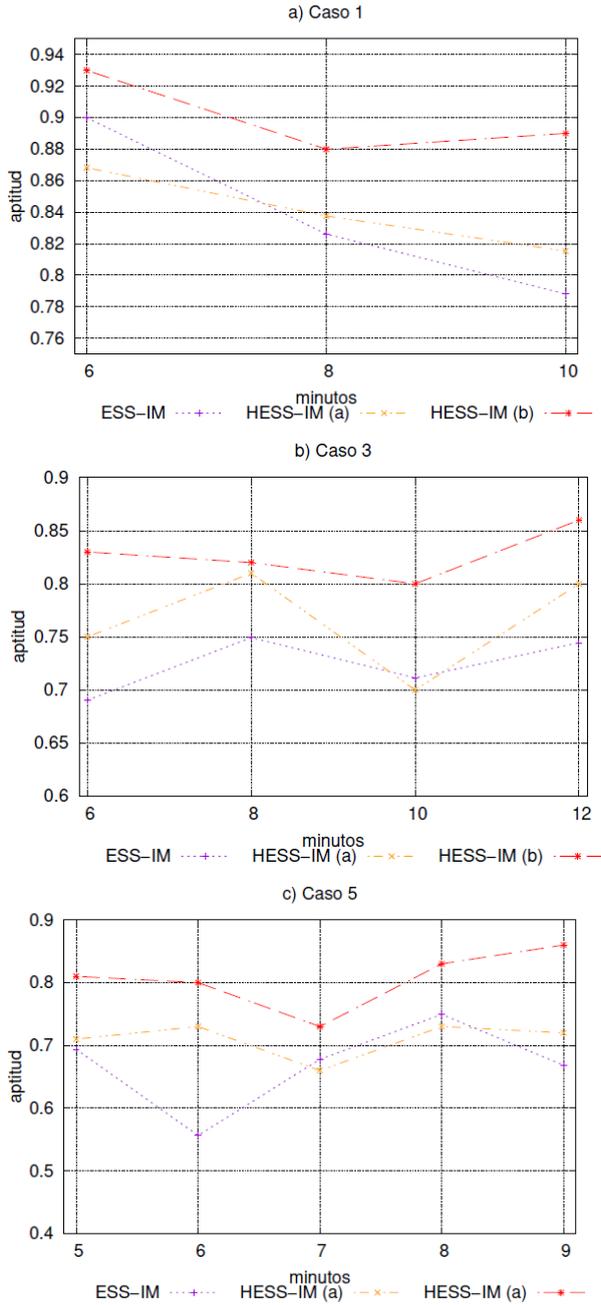


Figura 5.15: Comparación entre ESS-IM, HESS-IM (a: EA y DE), HESS-IM (b: EA, DE y PSO): Casos de quemas controladas 1, 3 y 5.

En las Fig. 5.16 se presentan los resultados de calidad de predicción para los casos de estudio correspondientes a incendios forestales (casos 1 y 2 de la Tabla 5.2). Al igual que los experimentos anteriores, se grafican tres líneas (correspondientes a ESS IM, HESS-IM (a), y HESS-IM (b), es decir HESS-IM con EA y DE, y HESS-IM con EA, DE y PSO, respectivamente). Los valores del tiempo corresponden a intervalos de 1440 minutos, equivalentes a un día. En el caso 1 (Fig. 5.16.a), se puede observar que la calidad de predicción de HESS-IM (b) supera completamente al resto de las evaluaciones.

Es importante tener en cuenta que, en incendios de grandes dimensiones, aumenta la dificultad para obtener valores de calidad de predicción del orden de los casos correspondientes a quemas controladas. Esto se debe tanto por la dimensión de las áreas bajo estudio, la resolución con la que se aborda el fenómeno, la falta de incorporación de otras variables que deberían ser consideradas, como: ríos, montañas, grandes desniveles, áreas sin combustible, zonas urbanas, entre otras, y además debido a la baja periodicidad de calibración con los frentes de fuego reales (LFRs).

En la Tabla 5.14, se puede observar que el mejor valor de predicción se obtiene en el primer instante de predicción (valor 0,740158) para HESS-IM (b). Otro aspecto a considerar es que la calidad de predicción, va decreciendo desde el comienzo de la experimentación, esto puede tener relación con la complejidad de la morfología del área quemada para cada instante de tiempo. Se puede observar en la Fig. 5.14.a, que el área quemada en los minutos 2.880 y 4.320, donde los valores de

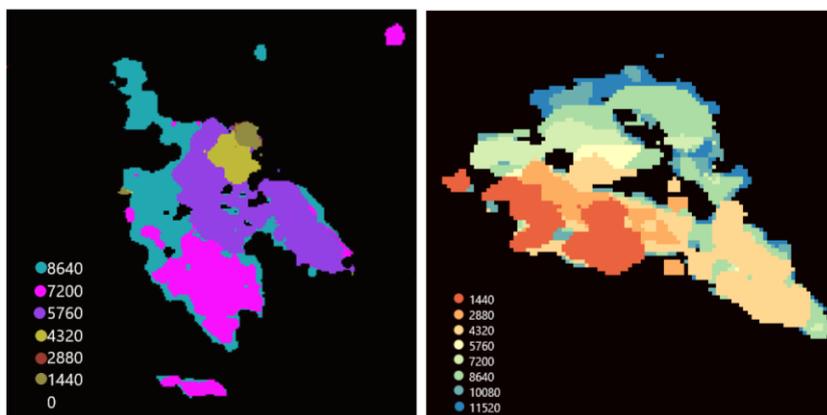


Figura 5.14: Mapas finales del desarrollo del incendio: a la izquierda, a) Caso de incendio forestal 1, y a la derecha, b) Caso de incendio forestal 2.

predicción son más elevados, la forma del área quemada presenta mayor homogeneidad que en los pasos posteriores. En el caso 2, en el desempeño de calidad de predicción, se pueden observar características similares al caso anterior (ver Fig. 5.16.b). Por un lado, al igual que en el caso 1, HESS-IM (b) obtiene el mejor rendimiento en todos los pasos de predicción, con un valor máximo alcanzado equivalente a 0,591102 (ver Tabla 5.15), y por el otro, si se contrasta con la Fig. 5.14.b, se puede observar que los mayores niveles de calidad de predicción coinciden con momentos del fenómeno donde el área quemada posee una morfología un tanto más uniforme en contraste con el resto de los instantes de tiempo.

5.2.3.2. Evaluación de Rendimiento

Experimentación ESS-IM y HESS-IM (EA y DE):

En la Tabla 5.13 se presenta el tiempo de ejecución promedio de la experimentación comparativa entre ESS-IM y HESS-IM (EA y DE). Tal como se puede observar, ESS-IM presenta mejor rendimiento, con una diferencia entre 2 y 4 minutos.

Es importante tener en cuenta que en esta experimentación se utilizó para ambos métodos una cantidad de 6 islas, en donde ESS-IM al igual HESS-IM operan en la modalidad EC-Monitor-1, es decir, utilizan el mejor valor K_{ign} obtenido en todas las islas.

Tabla 5.13: Tiempo de ejecución promedio obtenido por cada método ESS-IM y HESS-IM (EA y DE).

Experimento	ESS-IM	HESS-IM (EA y DE)
Caso 1	00:54:31	00:56:22
Caso 3	01:08:21	01:12:21
Caso 5	01:22:18	01:24:46

ESS-IM y HESS-IM (EA, DE y PSO):

Con el propósito de mejorar el rendimiento de HESS-IM se llevó a cabo una implementación del método basada en CPU-GPU. Para ello se realizó una adaptación del trabajo de [158], el cual consiste en una implementación del simulador fireLib para GPU. La adaptación fue necesaria para poder ser introducido en la etapa de Simulación de Incendios (SI) del método. De esta manera, se añade en cada isla un nodo con capacidad de GPU, el cual opera procesando soluciones

candidatas a demanda. En esta etapa de desarrollo del trabajo el procesamiento GPU es utilizado únicamente en la etapa SI, es decir en evaluación de los individuos.

Con el objetivo de cuantificar la ganancia obtenida con esta implementación, durante las experimentaciones previamente presentadas se ha registrado el tiempo de cómputo de cada uno de los casos. En la Tabla 5.16 se presenta el promedio de tiempo de ejecución para cada experimento. La primera columna indica el experimento, considerando los 5 casos de quemas controladas y los 2 experimentos de incendios reales. Las tres filas siguientes presentan los resultados de tiempo medidos en horas, minutos y segundos del promedio de ejecución para cada método. Tanto para ESS, ESS-IM, HESS-IM con su implementación sólo para CPU y los resultados de HESS IM para CPU-GPU.

Tabla 5.14: Caso 1 real: Calidad de predicción y promedio general.

Intervalo (min)	ESS-IM	HESS-IM (a)	HESS-IM (b)
0 a 1440	-	-	-
1440 a 2880	0,2405400	0,6305840	0,740158*
2880 a 4320	0,6158400	0,5605800	0,6332150
4320 a 5760	0,3198060	0,3658820	0,5054600
5760 a 7200	0,4015890	0,4201500	0,4602350
7200 a 8640	0,2398500	0,3036500	0,4023500
Promedio	0,3635250	0,4561692	0,5003150

Los resultados presentados claramente evidencian una pérdida de rendimiento entre ESS-IM y HESS-IM(CPU), a favor del primero, esto se debe a un posible desacoplamiento de velocidades producido entre las diferentes islas, debido a la utilización de diferentes metaheurísticas. Es importante tener en cuenta que los parámetros de configuración de HESS-IM no han sido calibrados, se estima que mediante un análisis de calibración estática sea posible mejorar el rendimiento de HESS-IM, tanto en calidad de predicción como así también en tiempo de ejecución. En la implementación CPU-GPU de HESS-IM se puede observar un incremento de rendimiento muy importante del orden del 60% en prácticamente la totalidad de los experimentos. Esto es posible mediante la ganancia de rendimiento obtenida tras realizar gran parte de las evaluaciones de individuos en la etapa SI sobre las GPUs. Para los experimentos se utilizaron cuatro GPUs: dos NVIDIA Tesla C2075 y dos NVIDIA GeForce GTX 480.

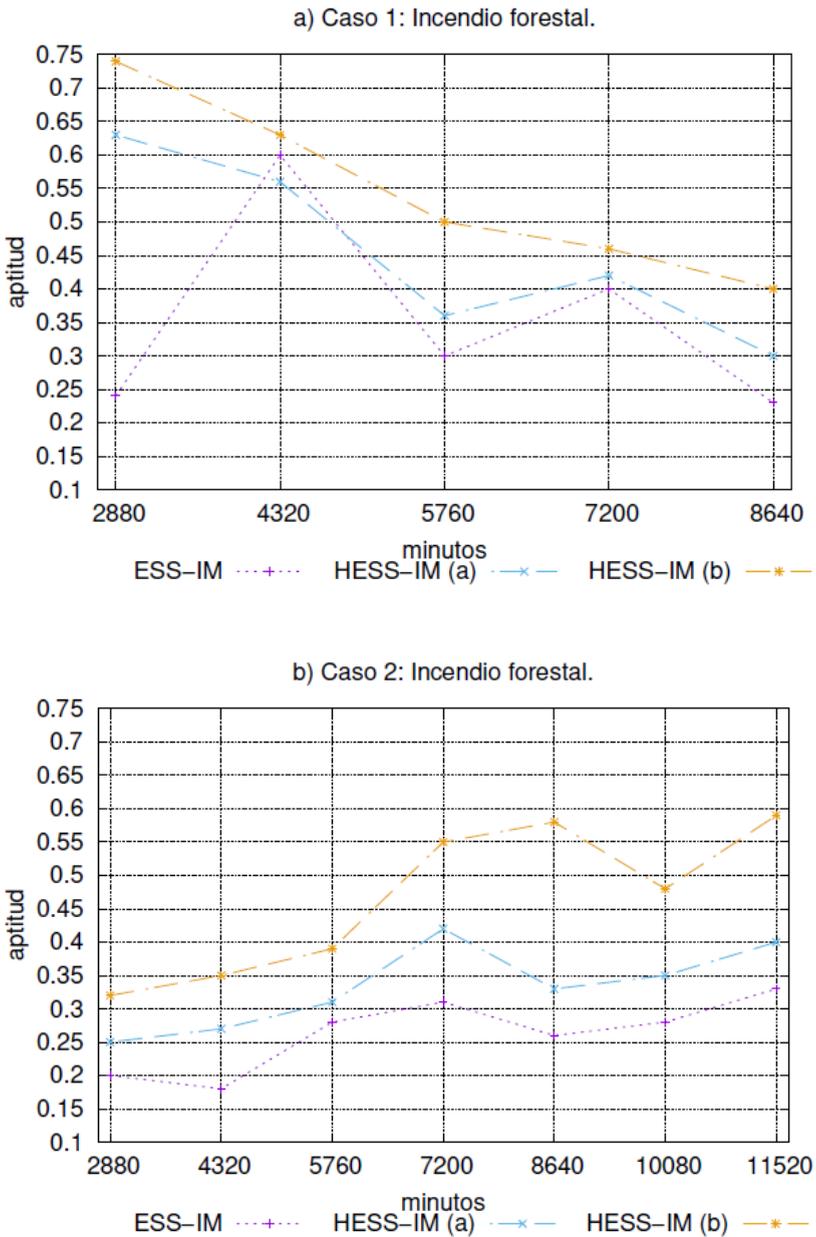


Figura 5.16: Comparación entre ESS-IM, HESS-IM (a: EA y DE), HESS-IM (b: EA, DE y PSO): Casos de incendios forestales 1 y 2.

Es importante mencionar que si bien el equipamiento sobre el cual se llevó a cabo el conjunto de experimentos es muy limitado, ha permitido demostrar que los métodos desarrollados, principalmente el de HESS-IM en su implementación CPU GPU, puede dar un paso adelante hacia los sistemas de *urgent computing*. Esto se debe a que, además de ofrecer buenos resultados de calidad de predicción, obtiene buen desempeño en término de tiempos de procesamiento. Ya que para predecir un único instante de tiempo de un incendio forestales en condiciones reales sólo necesita de 7 minutos. Si consideramos que se realizará una única predicción por día, es posible aplicar este método a situaciones reales. Además, si consideramos que este tipo de sistemas, al estar en producción suelen ser implementados sobre plataformas de hardware de última tecnología, seguramente el tiempo de respuesta pueda verse incrementado.

Tabla 5.15: Caso 2 real: Calidad de predicción y promedio general.

Intervalo (min)	ESS-IM	HESS-IM (a)	HESS-IM (b)
0 a 1440	-	-	-
1440 a 2880	0,21028000	0,25088700	0,32110200
2880 a 4320	0,18025000	0,27100000	0,35330210
4320 a 5760	0,28468000	0,31025700	0,39111400
5760 a 7200	0,31045300	0,42253600	0,55221000
7200 a 8640	0,26005800	0,33153000	0,58203300
8640 a 10080	0,28025000	0,35032190	0,48193700
10080 a 11520	0,33170300	0,41143000	0,5911025*
Promedio	0,26538200	0,31724200	0,43995222

Tabla 5.16: Tiempo de ejecución promedio obtenido por cada método.

Experi- mento	Promedio de tiempo de ejecución				Ganancia porcentual
	ESS	ESS-IM	HESS-IM (CPU)	HESS-IM (CPU/GPU)	
Caso1	00:48:45	00:50:10	00:55:20	00:20:02	63,80%
Caso2	00:56:40	00:57:20	00:59:12	00:23:12	60,81%
Caso3	00:59:45	01:01:15	01:10:00	00:27:01	61,40%
Caso4	00:56:55	00:58:35	01:05:25	00:25:05	61,66%
Caso5	01:16:04	01:18:06	01:22:15	00:33:06	59,76%
Caso1 real	-	01:50:00	02:05:33	00:50:11	60,03%
Caso2 real	-	01:59:28	02:17:21	00:53:20	61,17%

5.3. Síntesis

La hibridación implementada en HESS-IM utilizando tres metaheurísticas poblacionales evolutivas ha permitido mejorar el optimizador del método incrementando la calidad de predicción del método. La hibridación se ve mejorada utilizando un esquema de migración basado en topología de anillo entre diferentes metaheurísticas utilizando una configuración de parámetros clásica. La metodología es validada aplicando la misma sobre casos de quemas controladas como así también en casos de incendios forestales. La implementación CPU-GPU del método ha permitido mejorar también el rendimiento del método en términos de tiempo de ejecución, siendo este un factor de suma importancia en los sistemas abocados a resolver problemáticas con gran impacto social.

Calibración de Métodos

En esta sección se presentan los estudios de calibración más significativos realizados durante la presente tesis con el propósito de incrementar el rendimiento de los métodos mediante la elección adecuada de parámetros de configuración de las metaheurísticas intervinientes. Puntualmente se trata de estudios de calibración y/o sintonización estática de los métodos.

6.1. Introducción

Los AEP pueden obtener diferentes resultados, en función de la configuración de parámetros evolutivos utilizada [168], haciendo que operen de mejor o peor manera en función de dicha configuración. Esto impacta tanto en términos de la calidad de las soluciones encontradas, como así también en relación a la cantidad de tiempo y recursos utilizados.

6.1.1. Tipos de calibración

Para obtener una configuración óptima de estos parámetros, es necesario someter el algoritmo a un proceso de optimización, análisis y/o ajuste [66]. Dicho proceso de optimización se conoce como *calibración*, *sintonización* o *análisis de parámetros*. Debido a que las metaheurísticas son aplicadas a diferentes tipos de problemas, con características disímiles, es difícil definir una configuración genérica de parámetros para cada metaheurística. Según [66], dos son las técnicas más difundidas para realizar un análisis de calibración de parámetros:

a) Calibración de parámetros Off-line:

Esta técnica consiste en definir diferentes configuraciones de parámetros para ser ejecutadas de forma independiente. Los valores de cada parámetro son establecidos antes de ejecutar el algoritmo y se mantienen estáticos durante toda la ejecución. Cuando todas las experimentaciones fueron llevadas a cabo, se analizan los resultados obtenidos y se identifica la configuración de parámetros que permite

obtener mejores resultados. Uno de los principales inconvenientes de este tipo de enfoques se relaciona a su alto costo computacional. Cuando el estudio incluye una reducida cantidad de experimentos o cuando se trata de estudios de sondeo, los experimentos pueden ser realizados mediante un Diseño por Hipercubo Latino (Latin Hypercube Design, LHD) [169], diseño fraccional [170], entre otros. Sin embargo, existen técnicas de calibración estática que consisten en plantear la búsqueda de parámetros como un problema de optimización, esto se conoce como metaoptimización, algunos ejemplos son irace [171], ParamILS [172] y SMAC [173].

b) Calibración de parámetros On-line:

En la calibración on-line los valores de cada parámetro son determinados de manera dinámica en tiempo de ejecución. Este enfoque, permite adaptar los parámetros para mejorar la efectividad de una configuración de parámetros que pueda cambiar en diferentes momentos de la búsqueda.

La implementación de cada una de estas técnicas posee diferentes niveles de complejidad, ventajas y desventajas que deben ser analizadas en detalle, junto con el tipo de problema a tratar. Cabe destacar que, debido al tipo de fenómeno tratado, la complejidad de la arquitectura interna del método, el grado de utilización de recursos necesario y la cantidad de parámetros involucrados, los análisis de calibración de parámetros realizados en el presente trabajo son del tipo Off-line (o estático) con propósito de sondeo basado en LHDs.

6.2. Calibración de ESS-IM

En esta sección se presentan dos tipos de análisis de calibración llevadas a cabo en ESS-IM. En el primero, el estudio se realizó utilizando una configuración homogénea de parámetros, es decir, se analizaron diferentes configuraciones de parámetros, instanciando todas las islas con la misma configuración. En el segundo análisis, se realizó un estudio de calibración utilizando una configuración heterogénea de parámetros, es decir, utilizando una lista de buenas configuraciones e instanciando cada isla con una configuración diferente (escogiendo entre las mejores configuraciones de parámetros).

6.2.1. Parámetros de los AEP

En primer lugar, es necesario determinar los parámetros que serán sometidos al análisis de calibración y qué rangos de valores serán considerados para cada uno de ellos. A continuación, se listan los parámetros más significativos de todo AEP, es importante tener en cuenta que no se listan aquellos parámetros específicos de cada metaheurística. Además, es importante mencionar que no todos los parámetros han sido calibrados, igualmente dichos parámetros se describen para dejar en claro los valores que fueron utilizados en dichos casos:

a) Umbral de fitness:

Determina el valor de aptitud mínimo que debe ser obtenido por al menos un individuo. Un valor bajo implica menor tiempo para encontrar una la solución, pero con mayor probabilidad de ser un valor distante del óptimo. En cambio, un valor elevado, dificultará la convergencia a una solución, incrementando el tiempo de procesamiento.

b) Cantidad máxima de iteraciones:

Determina la cantidad máxima de generaciones evolutivas que se llevará a cabo, suele ser utilizada como condición de finalización junto con el umbral de fitness³. Pocas iteraciones harán que el algoritmo finalice más rápido, con un pobre nivel de exploración del espacio de búsqueda. Por el contrario, mayor cantidad de iteraciones puede generar pérdidas considerables de tiempo.

c) Cantidad de islas:

Establece la cantidad total de poblaciones o subpoblaciones que participarán del AEP. Mientras más poblaciones sean utilizadas, mejor se explorará el espacio de búsqueda incrementando la cantidad de recursos necesarios, como así también el tiempo de procesamiento.

d) Cantidad de individuos por isla:

Determina la cantidad total de individuos que cada población mantendrá durante toda la ejecución del algoritmo (generalmente suele ser fija). Éste es un parámetro importante que afecta el rendimiento y la eficiencia del AEP, ya que si las poblaciones son muy pequeñas existirá poca diversidad, haciendo que no se puedan alcanzar soluciones de

³ Ambos parámetros (umbral de fitness y cantidad máxima de iteraciones) se utilizan en conjunto como condición de parada del componente de optimización de los métodos presentados.

calidad. En cambio, a mayor tamaño aumentará la diversidad, pero será necesario realizar mayor cantidad de evaluaciones.

e) Cantidad de individuos a migrar:

El proceso de migración consiste en intercambiar soluciones candidatas entre las diferentes islas, realizando la selección de individuos en cada isla. Este parámetro especifica cuántos individuos deben ser copiados hacia las islas vecinas al producirse una migración.

f) Criterio migración:

Este parámetro indica la forma en que serán seleccionados los individuos que migrarán (por ejemplo, si se escogen los mejores individuos, o si la selección se produce de forma aleatoria, o una combinación entre ambas, etc.). También indica qué va a ocurrir con los nuevos individuos que se incorporen de otras poblaciones (por ejemplo, si reemplazarán a los peores, si el reemplazo se realizará de forma aleatoria, etc.).

g) Frecuencia de migración:

Este parámetro establece cuándo ocurre la migración, ésta puede realizarse de manera periódica, de manera determinista o a partir de un evento determinado. Una configuración habitual de este parámetro es en función al número de iteraciones. Es importante tener en cuenta que la migración introduce una sobrecarga de comunicación, por lo que este parámetro se debe determinar cuidadosamente para que la comunicación no domine a la computación.

h) Probabilidad de cruzamiento:

Parámetro que controla el accionar del operador de cruzamiento. Esto lo logra determinando con que probabilidad se realizará el cruzamiento entre dos individuos. Se espera que este operador explote las mejores características de la población actual.

i) Probabilidad de mutación:

Parámetro que define con que probabilidad se mutará cada elemento de la población. La mutación actúa sobre un único individuo. Generalmente, se recomiendan valores pequeños para este parámetro.

j) Topología de comunicación:

Especifica la relación de comunicación entre las diferentes islas para el proceso de migración, una de las estrategias más comunes es la topología de anillo.

6.2.2. Calibración homogénea

En la Tabla 6.1 se listan los parámetros incluidos en el estudio de calibración junto con sus rangos asociados. Los parámetros con valores fijos no han sido calibrados, pero los valores indicados (para tales parámetros) son los utilizados en todos los experimentos. Es importante tener en cuenta que los rangos utilizados para cada parámetro han sido seleccionados en función de la literatura y trabajos previos realizados [174].

Tabla 6.1: Parámetros evolutivos y rangos utilizados.

Parámetro	Fijo / dinámico	Valor / rango
Umbral de fitness	dinámico	0,55 a 0,85
Cantidad máxima de de iteraciones	dinámico	100 a 500
Cantidad de islas	fijo	4
Cantidad de individuos por isla	dinámico	10 a 200
Cantidad de individuos a migrar	dinámico	5 a 20
Criterio de migración	fijo	elitista
Frecuencia de migración	dinámico	1 a 10
Probabilidad de cruzamiento	dinámico	10 al 60%
Probabilidad de mutación	dinámico	0,1 a 2,5%
Topología de comunicación	fijo	anillo

El experimento se desarrolló mediante un Diseño por Hipercubo Latino (es decir, cuántos experimentos serán llevados a cabo y que configuración de parámetros tendrá cada uno de ellos), mediante el software estadístico R [175]. En la Tabla 6.2 se observa cada uno de los experimentos del LHD con sus valores de parámetros. La fila 0 corresponde a la configuración de parámetros clásica, es decir una configuración de parámetros obtenida en experimentaciones previas (no generada por el LHD). El resto de las filas (1 a 20) corresponden a los experimentos del LHD, donde se han marcado con asterisco (*) las mejores configuraciones de parámetros obtenidas mediante el ranking de parámetros (ver sección 6.2.3.1).

Los casos de estudio utilizados para evaluar los experimentos corresponden a dos incendios controlados denominados caso A y B (correspondientes a los casos 4 y 2 de la Tabla 5.1, respectivamente).

Tabla 6.2: Lista de experimentos según LHD.

Exp.	Umbral	It.	Ind.	Ind. Migrar	Frec. M.	Cruce (%)	Mut. (%)
0*	0,70	200	200	5	5	60	0,5
1	0,78	373	68	3	2	44	0,7
2*	0,75	419	194	17	4	37	1,7
3*	0,74	198	175	3	7	19	1,1
4*	0,71	302	41	11	10	42	0,3
5*	0,81	172	25	9	3	22	2,0
6	0,56	359	125	4	9	27	2,1
7*	0,68	129	145	20	7	24	2,2
8	0,69	117	91	5	7	58	1,7
9*	0,59	429	187	7	5	30	0,4
10	0,57	144	51	6	4	28	0,6
11*	0,65	249	97	19	8	10	0,9
12*	0,73	454	39	9	12	40	0,35
13*	0,84	465	110	12	3	13	1,5
14*	0,62	263	15	4	6	53	1,4
15	0,64	328	62	16	3	17	1,3
16*	0,80	293	123	13	9	33	0,5
17*	0,67	486	162	15	6	56	1,1
18*	0,82	234	139	9	5	48	2,4
19	0,77	398	37	8	9	39	1,8
20*	0,161	304	84	14	1	45	2,4

La ejecución de los experimentos se llevó a cabo evaluando tanto el tiempo de ejecución promedio de cada configuración de parámetros, como así también la calidad de predicción para cada paso de predicción (t_i). Debido a que ESS-IM posee un comportamiento estocástico, se fijaron 30 semillas para ser utilizadas en cada experimento. Por lo tanto, por cada configuración de parámetros del LHD se realizaron 30 ejecuciones, una con cada semilla, y se tomó como resultado final (para cada paso de predicción) el promedio de éstas. Es importante tener en cuenta que el conjunto de 30 semillas fue el mismo que se utilizó en cada experimento; los cuales fueron ejecutados en un clúster Linux, bajo un entorno MPI y Gigabit Ethernet al igual que los experimentos del Capítulo 5.

Tabla 6.3: Caso A: Resultados promedios de calidad de predicción y tiempo de ejecución.

Exp.	Paso1	Paso2	Paso3	Promedio	Tiempo (seg)
0	0,8454280	0,8436720	0,769039	0,8193795000	2988,665519
1	0,1060040	0,1085770	0,772652	0,3290776667	2612,619676
2	0,1505860	0,8013100	0,610826	0,5209073333	2244,432355
3	0,2862210	0,8132780	0,806390	0,6352963333	1429,851508
4	0,8626640	0,7790930	0,800682	0,8141463333	1206,655527
5	0,6725220	0,5769640	0,801539	0,6836750000	1713,710328
6	0,2000710	0,7937920	0,682604	0,5588223333	1382,225014
7	0,2555400	0,8010320	0,781745	0,6127723333	1253,853691
8	0,0909530	0,8120400	0,730489	0,5444940000	2212,788611
9	0,1362760	0,8064392	0,816010	0,5862417333	2008,577016
10	0,2392040	0,6424780	0,793625	0,5584356667	1696,546906
11	0,0740595	0,7323470	0,500000	0,4354688333	1205,362326
12	0,0742444	0,7995600	0,824750	0,5661848000	1233,445413
13	0,2450130	0,7335930	0,763948	0,5808513333	4100,662664
14	0,1583520	0,7837970	0,780156	0,5741016333	870,643627
15	0,0740207	0,6402120	0,798640	0,5042909000	1105,949597
16	0,0740498	0,7904000	0,808626	0,5576919333	2261,608520
17	0,1154470	0,7155060	0,553548	0,4615003333	2756,909569
18	0,2280060	0,7243200	0,804320	0,5855486667	4860,861947
19	0,0979542	0,6428020	0,806483	0,5157464000	2441,529033
20	0,1616130	0,6282050	0,815780	0,5351993333	1830,742101

Los resultados obtenidos tras el análisis de calibración han sido comparados con los resultados de trabajos previos (resultados de fila 0 en Tablas 6.3 y 6.4), donde se utilizó la configuración clásica de parámetros sobre los mismos casos de estudio. En las Tablas 6.3 y 6.4, se pueden observar los resultados de calidad de predicción de cada experimento para cada paso de predicción, junto con el tiempo consumido por cada configuración de parámetros (en negrita se han resaltado los mejores valores de cada paso de predicción y el menor tiempo). Con el objetivo de cuantificar de forma global cada configuración de parámetros, se ha incluido una métrica adicional un valor promedio de calidad de predicción para cada experimento (columna promedio en Tablas 6.3 y 6.4). De esta manera se puede

determinar, de mejor manera, aquellas configuraciones de parámetros que obtienen mejor rendimiento en términos de calidad de predicción.

Tabla 6.4: Caso B: Resultados promedios de calidad de predicción y tiempo de ejecución.

Exp.	Paso1	Paso2	Paso3	Paso4	Paso5	Promedio	Tiempo (seg)
0	0,7837408	0,73424225	0,785167	0,7625736	0,7627824	0,76570	3192,0
1	0,1170257	0,8760320	0,718751	0,5892480	0,6666530	0,59354	1498,1
2	0,2778965	0,6598605	0,648686	0,8150000	0,7631790	0,63292	2449,9
3	0,2393410	0,8342910	0,721708	0,6763865	0,7297110	0,64020	1664,0
4	0,0826920	0,8067395	0,759430	0,7647505	0,7076375	0,62424	1601,1
5	0,4682175	0,3455755	0,162250	0,4615845	0,7904290	0,44560	2794,5
6	0,0816627	0,7454375	0,726323	0,7599560	0,7044820	0,60357	2350,8
7	0,3033375	0,7232620	0,750490	0,7276985	0,6625795	0,63340	1566,0
8	0,1172037	0,6676135	0,631893	0,8344765	0,6930670	0,58885	1992,5
9	0,3383910	0,7896695	0,772685	0,8092990	0,7965525	0,70130	2219,0
10	0,1267454	0,7686335	0,665308	0,4499535	0,6498410	0,53209	1780,4
11	0,0613998	0,8829155	0,653462	0,8386590	0,7605645	0,63940	1680,9
12	0,1326515	0,7376670	0,718484	0,6731375	0,7376750	0,59992	1336,6
13	0,1891811	0,1465973	0,759166	0,7312195	0,6653610	0,49830	6476,5
14	0,2702287	0,7780405	0,705941	0,6163610	0,7913965	0,63239	543,8
15	0,1213762	0,7192760	0,611154	0,6411255	0,7299095	0,56456	2054,5
16	0,0388287	0,6261275	0,776841	0,7649100	0,5994470	0,56123	3149,3
17	0,1097427	0,8154585	0,744599	0,7355995	0,7801650	0,63711	2620,3
18	0,2101255	0,8628810	0,685358	0,6809360	0,7729920	0,64240	5334,4
19	0,0856205	0,6627430	0,685097	0,6971825	0,4949295	0,52510	1797,2
20	0,2460463	0,8789670	0,642796	0,7950940	0,6981135	0,65220	1778,7

Para comparar de forma clara los resultados obtenidos, en la Fig. 6.1 se han graficado las curvas de calidad de predicción para cada caso de estudio. Con el objetivo de simplificar la interpretación de dichos resultados, sólo se han graficado 5 de los 20 resultados de interés por cada incendio, estos son:

a) La línea rotulada como *Clásica* representa los valores de calidad de predicción obtenidos utilizando la configuración clásica de parámetros (es decir, fila 0 de la Tabla 6.2).

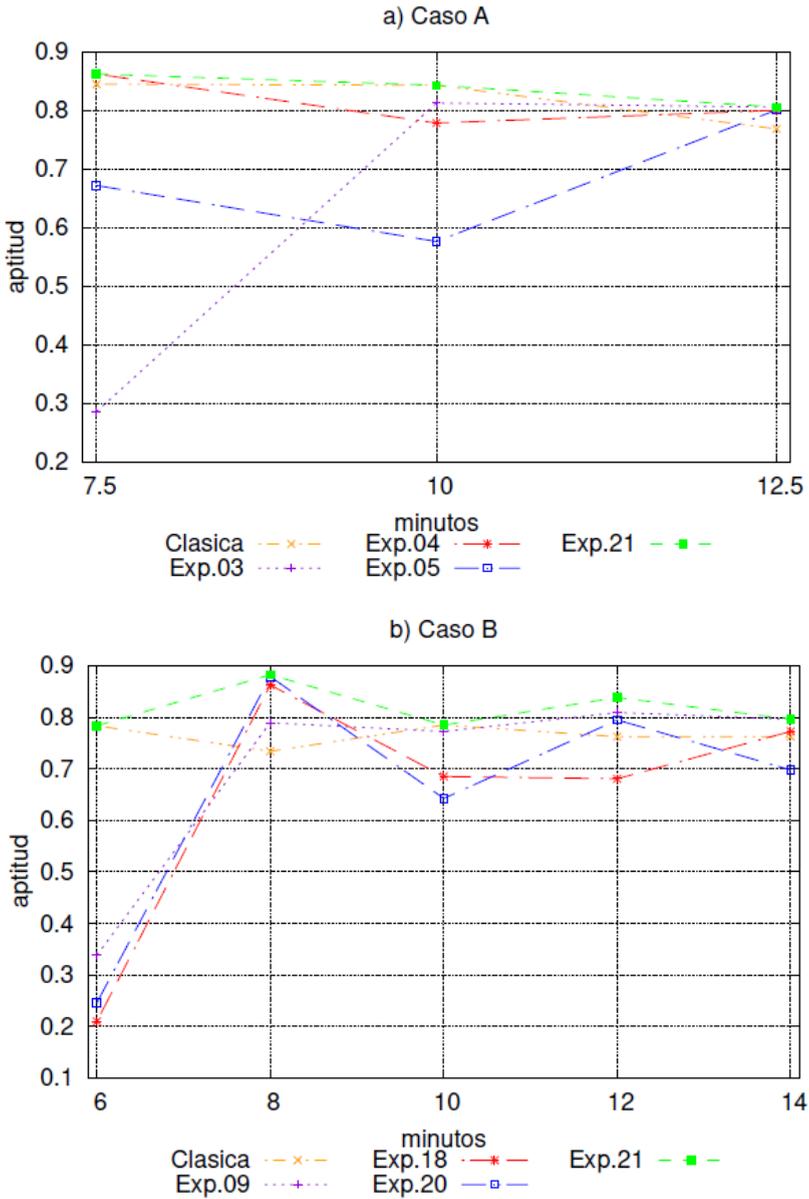


Figura 6.1: a) Comparación de calidad de predicción utilizando configuración clásica de parámetros, las tres mejores configuraciones y la configuración ideal (Exp.21). Caso A y Caso B.

b) Los tres mejores resultados en función de la métrica promedio:

- Caso A: experimentos: 3, 4 y 5.

- Caso B: experimentos: 9, 18 y 20.

c) La configuración de parámetros óptima seleccionando manualmente los mejores resultados de cada experimento y cada paso de predicción. La línea bajo el nombre de *Exp. 21*, representa los mejores resultados de calidad de predicción obtenidos tras comparar los 20 experimentos del LHD y la configuración clásica. En este caso los valores de calidad de predicción no corresponden a una única configuración de parámetros, es decir, los valores corresponden a los resaltados en negrita en las Tablas 6.3 y 6.4. La incorporación de esta curva nos permite valorar la dificultad existente para obtener una única configuración parámetros óptima, que brinde el mejor desempeño en todos los pasos de predicción.

Es importante tener en cuenta que, para el Caso A, los resultados de calidad de predicción previos a la calibración se encontraban por encima del 0,76903, lo que deja un margen de mejora bastante reducido, equivalente a 0,23096 (ya que, debido a las características del fenómeno, difícilmente se pueda alcanzar un valor de fitness próximo a 1). Por este motivo, y debido a que se está tratando de predecir un fenómeno natural de alta complejidad, se puede decir que los valores obtenidos son considerablemente importantes. En el Caso B, se produce una situación similar, el margen de mejora es igual a 0,26575. Sin embargo, en este caso los incrementos de calidad de predicción son un tanto superiores, principalmente en el minuto 8 y 12.

Si se observa el comportamiento de los resultados en función del tiempo de ejecución (ver Fig. 6.2), se puede observar que las configuraciones de parámetros con valores elevados consumen mayor cantidad de tiempo. Por ejemplo, el experimento 13 (para ambos casos de estudio) posee el mayor umbral de fitness de todo el diseño, igual a 0,84, y una alta cantidad de iteraciones equivalente a 465. Estos valores obligan a la metaheurística a iterar hasta que se alcance dicho valor de umbral o a llegar a la cantidad de 465 iteraciones evolutivas. En contraste, la configuración que menor tiempo consume es la 14, la cual tienen valores intermedios de umbral de fitness y cantidad de iteraciones, pero una reducida cantidad de individuos, igual a 15. Finalmente, también es interesante observar el tiempo consumido por las dos mejores configuraciones. Por ejemplo, para el Caso A, la mejor configuración (experimento 0) emplea un total de 2988,66 segundos, con un valor promedio igual a 0,81937, y la segunda mejor configuración (experimento 4) emplea 1206,65 segundos, con un valor promedio igual a 0,81414. En el Caso B, la mejor configuración (experimento 0) utiliza 3192,04 seg, con un promedio igual a 0,76570, y la siguiente (experimento 9) consume 2219,04 seg con un promedio de calidad de

predicción equivalente a 0,70131. En el Caso A, la reducción de tiempo es superior al 50% con una diferencia mínima en la calidad promedio, en cambio en el Caso B, la reducción no es tan significativa, y la diferencia en términos de calidad de predicción es bastante más notoria, lo cual puede justificar el mayor uso de recursos obteniendo una mejor calidad de predicción. Estos resultados evidencian que una configuración única de parámetros difícilmente logre obtener el mejor resultado en todos los pasos de predicción.

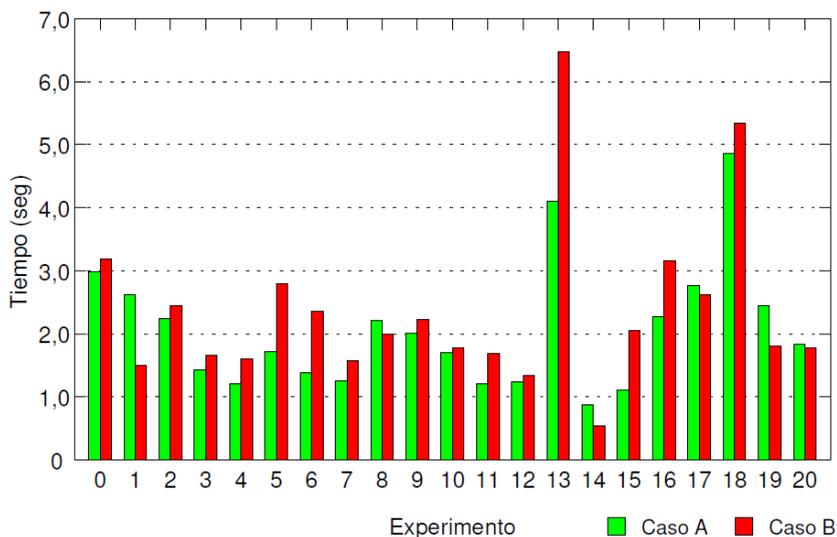


Figura 6.2: Comparación de tiempo de ejecución para cada experimento y caso de estudio.

Tal como ya se ha mencionado, la implementación de ESS-IM de esta experimentación utiliza una configuración de parámetros evolutivos de forma global u homogénea, es decir que, si bien cada isla inicializa su población utilizando una semilla diferente, todas las islas evolucionan siguiendo la misma configuración de parámetros (mismo criterio de búsqueda). En función de los resultados obtenidos se puede observar que la calibración de parámetros permite incrementar la calidad de predicción del método en más del 60 % de los pasos de predicción (considerando los resultados del experimento 21 de la Fig. 6.1), destacando la importancia de todo análisis de calibración. Además, que no es posible determinar una única configuración de parámetros óptima, debido que los mejores valores encontrados (para cada paso de predicción) no son generados por una misma configuración de parámetros. Por lo tanto, una determinada configuración puede brindar

en t_i , un valor de alta calidad de predicción, y en t_{i+1} lo contrario, o viceversa. Además, es importante observar que, si bien se logra incrementar la calidad en diferentes pasos de predicción en ambos casos de estudio, los mejores valores en la medida promedio son obtenidos por la configuración clásica. En este sentido, tal como se menciona en [176] y en función de los resultados presentados, se observa que la robustez del método puede ser incrementada si se utiliza una configuración heterogénea de parámetros evolutivos. Es decir, si el método utiliza diferentes configuraciones de parámetros en cada isla.

6.2.3. Calibración heterogénea

En [176] y [177], se ha demostrado que la combinación de varios procesos de búsqueda utilizando configuraciones heterogéneas de parámetros, incrementa la robustez del mecanismo de búsqueda, permitiendo explorar el espacio de soluciones de manera más eficiente, facilitando así la obtención de mejores resultados. Desde el punto de vista de configuración de parámetros evolutivos, un AEP basado en islas puede implementarse tanto de forma homogénea como heterogénea. La primera consiste en que la totalidad de las islas utilicen la misma configuración de parámetros evolutivos (a excepción de las poblaciones inicializadas de forma independiente), es decir, los mismos valores de tamaño de población, probabilidad de cruce, mutación, etc. En cambio, el segundo consiste en utilizar diferentes configuraciones de parámetros en cada isla, ya sean configuraciones previamente conocidas, obtenidas por medio calibración, o incluso utilizando valores aleatorios dentro de rangos preestablecidos.

En una implementación heterogénea existen parámetros que suelen ser compartidos entre las islas, como: a) la cantidad de individuos a migrar y b) la frecuencia de migración. Con el propósito de evitar incompatibilidades entre estos parámetros, en este estudio de calibración se plantean cuatro escenarios diferentes, representados gráficamente en la Fig. 6.3, los cuales se describen brevemente a continuación:

a) *Menor cantidad de parámetros*: este esquema plantea que los parámetros compartidos (frecuencia de migración y cantidad de individuos a migrar) no utilicen valores heterogéneos en cada isla, es decir, que sean instanciados con valores fijos. El resto de los parámetros utilizarán, en cada isla, los valores correspondientes a las configuraciones de parámetros establecidas en cada isla. Por ejemplo, en la Fig. 6.3.a la isla 0 utiliza la configuración de parámetros 16,

mientras que la isla 1 utiliza la configuración 4. De esta manera, cada isla utilizará todos los valores establecidos en la Tabla 6.2 para cada configuración, a excepción de los parámetros de migración. Los cuáles serán obtenidos a partir de la configuración 0 de la Tabla 6.2.

b) *Agrupamiento de islas*: en este esquema intervienen todos los parámetros, pero aquellas islas que utilizan la misma configuración de parámetros se agrupan, generando una migración entre grupos de islas que poseen diferentes parámetros (existiendo de esta manera dos jerarquías de migración). En la Fig. 6.3.b se ha graficado este concepto. Como se puede observar, las islas 0–1 y 2– n (que poseen la misma configuración de parámetros entre sí) son agrupadas, evitando de esta manera la incompatibilidad de parámetros de migración.

c) *Promedio de parámetros*: en este escenario las islas operan con parámetros diferentes, pero al momento de realizar la migración, se utiliza un valor promedio para aquellos parámetros en conflicto. Es decir, si una isla posee en su parámetro *cantidad de individuos a migrar* un valor n , y la otra isla un valor m , se utilizará como valor en las migraciones el promedio de ambos. En la Fig. 6.3.c se han representado dos islas (con diferentes configuraciones de parámetros), las cuales en el proceso de migración utilizan el promedio de los parámetros de ambas islas.

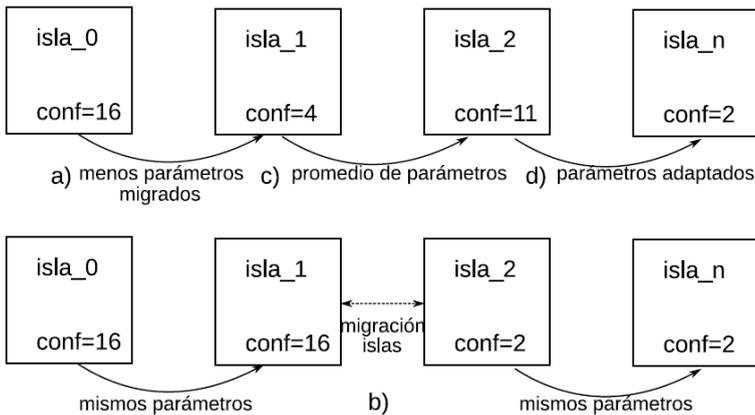


Figura 6.3: Escenarios de configuración de parámetros compartidos: a) menor cantidad de parámetros, b) agrupamiento de islas, c) promedio de parámetros, y d) parámetros adaptativos.

d) *Adaptativo*: en este esquema cada isla utiliza una configuración diferente de parámetros, en donde las islas se adaptan a la configuración de parámetros de migración de la isla emisora con la cual se encuentra

comunicando en cada instante de migración. Por ejemplo, en la Fig. 6.3.d, para el parámetro *cantidad de individuos a migrar*, donde cada isla posee un valor diferente, migración entre dos islas con diferentes parámetros una isla puede enviar n individuos y recibir m individuos.

La experimentación llevada a cabo consistió en tomar los resultados obtenidos en la sección 6.2.2, los cuales permitieron determinar configuraciones que ofrecían valores prometedores de calidad de predicción, sin poder determinar una única configuración, ya que los resultados varían por cada paso de predicción. En dicha calibración se evaluaron las 21 configuraciones de parámetros (Tabla 6.2), 20 de ellas generadas a través de un Diseño por Hipercubo Latino, además de un experimento adicional obtenido en base a una configuración clásica de parámetros (configuración No. 0 en Tabla 6.2).

6.2.3.1. Ranking

Con el objetivo de utilizar aquellas configuraciones que ofrecen mejores resultados en términos de calidad de predicción, se decidió elaborar un ranking que permitiese llevar a cabo esta identificación, y además reducir la cantidad de experimentaciones. El ranking se elaboró analizando el resultado de las configuraciones de la Tabla 6.2, es decir, los resultados obtenidos para cada configuración de parámetros. La clasificación de las configuraciones de parámetros se realizó para ambos incendios (columnas Caso A y Caso B en Tabla 6.5) como para el promedio de ambos (columna Promedio en Tabla 6.5). El ranking generado puede observarse en la Tabla 6.5; los valores de las columnas Caso A, Caso B y Promedio Total indican el número de configuración de parámetros de la Tabla 6.2. Por lo tanto, de las 21 configuraciones iniciales, se logró reducir, mediante un ranking de 11 posiciones, a un total de 15 configuraciones posibles de parámetros finales, las cuales han sido marcadas con “*” en la Tabla 6.2. Es importante remarcar que las configuraciones del ranking (Tabla 6.5) son las que ESS-IM y asigna a cada una de las islas al momento de instanciar el método. La selección de cada configuración, a nivel de isla, se realiza de forma aleatoria en el proceso master de cada isla (previamente la totalidad de las configuraciones son enviadas desde el proceso monitor).

Tabla 6.5: Ranking de configuración de parámetros. Cada valor de las columnas: Caso A, Caso B y Promedio corresponden a números (No.) de configuraciones de parámetros de la Tabla 6.2.

Ranking	Caso A	Caso B	Promedio
1°	No. 0	No. 0	No. 0
2°	No. 5	No. 9	No. 9
3°	No. 18	No. 20	No. 18
4°	No. 3	No. 18	No. 3
5°	No. 14	No. 3	No. 14
6°	No. 13	No. 11	No. 7
7°	No. 7	No. 17	No. 4
8°	No. 9	No. 7	No. 20
9°	No. 4	No. 2	No. 12
10°	No. 16	No. 14	No. 2
11°	No. 12	No. 4	No. 5

6.2.3.2. Resultados

Para la experimentación se utilizaron los mismos casos de quemas controladas del estudio anterior, es decir Caso A y B. Los resultados obtenidos, tanto en términos de calidad de predicción como en tiempo de procesamiento, han sido comparados con los obtenidos por ESS-IM utilizando una configuración homogénea de parámetros. Con el objetivo de clarificar las gráficas y tablas de resultados, se ha nombrado como: ESS-IM y ESS-IMh, el primero hace referencia a la implementación original de ESS-IM basada en una configuración homogénea de parámetros evolutivos, y el segundo a la versión de configuración heterogénea. Es importante remarcar que la configuración de parámetros utilizada por ESS-IM en dichos experimentos corresponde a la configuración No. 0 de la Tabla 6.2 (dado que en los resultados promedios es la configuración que obtiene mejor desempeño), mientras que en ESS-IMh cada isla utiliza algunas de las 15 configuraciones generadas por el ranking. Ambos métodos poseen la misma configuración de parámetros para todas las características comunes relacionadas con la configuración operativa del AEP (es decir, número de islas, workers por isla, etc.). Es importante tener en cuenta que las diferentes configuraciones evaluadas se representan en la Tabla 6.6. Tal como se puede observar, se llevaron a cabo 12 experimentos, donde se varía el tipo de migración (tipos a, b, c o d de la Fig. 6.3) y la existencia de cambios de parámetros por isla en

tiempo de ejecución (modificación realizada en cada paso de predicción). Los valores posibles son: a) *si*, b) *no*, y c) *probabilidad*, donde este último indica la probabilidad con la que se cambiarán los parámetros de cada isla en cada paso de predicción. En todas las ejecuciones se utilizó un valor de probabilidad igual a 0,2. Y finalmente, la última columna indica (sólo para el tipo de migración *b*) el porcentaje de migración de individuos (entre grupos de islas) y la frecuencia. Al igual que el parámetro mencionado previamente, tampoco es sintonizado.

Tabla 6.6: Configuración de experimentos realizados.

Exp.	Tipo de Migración	Cambio por paso de predicción	Migración jerárquico
1	a	si	–
2	b	si	10% cada iteración
3	c	si	–
4	d	si	–
5	a	no	–
6	b	no	10% cada iteración
7	c	no	–
8	d	no	–
9	a	probabilidad	–
10	b	probabilidad	10% cada iteración
11	c	probabilidad	–
12	d	probabilidad	–

En la Fig. 6.4 se ha graficado la calidad de predicción para los N pasos de predicción de cada experimento. Dicha calidad de predicción obtenida para cada método ha sido generada a partir del promedio de 20 ejecuciones. Es importante mencionar que se han graficado únicamente los cuatro mejores resultados obtenidos, siendo éstos los producidos por los experimentos: 5, 6, 9 y 10. Tal como se puede observar, el hecho de modificar la configuración de parámetros en cada paso de predicción no aporta mejoras en los resultados obtenidos (experimentos 1 a 4). Entre los experimentos 5 y 9 prácticamente existe poca diferencia, estos experimentos implementan el esquema de migración a, el cual utiliza parámetros diferentes sin incluir los parámetros de migración, en su lugar se utiliza la configuración de ESS-IM (exp. 0, de la Tabla 6.2). Los mejores resultados corresponden a los experimentos 6 y 10, los cuales utilizan un esquema de migración basado en agrupamiento de islas, donde entran en juego todos los

parámetros, y donde es necesario incluir una migración de mayor jerarquía entre los grupos de islas. Este esquema permite alcanzar mejores niveles de calidad de predicción en ambos casos de estudio.

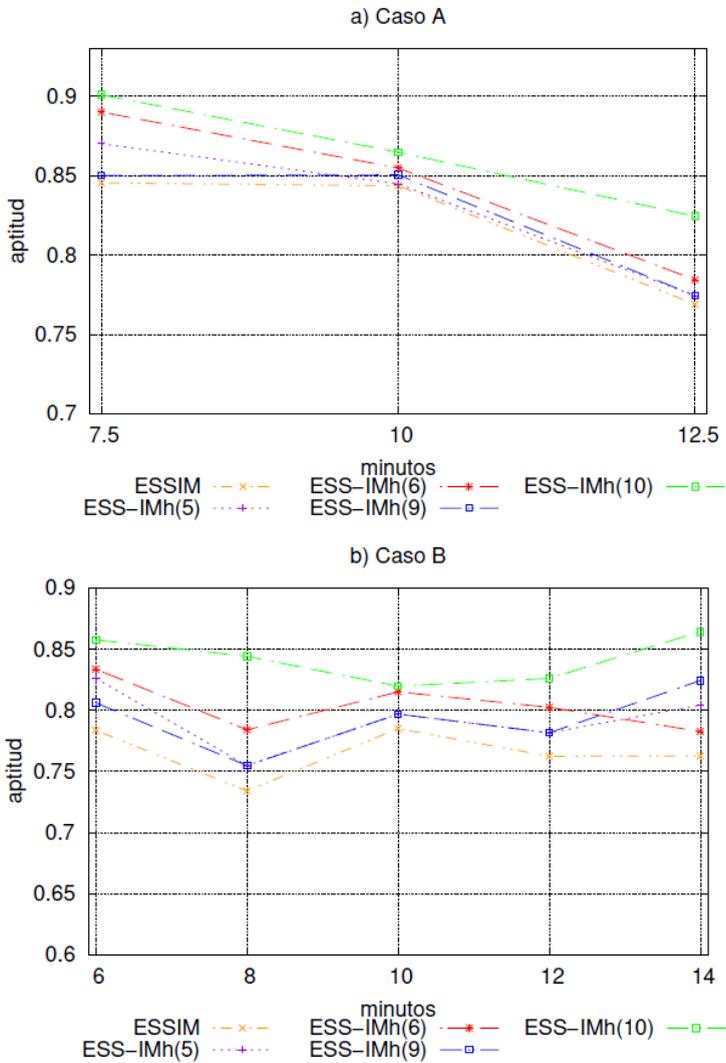


Figura 6.4: Comparación de calidad de predicción utilizando configuración clásica de parámetros, las tres mejores configuraciones y la ideal. Caso A y B.

6.3. Calibración de HESS-IM

En esta sección se presentan los resultados de un breve estudio de calibración del método HESS-IM. Es importante mencionar, que los parámetros utilizados en las experimentaciones de HESS-IM fueron establecidos a partir del estudio de valores de referencia hallados en la literatura, los cuales permitieron establecer una configuración de valores que ha permitido satisfacer las perspectivas esperadas dentro de rangos aceptables. Por este motivo, no se llevó a cabo un estudio paramétrico completo de todos los parámetros intervinientes. Tales parámetros de configuración se presentan en la Tabla 6.7. Estos son los utilizados en el Capítulo 5 en la sección 5.2.3, donde se presentaron los resultados de experimentación de HESS-IM.

Tabla 6.7: Configuración de parámetros HESS-IM, EA, DE y PSO.

Parámetro	Valor
Umbral de fitness	0,7
Iteraciones	200
Individuos por isla	200
Islas	5
Individuos a migrar	20,00%
Criterio	semi-elitista
Frecuencia	1 iteración
Workers por isla	7
(EA) Criadores	0,8
(EA) Mín. cruzamiento	0,2
(EA) Máx. cruzamiento	0,6
(EA) Mutación	0,5
(DE) f, Factor de mutación	0,9
(DE) Cr, Cruzamiento	0,3
(PSO) c1 y c2, Coeficientes de aceleración	0,5
(PSO) w, Factor inercial	0,6
(PSO) Tamaño del vecindario	incremental

Este estudio, al igual que los anteriores, se ha llevado a cabo con el propósito de identificar las configuraciones del método que puedan ofrecer el mejor desempeño posible. En este caso, el estudio se ha centrado en evaluar la mejor manera en que las islas deben utilizar a

cada metaheurística, relacionado con el proceso de migración dentro de la topología de comunicación, con el objetivo de lograr maximizar la colaboración e interacción entre las diferentes metaheurísticas. Por lo tanto, en este estudio no se evalúa el impacto de la variación de los parámetros principales de cada metaheurística. Sino en cambio, con una visión a nivel de islas, se experimenta con diferentes configuraciones de instanciación de metaheurísticas por isla. Estos esquemas se listan a continuación:

- Instanciación A: En este esquema, las islas pueden ser instanciadas con cualquier metaheurística, por lo que las migraciones pueden realizarse entre islas con iguales o diferentes metaheurísticas.
- Instanciación B: Aquí se genera aleatoriamente el orden o posición de las metaheurísticas, evitando que la misma metaheurística esté presente en dos islas lindantes. Es decir, la migración siempre ocurre entre diferentes metaheurísticas sin importar el orden.
- Instanciación C: Igual al esquema anterior, pero en este caso también se controla la secuencia de repetición de las metaheurísticas. Es decir, cada isla (en ambos sentidos del anillo) debe estar comunicada con islas que posean diferentes metaheurísticas.

6.3.1. Resultados

HESS-IM, con sus tres diferentes modalidades de instanciación de islas, ha sido aplicado a tres casos de quemas reales controladas, caso 1, 3 y 5 de la Tabla 5.1. Los resultados del promedio de 20 ejecuciones para cada experimento y caso se observan en la Fig. 6.5. Tal como se puede observar, la modalidad de instanciación de metaheurísticas que ofrece mejores resultados corresponde a la opción C, la cual superó en calidad de predicción al resto de las alternativas en los tres casos bajo análisis. Esta modalidad, al garantizar la migración entre diferentes técnicas de optimización permite mantener la diversidad en las soluciones y alcanzar mejor desempeño. En las gráficas también se puede observar que la configuración B en ciertos instantes de tiempo presenta resultados próximos a la mejor alternativa. Posiblemente este resultado se deba a que la configuración de instanciación permite que existan islas comunicadas (en ambos sentidos del anillo) a una misma metaheurística.

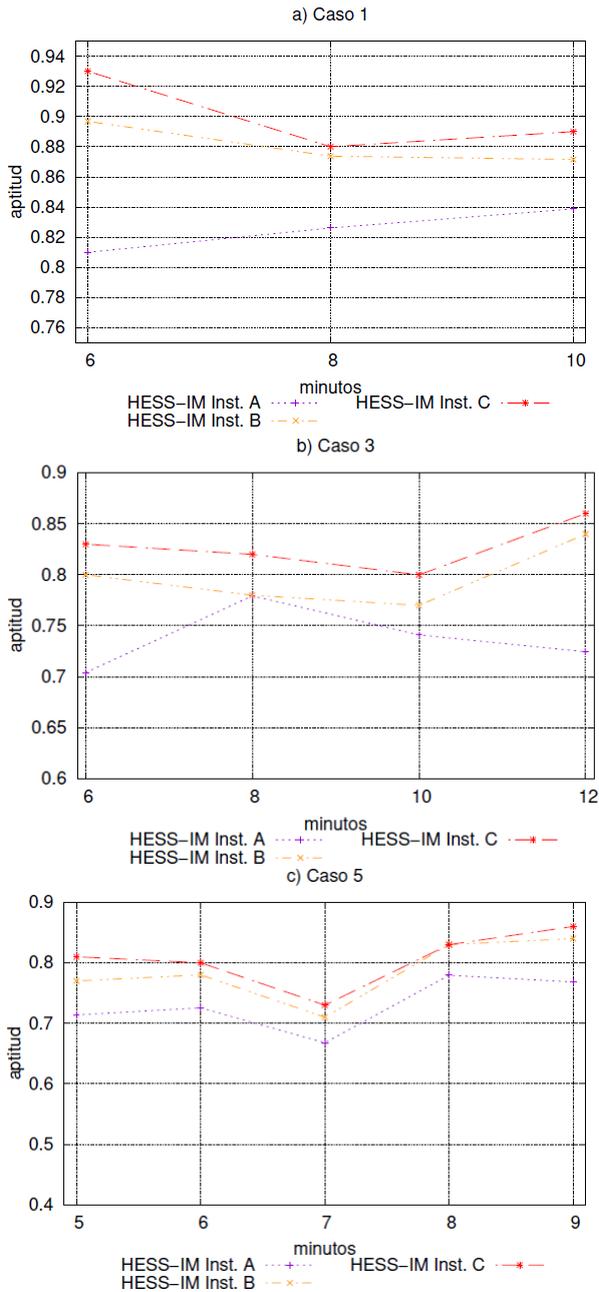


Figura 6.5: Comparación de calidad de predicción utilizando configuración clásica de parámetros, las tres mejores configuraciones y la ideal. Caso A y B.

6.4. Síntesis

Este capítulo ha presentado diferentes estudios de calibración llevados a cabo durante el desarrollo de la tesis. En ellos se ve reflejada la importancia de llevar a cabo este tipo de análisis, ya que sin realizar modificaciones estructurales a los métodos es posible lograr mejoras considerables en el desempeño de éstos. Producto de tales estudios se ha logrado determinar que un esquema de configuración heterogéneo de parámetros permite alcanzar mejores resultados en ESS-IM.

Imágenes Satelitales y Redes de Sensores Inalámbricos

En esta sección se presentan otras líneas de investigación en las cuales también se avanzó durante el desarrollo de la presente tesis y aún se continúa trabajando. Tales temáticas persiguen el objetivo de implementar los métodos de reducción de incertidumbre presentados en los capítulos previos en un sistema integral de detección, alerta temprana y predicción de incendios forestales con aplicación a situaciones reales. Puntualmente se trabajó con la utilización de imágenes satelitales y redes de sensores inalámbricos, esta última aún en desarrollo.

7.1. Introducción

Con el propósito de reducir la incidencia de los incendios forestales y mitigar sus efectos negativos, se vuelve necesario contar con herramientas que permitan: a) detectar en el menor tiempo posible la ocurrencia de estos fenómenos, y b) pronosticar con antelación el comportamiento de la línea de fuego.

En el primer caso, estas herramientas también se conocen como sistemas de alerta temprana, los cuales cumplen un rol muy importante en el proceso de lucha contra incendios debido a que permiten actuar antes de que el incendio se extienda demasiado, posibilitando así la utilización de menor cantidad de recursos en el proceso de extinción y minimizando los daños generados.

En el segundo caso, cuando los incendios ya se han magnificado, predecir su comportamiento puede ser la herramienta que marque la diferencia. La predicción del comportamiento de un incendio forestal puede permitir gestionar los generalmente “escasos” recursos disponibles, de manera eficiente y segura, facilitando la toma de decisiones complejas tales como evacuaciones, planificación de cortafuegos, extracción de combustible, entre otros.

En esta sección se presentan aspectos relacionados al diseño de un Sistema Integrado de Detección, Alerta Temprana y Predicción de

Incendios Forestales. Un esquema general de la propuesta se ha representado gráficamente en la Fig. 7.1, la cual incluye a uno de los métodos de reducción de incertidumbre desarrollados (HESS-IM), presentado en el Capítulo 4 (utilizado como sistema de predicción de comportamiento de incendios forestales). Y la integración con dos subsistemas: a) módulo de generación automática de mapas para su posterior tratamiento, y b) red (WSN/IoT) para la detección temprana de incendios forestales.

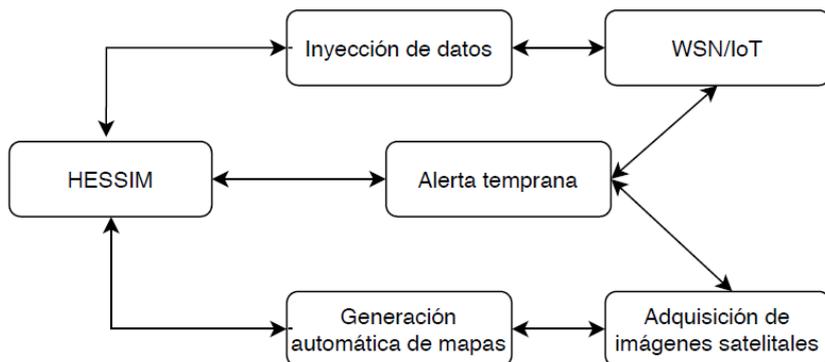


Figura 7.1: Esquema general del sistema integrado de detección, alerta temprana y predicción del comportamiento de incendios forestales.

7.1.1. Alerta Temprana y Monitoreo

En la actualidad existen diferentes enfoques para abordar este problema. Sin embargo, desde un punto de vista técnico, todavía no se ha resuelto completa y efectivamente [178]. Tradicionalmente, la tarea de monitoreo se realizaba por un equipo capacitado en una torre de observación ubicada en un punto alto [179]. Este método de monitoreo sigue siendo aplicable en algunos países, aunque debido a la falta de fiabilidad de las observaciones humanas, diferentes técnicas de visión han sido propuestas, las cuales sólo son válidas en pequeños bosques dado que dependen del alcance de visión de los instrumentos. Otro enfoque tiene que ver con el uso de imágenes satelitales. Esta tecnología si bien permite asegurar el monitoreo de grandes áreas, como método de alerta temprana presenta ciertas dificultades principalmente en términos de respuesta en tiempo real, ya que el tiempo de refresco es muy alto. También existe otro método ampliamente aplicado al monitoreo de incendios forestales que se basa en la obtención de datos

sobre la fuente de emisión utilizando sensores de adquisición de datos a bordo de aviones, vehículos no tripulados, drones, etc. Este enfoque requiere la disponibilidad de operadores y personal de mantenimiento de forma casi permanente, lo cual incrementa el costo de operación. Aunque su principal defecto, sin embargo, es que la resolución espacial y temporal se limita a una medición puntual en la ubicación actual del vehículo.

Debido a estos inconvenientes, y con el propósito de minimizar el costo de monitoreo de los incendios forestales, las Redes de Sensores Inalámbricos (WSN, Wireless Sensor Network) surgen como una alternativa apropiada dado que estos dispositivos son económicos, compactos y pueden ser desplegados en un territorio relativamente extenso (idealmente para áreas protegidas) funcionando de forma conjunta y autónoma [180]. En este contexto, el uso de WSNs para asegurar el monitoreo efectivo de fenómenos tales como incendios forestales cobra particular interés y significación en el ámbito científico [181]. En la literatura se pueden encontrar gran variedad de trabajos donde se utilizan este tipo de tecnologías como herramientas de alerta temprana de incendios forestales. Por ejemplo, en [182] se describe una WSN jerárquica dirigida a la detección temprana de incendios en áreas de riesgo, integrados con los centros de mando contra incendios, sistemas de información geográfica y simuladores de incendios. Esta propuesta incluye la utilización de un receptor GPS para definir la ubicación específica de cada sensor. Tal componente adicional por nodo afecta el desempeño energético de la red y aumenta de forma desmedida el costo de la misma, dado que la ubicación geográfica podría ser registrada al momento de instalar cada sensor evitando incurrir en el gasto que conlleva la utilización de receptores GPS. Esto cobra más sentido aún si tenemos en consideración que varios nodos se perderán en caso de producirse un acontecimiento de incendio forestal. En otro trabajo, [183], se propone un framework integral para el uso de redes de sensores inalámbricos para la detección y monitoreo de incendios forestales. El mismo incluye propuestas para la arquitectura de red de sensores inalámbricos, esquema de despliegue de sensores y protocolos de agrupación y comunicación. El objetivo principal de dicho trabajo consiste en poder detectar lo antes posible cualquier amenaza de incendio, aunque dicho trabajo no fue implementado en un ambiente real. En [184], llevan a cabo una serie de experimentos utilizando incendios controlados. La red está compuesta por diez nodos sensores con capacidad GPS. Cada nodo se despliega con rangos de hasta un kilómetro transmitiendo valores de temperatura, humedad y presión barométrica a una estación base. Sin embargo, debido a las grandes distancias entre los nodos sensores, los datos adquiridos no son

lo suficientemente valiosos como para detectar un incendio rápidamente y pronosticar la dirección de propagación del incendio. Además, con el crecimiento del fuego y la quema de algunos de los nodos sensores, la red podría fallar en la entrega de los datos de todos los nodos de sensor a la estación base. Si bien el uso de WSN puede ser una herramienta prometedora para poder detectar la ocurrencia temprana de incendios y posibilitar reducir la incertidumbre en los parámetros de entrada de los modelos, esta tecnología sólo puede ser utilizada en áreas protegidas (zonas delimitadas) donde la instalación de una red de sensores sea posible. Debido a esto, no es factible instalar una red extensa en grandes áreas siendo importante considerar su utilización junto a otras tecnologías (por ejemplo, imágenes satelitales).

7.1.2. Predicción de Incendios Forestales mediante Información Satelital

Tal como se describió en el Capítulo 5 los resultados de calidad de predicción obtenidos tras la evaluación de los métodos propuestos en casos de quemaduras controladas e incendios forestales han sido muy positivos. Tales resultados motivan a poder aplicar estos métodos a situaciones de incendios forestales en situaciones reales, para ello es necesario contar con herramientas de generación de los perímetros de área quemada inicial para posteriormente realizar tales predicciones. En este sentido, las imágenes satelitales conforman el recurso ideal para poder, en tiempo cuasi real, alimentar los sistemas de predicción con la información requerida.

Junto con esta información también podría ser posible incluir datos de redes de sensores inalámbricos, estaciones meteorológicas, datos históricos de incendios forestales, e inclusive podrían ser generados a partir de otros modelos (por ejemplo, modelos de viento, etc.). En el caso puntual de las imágenes satelitales, existe en la literatura numerosos antecedentes de su efectiva utilización en la detección de focos de calor, frentes de fuego activos, caracterización de incendios [185], utilizando diferentes plataformas satelitales tales como MODIS (Moderate-Resolution Imaging Spectroradiometer) [186], GOES (Geostationary Operational Environmental Satellite) [187], NOAA-AVHRR (National Oceanic and Atmospheric Administration - Advanced Very High Resolution Radiometer) [188], VIIRS (Visible Infrared Imaging Radiometer Suite) [189], [190], [191].

En este sentido, el objetivo planteado consiste en dar los primeros pasos hacia un sistema de alerta temprana, predicción y monitoreo de incendios forestales mediante el uso de información satelital y redes de sensores inalámbricos. El enfoque propuesto coincide con los Sistemas Dinámicos de Aplicaciones Guiadas por Datos (DDDAS, Dynamic Data-Driven Application Systems) [107], los cuales tratan de aprovechar al máximo la unión de los métodos de adquisición de datos, la computación y simulación en el desarrollo de sistemas. La creación de un DDDAS requiere la comprensión sobre el uso de sensores remotos, diferentes niveles de computación (por ejemplo, procesadores embebidos, supercomputadoras, redes, seguridad de datos, modelado matemático/computacional, problemas inversos, teoría de control, visualización, entre otros).

7.2. Detección de área quemada

Muchas son las metodologías que se han desarrollado para la detección de área quemada. Desde la utilización de aparatos físicos como el uso de sensores dispuestos en regiones forestales específicas para la detección y sensado en tiempo real, hasta la utilización de imágenes satelitales para detectar y fotografiar sectores de áreas quemadas tanto en el momento de su ocurrencia como los resultados evidenciados una vez finalizado el incendio. Debido a la diversidad de metodologías elaboradas durante los últimos años, en [192] se propone una clasificación que divide a las mismas en tres grupos importantes: a) estimaciones a partir de focos activos, b) estimaciones a partir del uso de bandas e índices espectrales, y c) estimaciones a partir de metodologías híbridas.

a) Estimaciones a partir de focos activos: En un principio, y a causa de la falta de datos de áreas quemadas a escala global, las estimaciones se realizaban mediante la detección de los diferentes focos activos [193]. Sin embargo, esta metodología presentaba muchas limitaciones, especialmente la dificultad de relacionar la cantidad de focos activos con el total del área quemada por éstos, ya que las causas de los mismos pueden deberse a diversos factores difíciles de reconocer, los que influyen directamente en la evolución de un incendio. La estimación indicaba que un área quemada era directamente proporcional al número de incendios [194], para mejorar esta estimación, diferentes estudios hicieron utilización del NDVI (Normalized Difference Vegetation Index) y la fracción de cobertura arbolada, logrando de esta manera

valores que demuestran mayor fidelidad en determinar la zona del incendio. Sin embargo, la precisión de este método por sí solo no era lo suficientemente exacto para el reconocimiento del área quemada por un incendio forestal.

b) Estimaciones a partir del uso de bandas e índices espectrales: Gracias al avance de la teledetección y del creciente lanzamiento de satélites capaces de tomar imágenes terrestres mediante el uso de instrumentos de medición de franjas espectrales determinadas en longitudes de onda, la estimación de áreas quemadas ha tenido un avance significativo. Las bandas del infrarrojo de onda corta (SWIR, Short-Wave Infrared), capaces de penetrar el humo e identificar fehacientemente los sitios con actividad de fuego; el infrarrojo cercano (NIR, Near InfraRed), con una gran capacidad para reflejar las vegetaciones del terreno [195], o el rojo visible (RED), son un ejemplo de aquellas bandas que han permitido la detección de las vegetaciones y del fuego. El uso y combinación de estas bandas ha permitido establecer índices espectrales que favorecen a la detección de la vegetación en sus diferentes estados, así como a una mejora significativa en la detección de fuego vivo y cenizas en sectores donde se producen incendios [196].

Entre los diferentes índices obtenidos por la combinación de las diversas bandas espectrales de estos satélites, los más importantes son el mencionado NDVI, o Índice de Vegetación de Diferencia Normalizada, usado para estimar la cantidad, calidad y desarrollo de la vegetación; y el índice NBR, o Índice de Quema Normalizado (Normalized Burn Ratio), resultado de normalizar la diferencia de reflectancias entre una banda del infrarrojo (IR) cercano y una del infrarrojo medio. El contraste entre las reflectancias de estas bandas está relacionado con la presencia de vegetación viva ya que ésta es muy reflectiva en la banda del IR cercano y no tanto en la del IR medio, mientras que la que ha sufrido daños tiene cada vez valores de reflectividad más bajos en la banda del IR cercano y más altos en la del IR medio. De esta forma, la manera simple para evaluar el área quemada y la severidad del incendio es calcular el índice NBR y asociar valores bajos del índice con alta severidad [197].

Una de las herramientas más populares actualmente es MODIS [198], la cual se encuentra integrada en los satélites Terra y Aqua. La órbita de Terra alrededor de la Tierra está sincronizada para recorrer de norte a sur, mientras que Aqua recorre de sur a norte, así ambos satélites están visualizando toda la superficie de la Tierra cada 1 o 2 días. Además de tratarse de un producto internacionalmente conocido y utilizado para diferentes tipos de biomasa, MODIS utiliza el algoritmo MOD14 para la detección de eventos de alta temperatura, facilitando la teledetección

de los incendios forestales. Debido a estas características, la herramienta MODIS es utilizada en el presente trabajo.

c) Estimaciones a partir de metodologías híbridas: Existen estudios que hacen uso de los incendios activos como una capa extra de información a las áreas quemadas. Principalmente, las detecciones de incendios activos son utilizadas como medio para seleccionar píxeles quemados en una o varias imágenes de incendios forestales, o bien, para identificar como incendios aquellos polígonos clasificados según la técnica utilizada [199].

7.2.1. Tipos de imágenes necesarias para estimar a partir del uso de bandas e índices espectrales

Los sensores utilizados por las plataformas satelitales obtienen información dentro de rangos específicos de longitud de onda dentro del espectro electromagnético. Esta información es registrada en distintos canales o bandas espectrales. Para visualizar las imágenes satelitales, se pueden combinar las distintas bandas mediante los tres colores primarios visibles por el ojo humano (azul, verde y rojo). De este modo, es posible representar y visualizar la energía de longitudes de onda, invisibles al ojo humano, como la luz infrarroja, las cuales son de gran utilidad para estudiar los distintos fenómenos.

Las imágenes satelitales están conformadas por pequeñas secciones de igual tamaño, llamados píxeles, los cuales integran la información sobre estas bandas espectrales representadas en el sistema RGB. El conjunto de estos píxeles forma una malla o ráster. Así, cada píxel del ráster tendrá un valor para cada color: rojo, azul y verde. Las imágenes satelitales se encuentran en formato digital, uno de los formatos más populares es el conocido GeoTIFF, utilizado como formato de intercambio para imágenes ráster georreferenciadas. El cual además de tener información de las bandas espectrales, también contiene otros datos como: coordenadas geográficas de la sección del mapa, tamaño del mapa, tamaño de cada celda y cantidad de celdas.

En este trabajo se hace uso del formato ráster, el cual conforma un mapa donde se visualiza el área quemada por el fuego. El mapa en cuestión se representa mediante una matriz de datos donde el tamaño de cada una de sus celdas corresponde a una porción del área total del bosque donde el incendio se ha producido, y cuyos valores son definidos por el instante de tiempo en el cual dicha celda ha sido alcanzada por el fuego, siendo el valor cero si dicha celda no ha sido quemada. En este sentido,

el proceso de automatización considera una etapa para la transformación de la información satelital al formato requerido.

7.2.2. Obtención de imágenes

Actualmente existen múltiples plataformas para la obtención de imágenes satelitales de incendios forestales, las cuales permiten tanto realizar descargas de las mismas para su procesamiento de forma local, como así también llevar a cabo la combinación de bandas espectrales desde las mismas plataformas on-line. Gran cantidad de estas herramientas suelen ser de pago. Sin embargo, algunas de ellas nos permiten trabajar de manera gratuita, aunque con ciertas limitaciones. Éste es el caso de Land Viewer, la plataforma web de EOS (Earth Observing System) [200] que en su capa gratuita permite crear imágenes de índices espectrales y la descarga de diez imágenes por día.

Dentro de las plataformas de uso gratuito se encuentra Earth Explorer, un servicio del USGS (United States Geological Survey) [202] que otorga herramientas para la descarga de información satelital, datos geográficos en tiempo real, así como el uso de sus APIs no sólo para incendios forestales, sino también diversos datos geográficos, marítimos y climatológicos. Otra plataforma importante por mencionar es el servicio EOSDIS Worldview de la NASA [203], la cual es una aplicación de visualización que permite operar con imágenes de los satélites MODIS. Esta plataforma a su vez permite el uso de combinaciones de bandas pre-diseñadas para distintos objetivos, tales como: calidad del aire, superficie del suelo, estado de la vegetación, temperatura de los océanos, tormentas de arena, incendios, entre otras.

En la Tabla 7.1 podemos observar las primeras 10 de las 36 bandas presentes para su utilización por los satélites MODIS. Las bandas número 1 y número 2, son usualmente aplicadas para detectar áreas quemadas y humo, así también, las bandas 1, 2, 6, 7, 20-25, 31 y 32, son utilizadas normalmente para la detección de fuego en actividad. Entre ellas, la banda número 7 es sensible a las altas temperaturas y no sensible a las medias, lo que la hace apta para verificar la intensidad del fuego [204]. La combinación utilizada en el presente trabajo corresponde a la 7-2-1, mejor conocida como MODIS (Terra/Aqua) Corrected Reflectance [195].

Tabla 7.1: Primeras 10 bandas MODIS, en nanómetros (nm). [201].

Banda	Longitud de onda (nm)
1	620 - 670
2	841 - 876
3	459 - 479
4	545 - 565
5	1230 - 1250
6	1628 - 1652
7	2105 - 2155
8	405 - 420
9	438 - 448
10	483 - 493

Por las características mencionadas de cada alternativa, en el presente trabajo se ha utilizado la plataforma Worldview para utilizar como fuente de datos para la búsqueda y descarga de imágenes satelitales. Dichas imágenes son utilizadas en el proceso de automatización para posteriormente alimentar el sistema de predicción del comportamiento de incendios forestales.

7.2.3. Proceso de Automatización

En la Fig. 7.2 se pueden observar las etapas que componen el proceso de con versión de imágenes satelitales a mapas de entrada para el método de predicción. Este proceso abarca desde la descarga del set de imágenes MODIS (del área bajo estudio) hasta el ingreso del mapa del incendio al sistema de predicción. Es importante mencionar que los métodos de predicción descritos no son capaces de procesar de forma directa las imágenes satelitales, por lo que éstas deben ser adaptadas al formato requerido (formato ráster ASCII).

Durante estos procesos se ha hecho uso de la biblioteca GDAL (Geospatial Data Abstraction Library, gdal.h) [205], [206]. La misma es una biblioteca de código abierto utilizada para la traducción y manipulación de formatos de datos geoespaciales ráster y vectoriales publicada por la Open Source Geospatial Foundation (OSGeo) [207] la cual es utilizada aquí para la conversión de mapas de incendios reales y algunas otras aplicaciones útiles como el redimensionamiento de los mapas.

En este sentido, tal como se puede observar (Fig. 7.2), la primera etapa consiste en la obtención de las imágenes satelitales del área bajo estudio, para ello se hace uso de la API de WorldView [208]. Ésta permite realizar la descarga de imágenes satelitales para ser procesadas o analizadas fuera de la plataforma on-line, es decir, de forma local en una computadora. Para ello es necesario conocer el área de interés y la fecha del incendio. Con estos datos se procede a descargar las imágenes correspondientes al incendio deseado⁴. Esta etapa se ejecuta para cada uno de los días del incendio (días transcurridos de los que se disponga imágenes satelitales), obteniendo un conjunto de mapas en formato GeoTIFF cuyos valores (tamaño total del mapa y resolución por pixel) deben ser idénticos para cada día.



Figura 7.2: Etapas del proceso de obtención y conversión de mapas.

La etapa siguiente hace uso de funciones de la librería de GDAL [209] para descomponer cada una de las bandas RGB como un mapa independiente. La función `GDALRasterBandH` (`GDALDatasetH hDS`, `int nBandId`), donde `hDS` corresponde al dataset a utilizar (correspondiente al mapa descargado) y `nBandId` será el número de banda a obtener. Por lo cual, es necesario llamar a esta función tres veces, una por cada banda existente, pasando al parámetro `nBandId` el

⁴ Es importante tener en cuenta que por el momento la información de entrada que requiere el proceso de conversión debe ser brindado por el usuario, se espera a futuro generar las solicitudes de forma automática a partir de sistemas de alertas tempranas.

valor 1 para la banda R, 2 para G y 3 para B, con el fin de obtener sus valores a partir del mapa del incendio. Como se observa en la Fig. 7.3, cada una de las bandas RGB corresponden a los valores de las bandas 7, 2, 1 respectivamente de los satélites MODIS. De esta forma la composición de las mismas en un sólo mapa permite ver el área incendiada, el incendio y el humo proveniente del mismo.

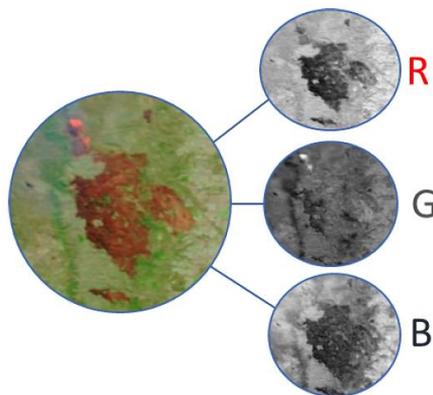


Figura 7.3: Descomposición de mapa RGB.

Una vez obtenidas las distintas bandas por separado, el tercer paso consiste en transformar cada uno de esos archivos en el formato ASCII correspondiente mediante la utilización de una segunda función de GDAL (`gdal_translate`) la cual realiza la conversión de datos ráster entre diferentes tipos de formatos y es capaz de ser usada para otras utilidades necesarias. Ésta convierte el archivo en una matriz de datos, donde el valor de cada celda corresponde a su valor en RGB para cada banda (entre 0 y 255), tales valores se utilizan para distinguir entre el área quemada y no quemada en cada instante de tiempo.

Cuando el área de trabajo posea grandes dimensiones puede ser necesario redimensionar el mapa encontrando la mejor relación de calidad y eficiencia entre el tamaño de celda y el tamaño total del mapa. Esta operación opcional, cuarta etapa, se realiza con la misma función de GDAL mencionada anteriormente, la cual permite redimensionar el tamaño de las celdas y la cantidad total de las mismas, sin modificar el área total del mapa. Este proceso permite disminuir el tiempo total de predicción ya que reduce el número de operaciones de cómputo que el sistema de predicción debe realizar. Es importante aclarar que, como todo proceso de redimensionamiento de una imagen de tipo ráster a un tamaño menor, implica la pérdida de información y, por ende, de calidad de dicha imagen. Por lo tanto, es necesario identificar

claramente la necesidad de su utilización teniendo en cuenta el equilibrio entre tiempo de trabajo y procesamiento y calidad de la imagen según su cantidad y resolución de píxeles.

A partir de la etapa 5, ya contando con el conjunto de mapas en formato ASCII para cada instante de tiempo, se procede a instanciar un algoritmo de comparación de mapas (ver Algoritmo 8). Donde **mapaR**, **mapaG**, **mapaB** corresponden a la matriz de mapas RGB, **ccSubestimacion** es un valor constante para controlar la subestimación del algoritmo, **instanteActual** corresponde al valor del instante (segundos, minutos, horas) desde el inicio del incendio hasta que la celda fue alcanzada por el fuego, y **mapaSalida** es el mapa resultado del algoritmo, el cual contendrá las celdas quemadas y no quemadas con los instantes de quema. El objetivo de esta etapa consiste en determinar aquellas celdas que han sido quemadas en la imagen actual (instante de tiempo x). Esto se logra identificando las celdas donde se encuentra predominancia del color rojo, ya que esto implica el área donde las celdas han sido alcanzadas por el fuego. Las celdas que sean identificadas como quemadas en el instante actual, serán marcadas (en el mapa de salida) con el valor de tiempo del instante actual (x). Y las celdas que no sean identificadas como quemadas, tendrán como valor en el nuevo mapa de salida cero, dando a entender que las llamas no alcanzaron dicha celda en el instante actual. Esto se logra comparando en cada celda el valor de R, si este es mayor que los valores de G y B en el instante de tiempo evaluado, la celda será marcada como quemada, caso contrario no. Terminado este proceso, se obtiene la totalidad de los mapas de salida, es decir, un nuevo conjunto de mapas con la delimitación del área quemada en su intervalo de tiempo específico.

En la etapa 6 se realiza una combinación del conjunto de estos mapas obtenidos. Para lograrlo, cada valor de celda de un instante de tiempo nuevo actualizará el valor del instante de tiempo anterior, con la excepción de aquellas celdas donde el nuevo instante sea igual a cero. Una vez finalizado, la salida es un mapa final con los valores del área de quema por instante en cada celda, dejando a la vista la dirección de propagación del incendio a lo largo del tiempo, desde que inició hasta que terminó. Finalmente, el paso número 7 corresponde al ingreso de este mapa final como parámetro de entrada (Línea de Fuego Real) en el sistema de predicción de incendios forestales.

Algoritmo 8: Pseudo-código del Algoritmo de Extracción de Incendio.

```
mapaR; /* Matriz de mapa Red (banda roja)*/
mapaG; /* Matriz de mapa Green (banda verde)*/
mapaB; /* Matriz de mapa Blue (banda azul)*/
ccSubestimacion; /* valor o rango de diferencia entre los valores de cada celda de los
mapaR y mapaG utilizado para asegurar que dicha celda fue alcanzada por el fuego */
instanteActual; /* valor del instante (segundos, minutos, horas) desde el inicio del
incendio hasta que la celda fue alcanza por el fuego. */
i; /* valor de fila de la matriz final */
j; /* valor de columna de la matriz final */
mapaSalida; /* Mapa final del incendio. */

repeat
  repeat
    if (mapaR[i][j] - ccSubestimacion) > mapaG[i][j] and mapaR[i][j] >
mapaB[i][j] then
      | mapaSalida[i][j] = instanteActual;
    else if mapaR[i][j] < mapaB[i][j] and mapaR[i][j] > mapaG[i][j] then
      | mapaSalida[i][j] = instanteActual;
    else
      | mapaSalida[i][j] = 0;
  until j == jmax;
until i == imax;
Output: mapaSalida; /* Mapa final del incendio. */
```

7.3. Generación de casos de prueba

A continuación, se describe la generación de dos casos de prueba que han sido utilizados para validar el desempeño del método de conversión y generación de mapas a partir de imágenes satelitales, como así también, de los métodos de reducción de incertidumbre desarrollados aplicados a grandes incendios forestales. Para la selección de estos casos se tuvo en cuenta incendios forestales que hayan tenido más de 5 días de duración (con el propósito de poder contar con al menos de 4 pasos de predicción), y que éstos cuenten con completa visibilidad durante todo el incendio, es decir, evitando la cobertura de nubes en alguna de las imágenes. Es importante tener en cuenta que los casos de prueba descritos a continuación son los utilizados en las experimentaciones de incendios forestales del Capítulo 5 (Tabla 5.2).

7.3.1. Caso 1

El primer caso se trata de un incendio ocurrido en el año 2008 en la localidad de Capilla del Monte, Provincia de Córdoba, Argentina. Las imágenes registradas del evento corresponden a 6 días desde el 29 de agosto al 3 de septiembre de dicho año. El incendio registró un total de 51.852,063 hectáreas quemadas [210, 211], en la Fig. 7.4 se puede observar el área quemada en el último día del incendio.

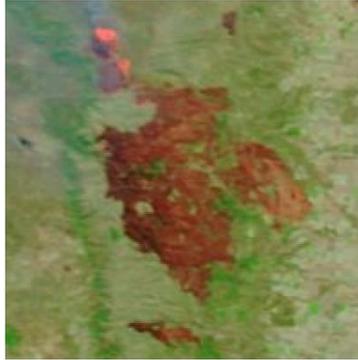


Figura 7.4: Día final del incendio en Capilla del Monte, Provincia de Córdoba, Argentina.

En la Fig. 7.5 se puede apreciar el conjunto de imágenes que componen el set de imágenes satelitales descargadas para este incendio (6 imágenes, una por cada día, ver Tabla 7.2).

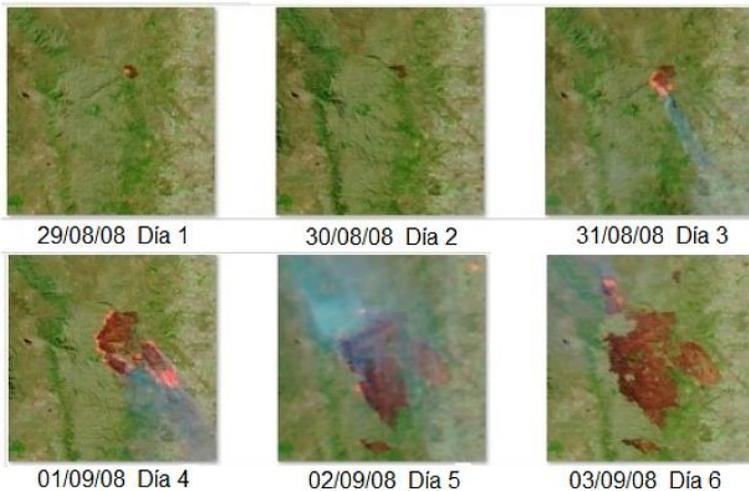


Figura 7.5: Días del incendio en Córdoba.

A su vez, deben estandarizarse los valores del tamaño total del mapa (tomando siempre los mismos puntos de coordenadas en el mapa), resolución por pixel y formato de archivo del mapa para trabajar con un conjunto de imágenes que se encuentren en la misma área, evitando de esta manera los errores que puedan surgir en aquellos procesos que precisen comparaciones entre ellas.

Tabla 7.2: Información de imágenes descargadas del incendio.

Cantidad de imágenes	6
Tamaño de imágenes (cantidad de celdas)	213px X 213px
Resolución (por pixel)	250 m
Formato de archivo	GeoTIFF

Para cada uno de estos seis mapas obtenidos se procedió a realizar el trabajo de descomposición de sus bandas RGB, teniendo así tres mapas por cada uno de ellos, ver Tabla 7.3. Una vez separadas las bandas, se llevó a cabo la conversión del formato GeoTIFF al formato ASCII precisado para cada uno de los 18 mapas nuevos. De esta manera, cada uno de los mapas corresponde a la matriz de 213x213 con un valor de 0 a 255 (rango de valores de color en formato RGB, siendo 0 el valor más claro y 255 valor más oscuro) en cada celda.

Tabla 7.3: Imágenes generadas en la descomposición de mapas RGB.

Imagen RGB por día	Mapa R	Mapa G	Mapa B
Mapa 29-08	29-08R	29-08G	29-08B
Mapa 30-08	30-08R	30-08G	30-08B
Mapa 31-08	31-08R	31-08G	31-08B
Mapa 01-09	01-09R	01-09G	01-09B
Mapa 02-09	02-09R	02-09G	02-09B
Mapa 03-09	03-09R	03-09G	03-09B

En la Fig. 7.6 pueden observarse las primeras tres filas del mapa 03-09R como una matriz de datos, donde cada celda implica su valor respecto de la banda R del sistema RGB. A su vez, en la parte superior se guarda la información de la cantidad de celdas, tamaño y coordenadas geográficas del mapa.

```

ncols      128
nrows     128
xllcorner  -123.172120000000
yllcorner  38.871830000000
cellsize   0.002197265625
NODATA_value -3.4028234663852886e+038
 66.0 80 93 107 112 115 114 104 90 83 75 61 47 36 34 37 41 43 49
F 52 48 41 36 30 27 28 29 33 39 43 44 43 42 40 40 41 43 44 45 47
I 43 54 57 60 61 56 55 57 60 68 65 63 63 66 68 71 75 72 78 90 98
L 106 105 99 91 79 71 66 59 55 51 49 53 53 58 56 59 67 71 58 38 24
A 22 22 25 32 38 44 47 52 51 47 42 41 41 45 49 68 64 60 68 81 90
 95 94 91 93 99 107 112 114 112 109 99 96 93 96 104 104 92 77 78
1 82 88 94 102 110 117
 67 78 96 111 116 118 115 104 87 85 75 62 48 38 34 37 42 44 53
F 58 53 46 41 39 38 37 39 41 43 45 45 44 42 44 43 40 41 42 41 42
I 42 48 53 58 59 57 55 54 57 56 58 60 64 69 71 71 73 72 82 95 108
L 117 116 110 102 85 78 71 63 57 51 49 49 49 51 54 61 77 86 72 49
A 41 35 31 30 29 32 37 43 53 53 47 43 41 44 49 55 63 59 60 76 96
 110 112 110 102 107 113 114 110 106 110 112 113 105 96 93 98 98
2 86 69 67 67 72 78 84 95 102
 81 93 109 121 119 114 107 90 71 60 55 52 46 41 37 40 42 40 40
F 42 40 43 47 43 38 45 44 44 46 48 48 48 47 38 36 33 31 32 31 33
I 35 42 50 56 61 60 58 56 55 45 50 58 65 70 72 70 69 73 85 99 115
L 123 122 114 105 93 85 76 67 59 53 49 45 36 36 41 52 74 89 78 55
A 36 33 32 29 27 30 40 49 58 55 48 42 39 43 48 53 52 59 70 86 105
 120 125 121 109 111 115 117 114 113 114 115 115 112 106 96 90 89
3 85 75 67 64 67 71 79 88 99

```

Figura 7.6: Datos del Mapa 03-09R en formato ASCII.

En el presente caso, debido al tamaño de las celdas en relación con el tamaño total del mapa y su repercusión en los tiempos de trabajo del sistema de predicción, se realizó una redimensión del mapa para obtener un tercio de la cantidad de celdas. Los nuevos valores de los mapas son los indicados en la Tabla 7.4, logrando con esto menor tiempo de procesamiento en el sistema de predicción.

Tabla 7.4: Información de mapas redimensionados.

Cantidad de imágenes	18
Tamaño de imágenes (cantidad de celdas)	71 px X 71px
Resolución (por pixel)	750 m
Formato de archivo	ASCII

Para el algoritmo de generación de área quemada, se realizó una comparación entre las imágenes de las bandas R, G, B, por cada día de quema (cada uno de los mapas descargados) de manera que aquellos valores de la misma celda cuya preponderancia se encuentre en el rojo, serán tomados como un nuevo valor para el nuevo mapa de salida y este valor corresponde al instante de quema en la celda. Es decir, si en una misma celda del mapa el valor de la banda R es mayor al de los dos restantes (G y B), se tomará como celda que ha sido alcanzada por el fuego y este valor (que estará en el rango de 0 a 255) será cambiado al

del instante de quema (minuto de quema desde que empezó el incendio) correspondiente, dejando en evidencia en qué instante de tiempo la celda fue quemada (esa sección del mapa fue alcanzada por el fuego). En el caso de que la preponderancia del valor de la celda no se encuentre en el rojo (fenómeno que nos indica que esa celda no ha sido quemada), el nuevo valor de esta celda será cero. Aunque este paso se repite para cada uno de los días solamente se presenta el mapa del día 3 de septiembre (Mapa 03-09) presentado en la Figura 7.7, generado a partir de la comparación de los mapas: 03-09R, 03-09G, 03-09B de la Tabla 7.3.

Tabla 7.5: Nuevos Mapas de Salida en Formato ASCII.

Nueva Imagen ASCII	Nuevo Instante de Quema
Mapa 29-08	1440 (minutos)
Mapa 30-08	2880 (minutos)
Mapa 31-08	4320 (minutos)
Mapa 01-09	5760 (minutos)
Mapa 02-09	7200 (minutos)
Mapa 03-09	8640 (minutos)

La Fig. 7.7 corresponde al nuevo mapa de salida en el formato ASCII requerido. Como se observa, todos aquellos valores de las celdas que no cumplieron con el valor adecuado del rojo, y cuyo nuevo valor equivale a cero, se encuentran en negro, en tanto que todas aquellas que si fueron alcanzadas por el fuego se encuentran en blanco. El valor de las celdas dependerá del instante de quema, en este caso, como este mapa corresponde al sexto día, su valor cargado en las celdas quemadas es de 8670 minutos (valor en minutos equivalente a seis días). Este proceso es realizado por separado para cada uno de los días, obteniendo un nuevo mapa de salida por cada combinación RGB, generando de esta manera seis nuevos mapas en el formato requerido (ver Tabla 7.5).

A partir de este momento ya es posible crear el mapa final, el cual se genera tomando cada mapa, desde el último hasta el primero y contrastando celda por celda. La comparación consiste en dejar el menor valor distinto de cero, de modo que pueda registrarse el avance del incendio temporalmente. La finalidad de realizarlo de esta manera (a la inversa de los sucesos) consiste en superponer las celdas de menores valores sobre las mayores, dejando a la vista en el nuevo mapa como se ha desarrollado el incendio en cada día.



Figura 7.7: Día final del incendio en Córdoba representado en el formato de salida ASCII.

Una vez obtenido el mapa de salida, cuyo resultado puede observarse en la Fig. 7.8, es posible introducirlo al simulador del comportamiento de incendios para ser procesado. En este mapa, cada celda tendrá el valor correspondiente al color indicado, por ejemplo, las celdas quemadas en el primer día tendrán el valor 1440, las celdas alcanzadas por el fuego el día 2 el valor 2880, etc. Es importante remarcar que este tipo de procesamiento (y mapa logrado), sirve para fines de evaluación, calibración, estudio y desarrollo de los métodos presentados, ya que contar con los sucesivos estados del incendio nos permite evaluar la aptitud lograda, y en consecuencia, la efectividad del método. Además, que el mapa presentado en la Fig. 7.8 no es utilizado para predecir, ya que las LFR no estarían disponibles durante el desarrollo del incendio.

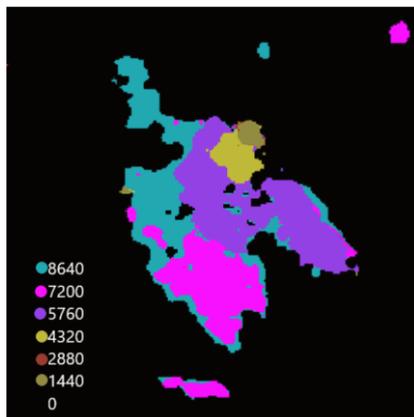


Figura 7.8: Mapa final del desarrollo del incendio en Córdoba en nuevo formato ASCII.

7.3.2. Caso 2

El caso número 2 corresponde al incendio denominado River Fire, parte del Mendocino Complex of Fires sucedido el año 2018, al este de Ukiah, California, Estados Unidos. Este incendio ocurrió durante el 29 de julio al 5 de agosto quemando más de 180.000 hectáreas [212], convirtiéndose en el incendio más extenso de California en toda su historia [213]. De ellas, cerca de 19.797 hectáreas corresponden al incendio de River Fire [214], presente en la zona oeste de Clear Lake. En la Fig. 7.9 puede observarse el total del complejo de incendios, incluyendo al River Fire, utilizado en el presente caso.



Figura 7.9: Mendocino Complex of Fires [215].

Debido a la progresión del incendio de River Fire en relación con el total del incendio, se ha trabajado desde el 28 de julio al 4 de agosto, fechas entre las que se produjo la totalidad del incendio en esta sección. Así, la cantidad de días excede en dos al del caso 1, con un total de 8 días de duración. De esta manera, es necesaria la descarga de 8 imágenes satelitales. En la Fig. 7.10 se pueden observar las imágenes resultantes de la progresión del incendio entre las fechas mencionadas.

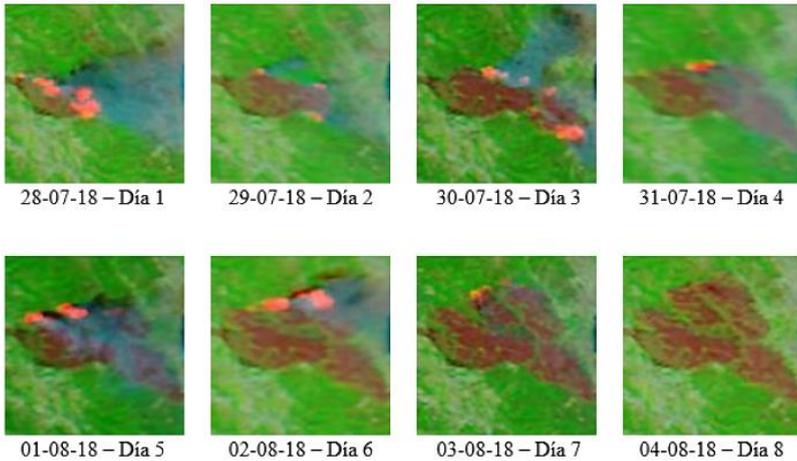


Figura 7.10: Días del incendio de River Fire.

En este caso, el valor del tamaño de incendio y del área del mapa a trabajar es significativamente menor al del caso número 1, como puede verse en la Tabla 7.8.

A partir de este set de mapas se realizó el proceso de descomposición por bandas para cada uno de ellos. La cantidad total de mapas resultantes para este proceso es 8 (cantidad de mapas) x 3 (cantidad de bandas por mapa). Es decir, 24 mapas (8R, 8G, 8B), uno de cada banda por cada día del incendio, ver Tabla 7.7.

Tabla 7.6: Información de Mapas.

Cantidad de imágenes	8
Tamaño de imágenes (cantidad de celdas)	104 px X 104 px
Resolución (por pixel)	250 m
Formato de archivo	GeoTIFF

Tabla 7.7: Descomposición de mapas RGB.

Imagen RGB por día	Mapa R	Mapa G	Mapa B
Mapa 28-07	28-07R	28-07G	28-07B
Mapa 29-07	29-07R	29-07G	29-07B
Mapa 30-07	30-07R	30-07G	30-07B
Mapa 31-07	31-07R	31-07G	31-07B
Mapa 01-08	01-08R	01-08G	01-08B
Mapa 02-08	02-08R	02-08G	02-08B
Mapa 03-08	03-08R	03-08G	03-08B
Mapa 04-08	04-08R	04-08G	04-08B

Con los nuevos mapas obtenidos, es necesario realizar la conversión de su formato GeoTIFF al ASCII. Este proceso genera 24 mapas con el mismo tamaños y celdas en el nuevo formato ASCII (Tabla 7.8).

Con los mapas convertidos, el proceso continúa con la generación del mapa final para cada día mediante el algoritmo de comparación de bandas. Este proceso es realizado para cada uno de los días del incendio, por lo cual, la salida será igual a 8 mapas donde cada uno de ellos mostrará la superficie del mapa que fue incendiada hasta ese día mediante el instante de quema en minutos. En la Tabla 7.9 se puede observar el valor de celda en minutos para cada día.

Finalmente, la última etapa genera el mapa final del incendio con las celdas marcadas con el valor del instante de tiempo en las que éstas fueron quemadas. Visualmente el mapa generado en este caso se puede observar en la Fig. 7.11.

Tabla 7.8: Información de Mapas.

Cantidad de imágenes	8
Tamaño de imágenes (cantidad de celdas)	104 px X 104 px
Resolución (por pixel)	250 m
Formato de archivo	ASCII

Tabla 7.9: Nuevos Mapas de Salida en Formato ASCII.

Nueva Imagen ASCII	Nuevo Instante de Quema
Mapa 28-07	1440 (minutos)
Mapa 29-07	2880 (minutos)
Mapa 30-07	4320 (minutos)
Mapa 31-07	5760 (minutos)
Mapa 01-08	7200 (minutos)
Mapa 02-08	8640 (minutos)
Mapa 03-08	10080 (minutos)
Mapa 04-08	11520 (minutos)

Es importante destacar que este procedimiento adquiere crucial importancia al momento de intentar utilizar los métodos desarrollados en situaciones reales, donde es necesario contar con información precisa en el menor tiempo posible, en este contexto nos referimos: el frente de fuego de un incendio forestal.

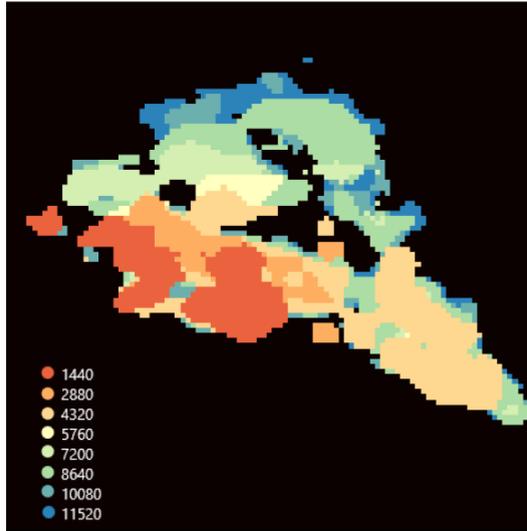


Figura 7.11: Mapa final del desarrollo del incendio River Fire en el nuevo formato ASCII.

7.4. Propuesta: detección de incendios con red de sensores inalámbricos

Las redes de sensores inalámbricos (WSN) se basan en dispositivos económicos y de bajo consumo llamados nodos. Estos, al estar dotados de sensores, son capaces de obtener información de su entorno, procesarla localmente, y comunicarla a través de enlaces inalámbricos hasta un nodo central de coordinación. Los nodos actúan como elementos de la infraestructura de comunicaciones al reenviar los mensajes transmitidos por nodos más lejanos. Las redes de sensores inalámbricos están formadas por numerosos dispositivos distribuidos espacialmente, los cuales utilizan sensores para controlar diversas condiciones en distintos puntos, como temperatura, humedad, presión entre otras. Los dispositivos son unidades autónomas que constan de un microcontrolador, una fuente de energía, un radio y n elemento/s sensor/es [216]. Existen diversas áreas de aplicación para las WSN, entre las que se destacan: monitoreo de energía, condiciones estructurales, transporte, monitoreo industrial y monitoreo ambiental. Enmarcado en esta última, el presente trabajo desarrolla un proyecto de utilización de una WSN con componentes de bajo costo para la detección de incendios forestales y su conexión con modelos de predicción para su prevención o seguimiento. En la actualidad existen

diferentes enfoques para abordar este problema. Este trabajo plantea la implementación de una WSN/IoT con nodos distribuidos en cuadrícula uniforme permitiendo regular el consumo energético [183]. Esto, además de facilitar la comunicación, permitirá la localización de los elementos evitando el uso de receptores GPS en cada nodo [217].

Algunos de los estándares para la comunicación entre nodos de una red de sensores son: Zigbee [218], IEEE 802.15.4 [218], IEEE 802.11 [219] y LoRa [221]. A su vez, existen diversas alternativas de hardware para utilizar como nodos. En este trabajo se ha optado por utilizar el chip ESP8266 D1. Cada nodo contará con dos sensores: a) DHT11 [222], para obtener valores en tiempo real de temperatura y humedad, y b) MQ-7 [223] utilizado para detectar la presencia de monóxido de carbono, permitiendo conocer la existencia de fuego. Asimismo, teniendo en cuenta la definición de esquema y despliegue descrita, mediante el nivel de monóxido de carbono detectado en un extremo de la red y considerando el tiempo en el que el mismo nivel sea detectado por un sensor próximo, se podrá estimar la velocidad y dirección del viento posiblemente pudiendo prescindir de otros elementos más costosos que cumplen esta función. Es importante tener en cuenta que la información obtenida por la WSN podrá ser utilizada para tres propósitos diferentes:

- a) Alertar anticipadamente ante la ocurrencia de un incendio forestal, permitiendo conocer además las zonas afectadas,
- b) Inyectar la lectura de variables en el método de predicción con el propósito de minimizar la incertidumbre, esto puede ser un complemento importante para reducir significativamente el tiempo de procesamiento necesario de los sistemas de predicción actuales. A fin de reducir la brecha de incertidumbre del modelo, podría alimentar al simulador con información real de un determinado parámetro, el cual, anteriormente era desconocido y calculado en base a estimaciones indirectas.
- c) Proporcionar en campo, al cuerpo de bomberos, información en tiempo real sobre el evento (por ejemplo, zonas afectadas, valores de temperatura, humedad, monóxido de carbono en el lugar y zonas aledañas, pronóstico de predicción, herramienta de comunicación, etc.)

7.4.0.1. Diseño de la red

Para el diseño de la red se considera utilizar dos tecnologías; por un lado, LoRa para los enlaces de larga distancia y por el otro, 802.11 para la comunicación entre nodos. Además, se evalúan dos alternativas para

los componentes del sistema, los nodos de la red WSN se encuentran representados por puntos de color rojo en el bosque. Dado que los nodos se alimentan por baterías y un panel solar, es importante minimizar la cantidad de comunicaciones entre nodos. Debido a la distribución de cuadrícula, cada nodo se conecta con 8 vecinos, por lo que es crucial garantizar la capacidad de regeneración de la red mediante los mecanismos de comunicación de datos necesarios para que a medida que los nodos se desconectan, la comunicación siga siendo efectiva entre todos los dispositivos que forman la red. Se llevaron a cabo pruebas mediante herramientas de simulación, con el fin de determinar el comportamiento de la red ante la caída de nodos en la WSN y de esta manera evaluar el grado de regeneración de la red. La regeneración de la red es un aspecto de gran importancia ya que, ante la presencia de fuego y la eventual caída de nodos, la red debe poder regenerarse en forma automática. También se han realizado experimentaciones preliminares de laboratorio, montando una red experimental compuesta por 4 nodos, donde se pudo corroborar la correcta regeneración de la red ante la eventual caída de algún nodo. Trabajos posteriores llevarán a cabo la experimentación de campo, para validar la correcta detección y la necesaria parametrización de la sensibilidad de los sensores para no subestimar las mediciones de los nodos. Finalmente es importante mencionar que la propuesta aquí descrita consiste en una línea abierta en la cual aún no se han obtenido resultados.

7.5. Síntesis

Repetidamente se observa, cuando los incendios forestales intervienen en las zonas de interfaz urbano-forestal⁵, la ocurrencia de eventos catastróficos con acontecimientos lamentables, donde además de las pérdidas materiales se producen pérdidas humanas. En la actualidad, grandes ciudades se encuentran amenazadas ante este tipo de eventos, dotar a estas áreas con redes de prevención (como WSN) para la detección temprana de incendios puede ser una solución factible, desde el punto de vista técnico como así también económico. Además, la vinculación de estas redes con una plataforma integral de alerta, monitoreo y predicción de incendios puede ser realmente beneficiosa. Este capítulo presentó avances relacionados a esta problemática,

⁵ Zonas donde se encuentran la unión de viviendas y otras estructuras con vegetación forestal y otros combustibles vegetales, lo que requiere un modelo de protección distinto del que se aplica en el terreno forestal.

utilizando además información satelital para dotar a un sistema de predicción con la capacidad de ser aplicado a incendios en tiempo real.

Conclusiones y Líneas de Trabajo Futuro

En este capítulo se exponen las conclusiones obtenidas tras el desarrollo de la presente tesis, como así también las posibles líneas de trabajo a futuro.

8.1. Introducción

En esta tesis se ha trabajado en el tratamiento de la reducción de incertidumbre en relación a los métodos de predicción y prevención de desastres naturales. Puntualmente se ha abordado como caso de estudio la problemática de los incendios forestales, uno de los más grandes problemas que afecta extensas regiones a lo largo del mundo. El problema en cuestión ha sido tratado desde diferentes perspectivas. Por un lado, el enfoque principal de este trabajo consiste en el desarrollo de métodos de reducción de incertidumbre que puedan ser utilizados para predecir el comportamiento de este fenómeno. Y por otro lado, se ha avanzado en otras aristas, como el diseño de un sistema integral de detección, alerta temprana y predicción de incendios forestales, el cual podría brindar asistencia en la lucha y mitigación de incendios, como así también en la prevención de los mismos. Este sistema integra los métodos propuestos, con el uso de redes inalámbricas de sensores y un sistema de generación automática de mapas alimentado por imágenes satelitales. Esta línea de trabajo podría hacer posible que este tipo de metodologías llegasen a ser implementadas y adoptadas por entes gubernamentales, permitiendo de esta manera minimizar los daños causados por este fenómeno.

En este trabajo se han propuesto y desarrollado dos métodos de reducción de incertidumbre capaces de asistir y participar en el proceso de resolución de problemas de alta magnitud y complejidad computacional, tal como la predicción y prevención de desastres naturales. En este sentido, los métodos desarrollados son: a) Sistema Estadístico Evolutivo con Modelo de Islas (ESS-IM), y b) Sistema Estadístico Híbrido Evolutivo con Modelo de Islas (HESS-IM), los cuales han logrado mejoras considerables en relación a la calidad de las predicciones generadas y así también en términos de tiempo de

respuesta, inclusive tras su aplicación con incendios forestales de gran escala.

Estos métodos han sido diseñados para poder aplicarse a otros fenómenos naturales con características de propagación (como avalanchas, deslizamientos de tierra, deslaves, inundaciones, entre otros). Debido a que estos fenómenos suelen causar gran impacto en las comunidades, la flora y la fauna, el ecosistema, entre otros, los sistemas de predicción deben proporcionar respuesta en el menor tiempo posible (sistemas altamente críticos). Por estos motivos, los métodos propuestos han sido desarrollados utilizando capacidades de alto rendimiento.

El primer método (ESS-IM), comenzó con el objetivo de lograr una mejora a una metodología previamente desarrollada denominada ESS (Sistema Estadístico Evolutivo). Este trabajo de tesis consistió en incrementar el paralelismo de la metaheurística interna de ESS, incorporando una arquitectura basada en modelo de islas con el propósito de obtener mejores resultados. Este desarrollo logró incrementar la capacidad de búsqueda de la metaheurística interna, impactando de forma directa en un incremento en la calidad de predicción del método. En la validación, ESS-IM fue aplicado en una serie de casos de quemas controladas e incendios forestales:

- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; TARDIVO, MARÍA LAURA; CAYMES SCUTARI, PAOLA. Predicción del Comportamiento de Incendios Forestales mediante aplicación de Modelo de Islas a ESS. Argentina. Ciudad Autónoma de Buenos Aires. 2013. Libro. Artículo Completo. Congreso. Noveno Encuentro del International Center for Earth Sciences (E ICES9). pp. 232–246. Comisión Nacional de Energía Atómica - CNEA. ISBN 978-987-1323-36-4.
- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; TARDIVO, MARÍA LAURA; CAYMES SCUTARI, PAOLA. Evolutionary-Statistical System with Island Model for Forest Fire Spread Prediction. Argentina. Mendoza. 2013. Libro. Artículo Breve. Simposio. Latin American Symposium on High Performance Computing HPCLatAm 2013.
- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; GIL COSTA, VERÓNICA. Método de reducción de incertidumbre basado en Algoritmos Evolutivos Paralelos orientado a procesos de predicción. Argentina. Ciudad Autónoma de Buenos Aires.

2014. Libro. Resumen. Jornada. Terceras jornadas de intercambio y difusión de los resultados de investigaciones de los doctorandos en ingeniería. Universidad Tecnológica Nacional.

- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; TARDIVO, MARÍA LAURA; CAYMES SCUTARI, PAOLA. Comparative analysis of performance and quality of prediction between ESS and ESS-IM. Uruguay. Montevideo, Uruguay. 2014. Libro. Artículo Completo. Congreso. XL Conferencia Latinoamericana en Informática (CLEI 2014). Universidad de la República (UdelAR). pp. 23–31. Universidad Católica del Uruguay, Universidad ORT Uruguay, Universidad de Montevideo y Universidad de la Empresa.
- MÉNDEZ-GARABETTI, MIGUEL; BIANCHINI, GERMÁN; TARDIVO, MARÍA LAURA; CAYMES-SCUTARI, PAOLA. Comparative analysis of performance and quality of prediction between ESS and ESS-IM. *Electronic Notes in Theoretical Computer Science*. Amsterdam: Elsevier. 2015 vol.314. pp. 45–60. ISSN: 1571-0661.
- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; GIL COSTA, VERÓNICA. Método de Reducción de Incertidumbre Basado en Algoritmos Evolutivos Paralelos Orientados a Procesos de Predicción. Argentina. Mendoza. 2015. Libro. Resumen. Encuentro. VIII Encuentro de Investigadores y Docentes de Ingeniería (ENIDI). UTN-FRM, UTN-FRSR, UMAZA, UM y UNCuyo.
- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; TARDIVO, MARÍA LAURA. Increase in the quality of the prediction of a computational wildfire behavior method through the improvement of the internal metaheuristic. *FIRE SAFETY JOURNAL*. Amsterdam. ELSEVIER SCI LTD. 2016 vol.82, pp. 49–62. ISSN: 0379-7112.

En trabajos posteriores, se llevaron a cabo diferentes investigaciones complementarias, tales como trabajos de sintonización de parámetros, análisis de redimensión de mapas, incorporación de diferentes modos de operación en la etapa de calibración del proceso monitor, entre otros. Estos desarrollos permitieron continuar mejorando el método:

- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; TARDIVO, MARÍA LAURA; CAYMES

SCUTARI, PAOLA. Predicción de Incendios Foresta les mediante ESS-IM: Análisis de Variación Dinámica de la Representación del Terreno. Argentina. Ciudad Autónoma de Buenos Aires. 2014. Revista. Resumen. Congreso. Décimo Encuentro del International Center for Earth Sciences (E-ICES10). pp. 409–420. Comisión Nacional de Energía Atómica - CNEA. ISBN: 978-987-1323-39-5.

- BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; MÉNDEZ GARA BETTI, MIGUEL; TARDIVO, MARÍA LAURA. Método de Reducción de Incertidumbre basado en Algoritmos Evolutivos Paralelos orientado a Procesos de Predicción. Argentina. Salta. 2015. Libro. Artículo Completo. Workshop. XVI Workshop de Investigadores en Ciencias de la Computación. Departamento de Informática de la Universidad Nacional de Salta.
- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; TARDIVO, MARÍA LAURA. Sistema Estadístico Evolutivo con Modelo de Islas: Calibración de parámetros evolutivos. Perú. Lima. 2015. Libro. Artículo Completo. Congreso. 6th International Symposium on Innovation and Technology (ISIT2015). pp. 77–81. International Institute of Innovation and Technology (IIITEC) y Universidad FASTA. ISBN: 978-612-45917- 9-2.
- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; TARDIVO, MARÍA LAURA; CAYMES SCUTARI, PAOLA. Sintonización Estática de Parámetros Evolutivos en Sistema para la Predicción de Incendios Forestales. Argentina. Mendoza. 2015. Libro. Resumen. Encuentro. VIII Encuentro de Investigadores y Docentes de Ingeniería (ENIDI). UTN-FRM, UTN- FRSR, UMAZA, UM y UNCuyo. 2015.
- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; TARDIVO, MARÍA LAURA. Método de Reducción de Incertidumbre Aplicado a la Predicción de Incendios: Análisis de Sintonización Estática. Matemática Aplicada, Computacional e Industrial. Vol. 5. pp. 241– 244. Santa Fé: ASAMACI (Asociación Argentina de Matemática Aplicada, Computacional e Industrial). 2015.
- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; TARDIVO, MARÍA LAURA. ESS-IM: Incremento de la calidad de predicción mediante la sintonización de parámetros evolutivos.

Argentina. Mendoza. 2016. Libro. Artículo Completo. Congreso. 11vo. Encuentro del International Center for Earth Sciences (E-ICES11). pp. 154–164. Comisión Nacional de Energía Atómica - CNEA. ISBN: 978-987-1323-50-0.

- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; TARDIVO, MARÍA LAURA. ESS- IM applied to forest fire spread prediction: Parameters tuning for a heterogeneous configuration. IEEE, Estados Unidos de América. New York. 2016. Revista. Artículo Completo. Conferencia. 35th International Conference of the Chilean Computer Science Society (SCCC). Pontificia Universidad Católica de Valparaíso, Chile y Universidad Técnica Federico Santa María, Valparaíso, Chile. ISBN: 978-1- 5090-3340-9. DOI: 10.1109/SCCC.2016.7836007.

Finalmente, en la última etapa del desarrollo de esta tesis, se implementó una versión híbrida basada en metaheurísticas evolutivas bajo una estrategia colaborativa basada en islas. El método HESSIM, además se implementó de forma heterogénea (a nivel de hardware), logrando de esta manera hacer más eficiente el método, los resultados obtenidos incrementan la calidad de predicción y eficiencia del método:

- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; TARDIVO, MARÍA LAURA. Método híbrido paralelo de reducción de incertidumbre aplicado a la predicción del comportamiento de incendios forestales. MECÁNICA COMPUTACIONAL. Córdoba: Asociación Argentina de Mecánica Computacional (AMCA). 2016 vol. XXXI, pp. 2857– 2869. ISSN: 1666-6070.
- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; TARDIVO, MARÍA LAURA. Método híbrido de reducción de incertidumbre aplicado a la predicción del comportamiento de incendios forestales. Argentina. San Luis. 2016. Libro. Artículo Completo. Congreso. XXII Congreso Argentino de Ciencias de la Computación. pp. 149–158. Red de Universidades con Carreras de Informática (RedUNCI) y UNSL, Facultad de Ciencias Físico Matemáticas y Naturales, Departamento de Informática.
- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; TARDIVO, MARÍA LAURA; CAYMES SCUTARI, PAOLA; GIL COSTA, VERÓNICA. Hybrid-

Parallel Uncertainty Reduction Method Applied to Forest Fire Spread Prediction. Journal of Computer Science & Technology. La Plata: Red de Universidades Nacionales con Carreras de Informática (RedUNCI). 2017 vol.17 Nro. 1. pp, 12–19. ISSN 1666-6038.

- MÉNDEZ GARABETTI, MIGUEL; BIANCHINI, GERMÁN; CAYMES SCUTARI, PAOLA; TARDIVO, MARÍA LAURA. HESSIM: evaluación de calidad de predicción aplicada al incendio forestal ocurrido en Queiriga Portugal. Argentina. Mendoza. 2017. Libro. Artículo Completo. Encuentro. 12 Encuentro del International Center for Earth Sciences (E-ICES12). pp. 339–348. Comisión Nacional de Energía Atómica - CNEA. ISBN 978-987-1323-61-6.

Tal como se comentó al inicio de esta sección, también se avanzó en el desarrollo de un sistema integral centrado en el uso de WSN e información satelital, para poder aplicarse a eventos en tiempo real:

- MÉNDEZ GARABETTI, M., BIANCHINI G., CAYMES-SCUTARI P., EL GUETA R., CAYLÁ I.B., CÓRDOBA D., PEÑASCO A., MORELLI J., CARRIZO M., GIORLANDO A., LUJAN A., CARBALLO N. Predicción y alerta temprana de incendios forestales mediante integración de información satelital y de campo en un método paralelo de reducción de incertidumbre. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste). pp. 195–206. 2018. ISBN 978-987-1323-63-0.

8.2. Cumplimiento de los objetivos planteados

Esta tesis planteó como objetivo general: *el desarrollo de métodos de reducción de incertidumbre capaces de asistir y participar en el proceso de resolución de problemas de alta magnitud y complejidad computacional, tal como la predicción y prevención de desastres naturales*. A continuación, se detallan los específicos y se describe cómo se cumplieron los mismos:

- **Diseño, implementación y evaluación de métodos de reducción de incertidumbre aplicados a la predicción y prevención de desastres naturales. Este objetivo implica un**

estudio profundo de la literatura existente en este campo y la definición de los casos de estudio de análisis:

Este amplio objetivo es abordado entre los Capítulos 1, 2, 3 y 4. En ellos se describe los conceptos esenciales de las técnicas de reducción de incertidumbre, los beneficios de su aplicación en fenómenos naturales, como así también se describe su utilización en la predicción del comportamiento de incendios forestales.

- **Inclusión de capacidad de operar con múltiples metaheurísticas. La hibridación de metaheurísticas supone un incremento en la capacidad de búsqueda de toda técnica de optimización. Debido a esto se propone dotar al método con la capacidad de operar de forma des centralizada con múltiples metaheurísticas de forma colaborativa:**

En el Capítulo 2 se analizan diferentes metaheurísticas poblacionales y de trayectoria, incluidos también los diferentes esquemas de paralelización. Los detalles de la metaheurística híbrida implementada basada en un esquema de comunicación basado en islas son abordados en el Capítulo 4.

- **Identificación de los fenómenos naturales a trabajar, para así obtener datos de casos reales de acuerdo al modelo seleccionado. Adaptación de datos reales para su aplicación en los métodos desarrollados:**

Las posibles aplicaciones de los métodos de reducción de incertidumbre, junto con el fenómeno natural a trabajar se tratan en el Capítulo 1. Para la obtención de datos de casos reales se trabajó con el procesamiento de imágenes satelitales, tratado en el Capítulo 7.

- **Validación de los resultados obtenidos realizando una experimentación detallada aplicada a casos reales y contrastada con otras metodologías:**

En el Capítulo 5, se presenta una validación de los métodos descriptos aplicándolos a casos de quemas controladas como así también a incendios forestales de gran magnitud. El contraste con otras metodologías se lleva a cabo mediante su comparación con otros desarrollos previos, del mismo equipo de investigación.

En función de lo expuesto en este trabajo, y a lo arriba descrito se considera haber cumplido con el objetivo general y los objetivos específicos planteados al inicio de esta tesis. También, de acuerdo a los resultados obtenidos y a las contribuciones realizadas, se espera que este trabajo pueda continuar su desarrollo acercándolo cada vez más a

ser un producto implementado para la toma de decisiones que permita mejorar la gestión de desastres y la minimización de daños.

8.3. Líneas de trabajo futuro

En este apartado final, se describen algunas líneas de trabajo futuro que se desprenden de la presente tesis:

- Llevar a cabo estudios de sintonización que permitan sacar el máximo provecho a los métodos diseñados. En este sentido, es de interés llevar a cabo estudios complementarios de sintonización, analizando cómo afectan (tanto en calidad de predicción como en tiempo de procesamiento) la variación de la totalidad de los parámetros de forma simultánea (parámetros de cada metaheurística). También, evaluar el uso de técnicas de metaoptimización, sintonización dinámica, donde se evalúen los efectos de las variaciones de cada parámetro y las interdependencias entre ellos de forma automática [224].
- Explorar otras alternativas de hibridación entre metaheurísticas. En la literatura existen numerosos antecedentes donde se utilizan metaheurísticas híbridas basadas en métodos poblacionales y de trayectoria que han sido aplicados de forma exitosa en gran variedad de problemas.
- Optimizar el funcionamiento de los métodos para la operación de casos de incendios reales. Se considera importante enfocar los esfuerzos de trabajo en la optimización de los métodos para su funcionamiento en casos de incendios reales. Ya sea añadiendo nuevas funcionalidades, modelos, datos, o mediante estudios de sintonización. De esta manera se espera poder aproximar, en mejor medida, las predicciones a la realidad.
- Desarrollar plataforma de visualización: Pensando en herramientas para la toma de decisiones, es fundamental contar con un módulo que permita visualizar y/o monitorear el frente de fuego a medida que las predicciones se realizan.
- Culminar la integración entre los métodos de predicción y el sistema de alerta temprana y el módulo de generación automática de mapas. Esta línea se encuentra con un alto grado de avance y se considera oportuno culminar este desarrollo. La funcionalidad final permitirá que a partir de una alerta (ya sea proveniente de la WSN o a partir de la detección de un foco de calor MODIS) se inicie el proceso de seguimiento del evento. Generando mapas adaptados a la entrada del método de

predicción. De esta manera el método llevará a cabo la predicción de forma automática entregando la información de predicción a los tomadores de decisión.

Bibliografía

- [1] Stanford University and Center for the Study of Language and Information (U.S.), *Stanford encyclopedia of philosophy*. Stanford University, 1997.
- [2] T. Ishigami and T. Homma, “An importance quantification technique in uncertainty analysis for computer models,” in *Proceedings of the ISUMA 90’, First International Symposium on Uncertainty Modelling and Analysis*, pp. 398–403, IEEE Comput. Soc. Press, 1990.
- [3] T. Sano, T. Takeda, and M. Yamasaki, “A new uncertainty reduction method for fuel fabrication process with erbia-bearing fuel,” *Journal of Nuclear Science and Technology*, vol. 46, pp. 226–231, mar 2009.
- [4] Z. Hu, S. Mahadevan, and D. Ao, “Uncertainty aggregation and reduction in structure–material performance prediction,” *Computational Mechanics*, vol. 61, pp. 237–257, feb 2018.
- [5] X. Qian, W. Li, and M. Yang, “Two-stage nested optimization-based uncertainty propagation method for uncertainty reduction,” in *Communications in Computer and Information Science*, vol. 402, pp. 243–252, Springer, Berlin, Heidelberg, 2013.
- [6] M. Li, J. Hamel, and S. Azarm, “Optimal uncertainty reduction for multi disciplinary multi-output systems using sensitivity analysis,” *Structural and Multidisciplinary Optimization*, vol. 40, pp. 77–96, jan 2010.
- [7] A. Taghribi and S. Sharifian, “A Metaheuristically Tuned Interval Type 2 Fuzzy System to Reduce Segmentation Uncertainty in Brain MRI Images,” *Journal of Medical Systems*, vol. 41, p. 174, nov 2017.
- [8] I. Purica and D. Bae, “Climate Change events induced risk assessment and mapping and a potential insurance policy,” *Journal of Earth Science & Climatic Change*, vol. 07, no. 09, 2016.
- [9] E. Bartezzaghi, R. Verganti, and G. Zotteri, “A simulation framework for forecasting uncertain lumpy demand,” *International Journal of Production Economics*, vol. 59, pp. 499–510, mar 1999.
- [10] A. V. Kalyuzhnaya and A. V. Boukhanovsky, “Computational Uncertainty Management for Coastal Flood Prevention System,” *Procedia Computer Science*, vol. 51, pp. 2317–2326, jan 2015.

- [11] M. P. Thompson and D. E. Calkin, “Uncertainty and risk in wildland fire management: A review,” *Journal of Environmental Management*, vol. 92, pp. 1895–1909, aug 2011.
- [12] H. A. Simon, *Science seeks parsimony, not simplicity: searching for pattern in phenomena*, p. 32–72. Cambridge University Press, 2002.
- [13] G. Characteristics, “Confessions of a pragmatic statistician,” *The Statistician*, vol. 51, no. 2, pp. 40–45, 2001.
- [14] RAE, “Real Academia Española,” 2001.
- [15] M. R. Arahál, M. Berenguel Soria, and F. Rodríguez Días, *Técnicas de predicción con aplicaciones en ingeniería*. Sevilla: Secretariado de Publicaciones, Universidad de Sevilla, 2006.
- [16] L. Cornejo-Bueno, E. Garrido-Merchán, D. Hernández-Lobato, and S. Salcedo Sanz, “Bayesian optimization of a hybrid system for robust ocean wave features prediction,” *Neurocomputing*, vol. 275, pp. 818–828, jan 2018.
- [17] J. Zhang, S. Cui, Y. Xu, Q. Li, and T. Li, “A novel data-driven stock price trend prediction system,” *Expert Systems with Applications*, vol. 97, pp. 60–69, may 2018.
- [18] E. Howard, M. Meehan, and A. Parnell, “Contrasting prediction methods for early warning systems at undergraduate level,” *The Internet and Higher Education*, vol. 37, pp. 66–75, apr 2018.
- [19] X. Li, C. Yu, L. Ju, J. Qin, Y. Zhang, L. Dou, and S. Yuqing, “Position prediction system based on spatio-temporal regularity of object mobility,” *Information Systems*, vol. 75, pp. 43–55, jun 2018.
- [20] L. Oneto, E. Fumeo, G. Clerico, R. Canepa, F. Papa, C. Dambra, N. Maz zino, and D. Anguita, “Train Delay Prediction Systems: A Big Data Analytics Perspective,” *Big Data Research*, may 2017.
- [21] J. Sanz, D. Paternain, M. Galar, J. Fernandez, D. Reyero, and T. Belzunegui, “A new survival status prediction system for severe trauma patients based on a multiple classifier system,” *Computer Methods and Programs in Biomedicine*, vol. 142, pp. 1–8, apr 2017.
- [22] S. K. M. Hamouda, M. E. Wahed, R. H. Abo Alez, and K. Riad, “Robust breast cancer prediction system based on rough set theory at National Cancer Institute of Egypt,” *Computer Methods and Programs in Biomedicine*, vol. 153, pp. 259–268, jan 2018.
- [23] C. Soize, *Uncertainty Quantification: An Accelerated Course with Advanced Applications in Computational Engineering*. 2016.

- [24] P. J. Schoemaker, “Forecasting and Scenario Planning: The Challenges of Uncertainty and Complexity,” in *Blackwell Handbook of Judgment and Decision Making*, pp. 274–296, Malden, MA, USA: Blackwell Publishing Ltd, 2008.
- [25] P. J. Schoemaker, “Multiple scenario development: Its conceptual and behavioral foundation,” *Strategic Management Journal*, vol. 14, pp. 193–213, mar 1993.
- [26] A. Farguella, J. Moré, A. Cortés, J. Miró, T. Margalef, and V. Altava, “Reducing Data Uncertainty in Surface Meteorology Using Data Assimilation: A Comparison Study,” *Procedia Computer Science*, vol. 80, pp. 1846–1855, jan 2016.
- [27] T. G. Doeswijk, *Reducing Prediction Uncertainty of Weather Controlled Systems*. Ph.d. thesis, Wageningen University, The Netherlands, 2007.
- [28] D. W. Hyndman and D. W. Hyndman, *Natural Hazards and Disasters*. Brooks/Cole, Cengage Learning, 2014.
- [29] M. Banco, “Los desastres naturales empujan a 26 millones de personas a la pobreza y provocan pérdidas por USD 520 000 millones al año, según un nuevo análisis del Banco Mundial.” <http://www.bancomundial.org/es/news/pressrelease/2016/11/14/natural-disasters-force-26-million-people-into-poverty-and-cost-520bn-in-losses-every-year-new-world-bank-analysis-finds>, 2016. Accedido 2018-04-30.
- [30] T. Glade and D. E. Alexander, “Classification of natural disasters,” in *Encyclopedia of Earth Sciences Series*, pp. 78–82, Springer Netherlands, 2013.
- [31] D. Alexander, *Natural disasters / David Alexander*. New York, New York, USA: CRC Press, 1 ed., 2001.
- [32] D. Alexander, *Natural disasters*. UCL Press, 1993.
- [33] C. M. Lewis, “Albala-bertrand j. m., the political economy of large natural disasters with special reference to developing countries (oxford: Clarendon press, 1993), pp. vi 259 £35.00.,” *Journal of Latin American Studies*, vol. 26, no. 1, p. 272–273, 1994.
- [34] “Indicadores de Riesgo de Desastres Naturales y de Gestión de Riesgos,” tech. rep., BID, Banco Interamericano de Desarrollo, 2010.
- [35] W. Hough, “Fire and Human Civilization,” *Proceedings of the American Philosophical Society*, vol. 71, no. 6, pp. 403–406, 1932.

- [36] D. M. Molina, J. Blanco, M. Galán, E. Pous, J. B. García, and D. García, *Incendios forestales: fundamentos, lecciones aprendidas y retos de futuro*. AI FEMA, 2010.
- [37] S. Graham, “Global Fire Monitoring” <https://earthobservatory.nasa.gov/Features/GlobalFire/firoct> 1999. Accedido 2018-04-30.
- [38] W. Bond and J. Keeley, “Fire as a global ‘herbivore’: the ecology and evolution of flammable ecosystems,” *Trends in Ecology & Evolution*, vol. 20, pp. 387–394, jul 2005.
- [39] P. Morgan, C. C. Hardy, T. W. Swetnam, M. G. Rollins, and D. G. Long, “Mapping fire regimes across time and space: Understanding coarse and fine scale fire patterns,” *International Journal of Wildland Fire*, vol. 10, pp. 329–342, 2001.
- [40] *Estadísticas de Incendios Forestales 2016*. Buenos Aires: Presidencia de la Nación, 2016.
- [41] Greenpeace, “Cifras actualizadas: 2,5 millones de hectáreas quemadas en el Amazonas - Greenpeace Argentina,” 2019.
- [42] . Diario Los Andes, “El fuego ya arrasó 90.000 hectáreas de campos en Alvear y ahora afecta a San Rafael,” 2018.
- [43] . Diario La Voz, “El incendio en Cerro Mogote ya afectó más de 200 hectáreas y sigue el riesgo extremo” | Noticias al instante desde LAVOZ.com.ar | La Voz,” 2018.
- [44] . Diario La Nación, “Tras 11 días, logran controlar el incendio forestal en San Luis,” 2016.
- [45] G. D. Papadopoulos and F.-N. Pavlidou, “A Comparative Review on Wildfire Simulators,” *IEEE Systems Journal*, vol. 5, pp. 233–243, jun 2011.
- [46] L. G. Zárate López, *Estudio de las características físicas y geométricas de la llama en los incendios forestales*. PhD thesis, Universitat Politècnica de Catalunya, may 2004.
- [47] R. Rothermel, *A mathematical model for predicting fire spread in wildland fuels*, vol. II. Res. Pap. INT-115, US Dept. of Agric., Forest Service, Intermountain Forest and Range Experiment Station. (Ogden, UT.), 1972.
- [48] O. Nelles, *Nonlinear system identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Springer, 2001.

- [49] G. Wallace, “A Numerical Fire Simulation-Model,” *International Journal of Wildland Fire*, vol. 3, no. 2, p. 111, 1993.
- [50] M. A. Finney, “FARSITE : Fire Area Simulator — Model Development and Evaluation,” USDA *Forest Service Research Paper*, no. February, p. 47, 1998.
- [51] P. L. Andrews, C. D. Bevins, R. C. Seli, A. Andrews, P. L. . Bevins, and C. D. . Seli, “BehavePlus fire modeling system, version 4.0: User’s Guide Revised,” 2008.
- [52] F. A. Heinsch and P. L. Andrews, “BehavePlus fire modeling system, version 5.0: Design and Features,” Tech. Rep. 249, 2010.
- [53] A. Lopes, M. Cruz, and D. Viegas, “FireStation — an integrated software system for the numerical simulation of fire spread on complex topography,” *Environmental Modelling & Software*, vol. 17, pp. 269–285, jan 2002.
- [54] L. Ribeiro, D. Viegas, A. Lopes, P. Mangana, and P. Moura, “Operational application of a decision support tool in fire management in Portugal,” *Forest Ecology and Management*, vol. 234, no. Supplement 1, p. 243, 2006.
- [55] R. Buyya, *High Performance Cluster Computing: Architectures and Systems*. Prentice Hall, PTR, NJ, USA, 1999.
- [56] L. N. Long, “Parallel Computing,” *Seminar*, p. 134, 2008.
- [57] M. J. M. J. Quinn and M. J. M. J. Quinn, *Parallel computing : theory and practice*. McGraw-Hill, 1994.
- [58] S. H. Leong and D. Kranzlmüller, “Towards a General Definition of Urgent Computing,” *Procedia Computer Science*, vol. 51, pp. 2337–2346, jan 2015.
- [59] M. J. Flynn, “Selected Solutions to Computer Architecture: Pipelined and Parallel Processor Design,” tech. rep., 1995.
- [60] M. J. Flynn and K. W. Rudd, “Parallel Architectures,” tech. rep., 1996.
- [61] A. Grama, A. Gupta, G. Karypis, and V. Kumar, *Introduction to Parallel Computing*. Pearson/Prentice Hall, 2 ed., 2003.
- [62] B. Wilkinson and C. M. Allen, *Parallel Programming: Techniques and Ap plications Using Networked Workstations and Parallel Computers*. Pear son/Prentice Hall, 1999.

- [63] G. Bianchini, *Wildland fire prediction based on statistical analysis of multiple solutions*. Ph. d thesis, Universidad Autonoma de Barcelona, jul 2006.
- [64] K. Sörensen, M. Sevaux, and F. Glover, “A history of metaheuristics,” in *Handbook of Heuristics*, vol. 2-2, pp. 791–808, 2018.
- [65] P. Siarry, “Introduction,” in *Metaheuristics*, pp. 1–18, Cham: Springer International Publishing, 2016.
- [66] E.-G. Talbi, *Metaheuristics: from design to implementation*. John Wiley & Sons, 2009.
- [67] L. Bui, H. Abbass, and J. Branke, “Multiobjective optimization for dynamic environments,” *2005 IEEE Congress on Evolutionary Computation*, vol. 3, pp. 2349–2356, 2005.
- [68] M. R. Garey and D. S. Johnson, *Computers and intractability : a guide to the theory of NP-completeness*. W.H. Freeman, 1979.
- [69] X. Yu and M. Gen, *Introduction to Evolutionary Algorithms*, vol. 0 of *Decision Engineering*. London: Springer London, 2010.
- [70] C. Cruz, J. R. Gonzalez, and D. A. Pelta, “Optimization in dynamic environments: A survey on problems, methods and measures,” *Soft Computing*, vol. 15, pp. 1427–1448, jul 2011.
- [71] G. Chiandussi, M. Codegone, S. Ferrero, and F. Varesio, “Comparison of multiobjective optimization methodologies for engineering applications,” *Computers & Mathematics with Applications*, vol. 63, pp. 912–942, mar 2012.
- [72] I. H. Osman and J. P. Kelly, “Meta-Heuristics: An Overview,” in *Meta-Heuristics*, pp. 1–21, Boston, MA: Springer US, 1996.
- [73] E. L. Lawler, *Combinatorial Optimization : Networks and Matroids*. 1976.
- [74] Y. Wang, “A Sociopsychological Perspective on Collective Intelligence in Metaheuristic Computing,” *International Journal of Applied Metaheuristic Computing (IJAMC)*, vol. 1, no. 1, pp. 110–128, 2010.
- [75] Geem Z.W., Kim J.H., and G. Loganathan, “A New Heuristic Optimization Algorithm: Harmony Search,” *SIMULATION*, vol. 76, pp. 60–68, feb 2001.

- [76] C. Blum and A. Roli, “Metaheuristics in combinatorial optimization,” *ACM Computing Surveys*, vol. 35, pp. 268–308, sep 2003.
- [77] F. Glover and G. A. Kochenberger, *Handbook of Metaheuristics*, vol. 57. Kluwer Academic Publishers, 2003.
- [78] J. H. Holland, *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control and artificial intelligence*. MIT Press, 1975.
- [79] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. 1986.
- [80] Y. Wang, “On Contemporary Denotational Mathematics for Computational Intelligence,” in *Transactions on Computational Science II*, pp. 6–29, Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.
- [81] Y. Wang, “A Cognitive Informatics Reference Model of Autonomous Agent Systems (AAS),” *International Journal of Cognitive Informatics and Natural Intelligence*, vol. 3, no. 1, pp. 1–16, 2009.
- [82] T. Crainic and M. Toulouse, “Parallel strategies for metaheuristics,” *Handbook of Metaheuristics*, pp. 475–513, 2003.
- [83] J. Branke, *Evolutionary Optimization in Dynamic Environments*, vol. 1. Springer US, 2002.
- [84] C. Blum, D. Merkle, and (Eds.), *Swarm Intelligence. Introduction and Applications*. Springer, 2008.
- [85] C. Darwin, *On the origins of species by means of natural selection*. 1859.
- [86] A. E. Eiben and J. E. Smith, *Introduction to Evolutionary Computing*. Natural Computing Series, Berlin, Heidelberg: Springer Berlin Heidelberg, 2003.
- [87] R. Storn and K. Price, “Differential Evolution - A simple and efficient adaptive scheme for global optimization over continuous spaces,” tech. rep., TR- 95-012, ICSI, Berkeley, CA, 1995.
- [88] J. Kennedy and R. Eberhart, “Particle swarm optimization,” *Proceedings of ICNN’95 - International Conference on Neural Networks*, vol. 4, pp. 1942– 1948, 1995.

- [89] T. Bartz-Beielstein, D. Blum, and J. Branke, “Particle Swarm Optimization and Sequential Sampling in Noisy Environments,” in *Metaheuristics*, pp. 261–273, Boston, MA: Springer US, 2007.
- [90] V. W. V. W. Porto, R. C. IEEE Neural Networks Council., and C. Conference on Evolutionary Programming (7th : 1998 : San Diego, *Evolutionary programming VII : 7th international conference, EP98, San Diego, California, USA, March 25-27, 1998 : proceedings*. Springer, 1998.
- [91] Y. Shi and R. Eberhart, “A Modified particle swarm optimizer,” in *Proceedings of the IEEE Conference on Evolutionary Computation, ICEC*, pp. 69–73, IEEE, 1998.
- [92] M. Dorigo, *Optimization, Learning and Natural Algorithms*. PhD thesis, Politecnico di Milano, Italian, 1992.
- [93] E. Alba, G. Luque, and S. Nesmachnow, “Parallel metaheuristics: recent advances and new trends,” *International Transactions in Operational Research*, vol. 20, pp. 1–48, jan 2013.
- [94] E. Cantú-Paz, “A Survey of Parallel Genetic Algorithms,” *Calculateurs Parallèles, Réseaux et Systems Repartis*, vol. 10, 1998.
- [95] T. G. Mattson, B. A. Sanders, and B. Massingill, *Patterns for parallel programming*. Addison-Wesley, 2005.
- [96] E. Alba and J. M. Troya, “A survey of parallel distributed genetic algorithms,” *Complexity*, vol. 4, pp. 31–52, mar 1999.
- [97] Y.-J. Gong, W.-N. Chen, Z.-H. Zhan, J. Zhang, Y. Li, Q. Zhang, and J.-J. Li, “Distributed evolutionary algorithms and their models: A survey of the state-of-the-art,” *Applied Soft Computing*, vol. 34, pp. 286–300, sep 2015.
- [98] B. Abdalhaq, *A Methodology to enhance the prediction of forest fire propagation*. Universitat Autònoma de Barcelona, nov 2004.
- [99] K. Beven and A. Binley, “The future of distributed models: Model calibration and uncertainty prediction,” *Hydrological Processes*, vol. 6, pp. 279–298, jul 1992.
- [100] G. Bianchini, M. Denham, A. Cortés, T. Margalef, and E. Luque, “Wildland fire growth prediction method based on Multiple Overlapping Solution,” *Journal of Computational Science*, vol. 1, pp. 229–237, dec 2010.
- [101] G. Bianchini, P. Caymes-Scutari, and M. Méndez-Garabetti, “Evolutionary Statistical System: A parallel method for improving

forest fire spread prediction,” *Journal of Computational Science*, vol. 6, pp. 58–66, jan 2015.

[102] M. Méndez-Garabetti, G. Bianchini, P. Caymes-Scutari, and M. Tardivo, “In crease in the quality of the prediction of a computational wildfire behavior method through the improvement of the internal metaheuristic,” *Fire Safety Journal*, vol. 82, pp. 49–62, 2016.

[103] M. L. Tardivo, P. Caymes-Scutari, G. Bianchini, M. Méndez-Garabetti, A. Cencerrado, and A. Cortés, “A comparative study of evolutionary statistical methods for uncertainty reduction in forest fire propagation prediction,” *Procedia Computer Science*, vol. 108, pp. 2018–2027, jan 2017.

[104] M. Méndez-Garabetti, G. Bianchini, M. L. Tardivo, P. G. Caymes Scutari, and G. V. Gil Costa, “Hybrid-Parallel Uncertainty Reduction Method Applied to Forest Fire Spread Prediction,” *Journal of Computer Science and Technology*, ISSN-e 1666-6038, vol. 17, no. 1, pp. 12–19, 2017.

[105] F. Bouttier and P. Courtier, “Data assimilation concepts and methods,” 2002.

[106] F. Darema, “Introduction to the ICCS 2007 workshop on dynamic data driven applications systems,” in *Computational Science - ICCS 2007, 7th International Conference Beijing, China, May 27-30, 2007, Proceedings, Part I*, pp. 955–962, 2007.

[107] F. Darema, “Dynamic data driven applications systems: A new paradigm for application simulations and measurements,” in *Computational Science - ICCS 2004, 4th International Conference, Kraków, Poland, June 6-9, 2004, Proceedings, Part III*, pp. 662–669, 2004.

[108] D. Montgomery and G. Runger, *Applied Statistics and Probability for Engineers*, vol. 6th Edition. Limusa Wiley & Sons, New Jersey, 2014.

[109] Bevins, Collin D., “fireLib User Manual and Technical Reference,” tech. rep., Intermountain Research Station, Forest Service, U.S. Department of Agriculture., 1996.

[110] A. Torraza, *Desarrollo del motor de cálculo de un simulador de incendios forestales*. Tesis de maestría, Universidad Nacional de Córdoba, 2014.

[111] L. S. Bradshaw, J. E. Deeming, R. E. Burgan, and J. D. Cohen, “The 1978 National Fire-Danger Rating System: technical

documentation,” tech. rep., United States Department of Agriculture Forest Service, 1984.

[112] J. H. Scott, “A new set of standard fuel models for use with Rothermel’s spread model,” 2014.

[113] W. L. Fons, “Analysis of fire spread in light forest fuels,” *Journal of Agricultural Research*, vol. 72, no. 13, pp. 93–121, 1946.

[114] H. W. Emmons, “Fire in the forest,” *Fire Research Abstracts and Reviews*, vol. 5, pp. 163–178, 1964.

[115] A. G. McArthur, *Weather and Grassland Fire Behaviour*. Leaflet / Forestry and Timber Bureau, Canberra: Forest Research Institute, Forestry and Timber Bureau, 1966.

[116] C. Van Wagner, “Calculations on forest fire spread by flame radiation,” *Canadian Dept. of Forestry*, p. 18, 1967.

[117] P. H. Thomas, “Some aspects of the growth and spread of fire in the open,” *Forestry: An International Journal of Forest Research*, vol. 40, no. 2, pp. 139–164, 1967.

[118] H. E. Anderson, “Heat transfer and fire spread,” tech. rep., Forest Service, U.S. Dept. of Agriculture, Ogden, Utah, 1969.

[119] W. H. Frandsen, “Fire spread through porous fuels from the conservation of energy,” *Combustion and Flame*, vol. 16, no. 1, pp. 9–16, 1971.

[120] R. C. Rothermel, “A mathematical model for predicting fire spread in wildland fuels,” tech. rep., U.S. Department of Agriculture, Intermountain Forest and Range Experiment Station, Ogden, Utah, 1972.

[121] P. J. Pagni and T. G. Peterson, “Flame spread through porous fuels,” *Symposium (International) on Combustion*, vol. 14, no. 1, pp. 1099–1107, 1973.

[122] H. P. Telisin, “Flame radiation as a mechanism of fire spread in forests,” in *Heat transfer in flames* (N. Afgan and J. Beer, eds.), pp. 441–449, New York: John Wiley, 1974.

[123] F. R. Steward, “Fire spread through a fuel bed,” in *Heat transfer in fires: thermophysics, social aspects, economic impact* (P. Blackshear, ed.), pp. 317–378, scripta bo ed., 1974.

[124] H. M. Cekirge, “Propagation of fire fronts in forests,” *Computers and Mathematics with Applications*, vol. 4, no. 4, pp. 325–332, 1978.

- [125] L. Trabaud, “Étude du comportement du feu dans la Garrigue de Chêne kermès à partir des températures et des vitesses de propagation,” *Annales Sciences Forestières*, vol. 36, no. 1, pp. 13–38, 1979.
- [126] N. Fujii, J. Hasegawa, L. Pallop, and Y. Sakawa, “A nonstationary model of firespreading,” *Applied Mathematics and Modelling*, vol. 4, pp. 176–180, 1980.
- [127] I. R. Noble, G. A. B. Bary, and A. M. Gill, “McArthur’s fire-danger meters expressed as equations,” *Australian Journal of Ecology*, vol. 5, pp. 201–203, 1980.
- [128] A. M. Grishin, A. D. Gruzin, and V. G. Zverev, “Mathematical modelling of the spreading of high-level forest fires,” *Sov. Phys. Dokl*, vol. 28, pp. 328–330, 1983.
- [129] C. C. Hwang and Y. Xie, “Flame propagation along matchstick arrays on inclined base boards,” *Combustion Science and Technology*, vol. 42, pp. 1–12, dec 1984.
- [130] N. J. De Mestre, E. A. Catchpole, D. H. Anderson, and R. C. Rothermel, “Uniform Propagation of a Planar Fire front Without Wind,” *Combustion Science and Technology*, vol. 65, pp. 231–244, jun 1989.
- [131] R. O. Weber, “Thermal Theory for Determining the Burning Velocity of a Laminar Flame, Using the Inflection Point in the Temperature Profile,” *Combustion Science and Technology*, vol. 64, pp. 135–139, mar 1989.
- [132] A. M. A. M. Grishin, *Mathematical modeling of forest fires and new methods of fighting them*. Tomsk: Publishing House of the Tomsk State University, 1997.
- [133] J.-L. Dupuy, *Mieux comprendre et prédire la propagation des feux de forêts : expérimentation, test et proposition de modèles*. Thèse de doctorat en sciences, Université Claude Bernard, jan 1997.
- [134] R. Linn, *A transport model for prediction of wildfire behavior*. Master, University of California, United States, jul 1997.
- [135] P. A. Santoni and J. H. Balbi, “Modelling of two-dimensional flame spread across a sloping fuel bed,” *Fire Safety Journal*, vol. 31, pp. 201–225, oct 1998.
- [136] V. DX, R. PR, and M. L, “An empirical model for the spread of a fireline inclined in relation to the slope gradient or to wind direction,” in *Proceedings of the 3rd international conference on forest fire*

research. *14th Conference on Fire and Forest Meteorology*, (Luso, Portugal), pp. 325–342, 1998.

[137] N. P. Cheney, J. S. Gould, and W. R. Catchpole, “Prediction of fire spread in grasslands,” *International Journal of Wildland Fire*, vol. 8, no. 1, pp. 1–13, 1998.

[138] M. Larini, F. Giroud, B. Porterie, and J.-C. Loraud, “A multiphase formula tion for fire propagation in heterogeneous combustible media,” *International Journal of Heat and Mass Transfer*, vol. 41, pp. 881–897, mar 1998

[139] P. L. Andrews, “BEHAVE : Fire Behavior Prediction and Fuel Modeling System-BURN Subsystem, Part 1,” tech. rep., USDA Forest Service, Inter mountain Forest and Range Experiment Station, 1986.

[140] R. E. Burgan and R. C. Rothermel, “BEHAVE: Fire Behavior Prediction and Fuel Modeling Subsystem– FUEL Subsystem,” tech. rep., National Wildfire Coordinating Group Sponsored, 1984.

[141] P. L. Andrews, “BehavePlus fire modeling system: past, present, and future,” in *7th symposium on fire and meteorological society*, p. 13, 2007.

[142] R. Linn and F. Harlow, “Firetec: A transport description of wildfire behavior,” 12 1997.

[143] J. H. Furman and R. R. Linn, “What is FIRETEC (and why should I care)?,” *Fire Management Today*, vol. 76, no. 3, pp. 33–36, 2018.

[144] Forestry Canada Fire Danger Group, “Development of the Canadian Forest Fire Behavior Prediction System - Information Report ST-X-3,” *Forestry Canada Fire Danger Group*, vol. 3, p. 66, 1992.

[145] M. A. Finney, “An Overview of FlamMap Fire Modeling Capabilities,” in *USDA Forest Service Proceedings RMRS-P-41*, pp. 213–220, 2006.

[146] D. Carlton, P. Andrews, and C. Bevins, “BehavePlus Fire Modeling System, User’s Guide, Version 1.0,” *USDA Forest Service General Technical Report*, no. December, p. 108, 2001.

[147] P. L. Andrews, “Current status and future needs of the BehavePlus Fire Mode ling System,” *International Journal of Wildland Fire*, vol. 23, no. 1, pp. 21–33, 2014.

[148] J. Mandel, J. D. Beezley, J. L. Coen, and M. Kim, “Data Assimilation for Wildland Fires: Ensemble kalman filters in coupled atmosphere-surface mo dels,” *IEEE Control Systems*, vol. 29, pp. 47–65, jun 2009.

- [149] C. Skamarock, B. Klemp, J. Dudhia, O. Gill, D. Barker, G. Duda, X.-y. Huang, W. Wang, and G. Powers, “A Description of the Advanced Research WRF Version 3,” 2008.
- [150] T. L. Clark, J. Coen, and D. Latham, “Description of a coupled atmosphere fire model,” *International Journal of Wildland Fire*, vol. 13, no. 1, pp. 49–63, 2004.
- [151] J. L. Coen, “Simulation of the Big Elk Fire using coupled atmosphere - fire modeling,” *International Journal of Wildland Fire*, vol. 14, p. 49, apr 2005.
- [152] F. Charte, A. J. Rivera, M. J. del Jesus, and F. Herrera, “MLSMOTE: Approaching imbalanced multilabel learning through synthetic instance generation,” *Knowledge-Based Systems*, vol. 89, pp. 385–397, nov 2015.
- [153] J. Mandel, J. D. Beezley, and A. K. Kochanski, “Coupled atmosphere-wildland fire modeling with WRF 3.3 and SFIRE 2011,” *Geoscientific Model Development*, vol. 4, pp. 591–610, jul 2011.
- [154] E. G. Patton and J. L. Coen, “WRF-Fire: A Coupled Atmosphere-Fire Module for WRF,” tech. rep.
- [155] R. Real and J. M. Vargas, “The Probabilistic Basis of Jaccard’s Index of Similarity,” *Systematic Biology*, vol. 45, pp. 380–385, sep 1996.
- [156] W. Gropp, E. Lusk, and A. Skjellum, *Using MPI : portable parallel programming with the Message-Passing-Interface*. 2014.
- [157] D. Kirk and W.-m. Hwu, *Programming massively parallel processors : a hands on approach*.
- [158] F. Sousa, R. dos Reis, and J. Pereira, “Simulation of surface fire fronts using fireLib and GPUs,” *Environmental Modelling & Software*, vol. 38, pp. 167–177, dec 2012.
- [159] L. Dagum and R. Menon, “OpenMP: an industry standard API for shared memory programming,” *IEEE Computational Science and Engineering*, vol. 5, no. 1, pp. 46–55, 1998.
- [160] D. Guo, C. Zhou, and M. Liu, “A HERO EVOLUTIONARY ALGORITHM HYBRIDIZING FROM PSO AND GA,” in *Computational Methods*, pp. 1075– 1080, Dordrecht: Springer Netherlands, 2007.
- [161] H. Garg, “A hybrid PSO-GA algorithm for constrained optimization problems,” *Applied Mathematics and Computation*, vol. 274, pp. 292–305, feb 2016.

- [162] R. E. Johnson and B. Foote, “Designing reusable classes,” *Journal of Object Oriented Programming*, vol. 1, no. 2, pp. 22–30, 35, 1988.
- [163] D. X. Viegas, “Project Spread - Forest Fire Spread Prevention and Mitigation,” 2004.
- [164] L. LA NACIÓN, “Una nevada extinguió el fuego en Córdoba - LA NACION.”
- [165] B. Gabbert, “Mendocino Complex of Fires grows larger across three counties - Wildfire Today,” 2007.
- [166] M. N. Forest, “Mendocino Complex Public Information Map 9.14.2018 - Inci Web the Incident Information System,” 2018.
- [167] J. J. Dongarra, *The sourcebook of parallel computing*, vol. 41. Morgan Kaufmann Publishers, 2002.
- [168] F. G. Lobo, C. F. Lima, and Z. Michalewicz, *Parameter setting in evolutionary algorithms*. Springer, 2010.
- [169] M. D. McKay, R. J. Beckman, and W. J. Conover, “A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code,” *Technometrics*, vol. 21, p. 239, may 1979.
- [170] D. C. Montgomery, *Design and analysis of experiments*, vol. 3. New York: John Wiley & Sons, Ltd, jul 1987.
- [171] M. Loopez-Ibañez, J. Dubois-Lacoste, T. Stützle, and M. Birattari, “The irace Package : Iterated Race for Automatic Algorithm Configuration,” *Management*, no. February, p. 7288, 2011.
- [172] H. R. Lourenço, O. C. Martin, and T. Stützle, “Iterated Local Search: Framework and Applications,” in *Handbook of Metaheuristics* (M. Gendreau and J.-Y. Potvin, eds.), pp. 363–397, Boston, MA: Springer US, 2010.
- [173] F. Hutter, H. H. Hoos, and K. Leyton-Brown, “Sequential Model-Based Optimization for General Algorithm Configuration,” in *Proceedings of the 5th international conference on Learning and Intelligent Optimization*, pp. 507– 523, Springer-Verlag, 2011.
- [174] M. Mendez-Garabetti, G. Bianchini, P. Caymes Scutari, and M. L. Tardivo, “Sistema Estadístico Evolutivo con Modelo de Islas :,” in *6TH International Symposium on Innovation and Technology - ISIT2015 Mar del Plata – Argentina* (I. I. o. I. FASTA and T. I. y Universidad, eds.), (Mar del Plata), pp. 77–81, ISIT, 2015.

- [175] R Core Team, “R: A Language and Environment for Statistical Computing,” 1993.
- [176] E. Alba, F. Luna, and A. J. Nebro, “Parallel heterogeneous genetic algorithms for continuous optimization,” in *Proceedings - International Parallel and Distributed Processing Symposium, IPDPS 2003*, p. 7, Institute of Electrical and Electronics Engineers Inc., 2003.
- [177] E. Alba, F. Luna, and A. J. Nebro, “Parallel heterogeneous genetic algorithms for continuous optimization,” in *Proceedings - International Parallel and Distributed Processing Symposium, IPDPS 2003*, vol. 14, pp. 317–333, 2003.
- [178] A. Volokitina, M. Sofronov, and T. Sofronova, “Topical scientific and practical issues of wildland fire problem,” *Mitigation and Adaptation Strategies for Global Change*, vol. 13, pp. 661–674, aug 2008.
- [179] Fleming and Joe, “Technology & Development Program Fire Management TechTips,” tech. rep., 2003.
- [180] P. Baronti, P. Pillai, V. W. Chook, S. Chessa, A. Gotta, and Y. F. Hu, “Wireless sensor networks: A survey on the state of the art and the 802.15.4 and ZigBee standards,” may 2007.
- [181] B. Leblon, “Monitoring forest fire danger with remote sensing,” jul 2005.
- [182] A. Molina-Pico, D. Cuesta-Frau, A. Araujo, J. Alejandro, and A. Rozas, “Forest Monitoring and Wildland Early Fire Detection by a Hierarchical Wireless Sensor Network,” *Journal of Sensors*, vol. 2016, pp. 1–8, feb 2016.
- [183] Y. E. Aslan, I. Korpeoglu, and Ö. Ulusoy, “A framework for use of wireless sensor networks in forest fire detection and monitoring,” *Computers, Environment and Urban Systems*, vol. 36, pp. 614–625, 2012.
- [184] D. M. Doolin and N. Sitar, “Wireless sensors for wildfire monitoring,” in *Smart Structures and Materials 2005: Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, vol. 5765, p. 477, 2005.
- [185] M. A. Fischer, C. M. Di Bella, and E. G. Jobbágy, “Fire patterns in central semiarid Argentina,” *Journal of Arid Environments*, vol. 78, pp. 161–168, mar 2012.
- [186] G. Hmimina, E. Dufrêne, J. Y. Pontailleur, N. Delpierre, M. Aubinet, B. Caquet, A. de Grandcourt, B. Burban, C. Flechard, A.

Granier, P. Gross, B. Hei nesch, B. Longdoz, C. Moureaux, J. M. Ourcival, S. Rambal, L. Saint André, and K. Soudani, “Evaluation of the potential of MODIS satellite data to predict vegetation phenology in different biomes: An investigation using ground based NDVI measurements,” *Remote Sensing of Environment*, 2013.

[187] W. Xu, M. J. Wooster, G. Roberts, and P. Freeborn, “New GOES imager algorithms for cloud and active fire detection and fire radiative power assessment across North, South and Central America,” *Remote Sensing of Environment*, 2010.

[188] A. A. Van De Griend and M. Owe, “On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces,” *International Journal of Remote Sensing*, 1993.

[189] W. Schroeder, E. Prins, L. Giglio, I. Csiszar, C. Schmidt, J. Morisette, and D. Morton, “Validation of GOES and MODIS active fire detection products using ASTER and ETM+ data,” *Remote Sensing of Environment*, vol. 112, pp. 2711–2726, may 2008.

[190] C. F. Waigl, M. Stuefer, A. Prakash, and C. Ichoku, “Detecting high and low-intensity fires in Alaska using VIIRS I-band data: An improved operational approach for high latitudes,” *Remote Sensing of Environment*, vol. 199, pp. 389–400, sep 2017.

[191] T. J. Hawbaker, V. C. Radeloff, A. D. Syphard, Z. Zhu, and S. I. Stewart, “Detection rates of the MODIS active fire product in the United States,” *Remote Sensing of Environment*, vol. 112, pp. 2656–2664, may 2008.

[192] L. Giglio, T. Loboda, D. P. Roy, B. Quayle, and C. O. Justice, “Remote Sensing of Environment An active- fire based burned area mapping algorithm for the MODIS sensor,” *Remote Sensing of Environment*, vol. 113, no. 2, pp. 408–420, 2009.

[193] G. Van Der Werf, J. Randerson, G. Collatz, L. Giglio, P. Kasibhatla, A. Arellano, S. Olsen, and E. Kasischke, “Continental-scale partitioning of fire emissions during the 1997 to 2001 el niño/la niña period,” *Science*, vol. 303, pp. 73– 76, 1 2004.

[194] L. Giglio, G. R. V. D. Werf, J. T. Randerson, G. J. Collatz, and P. Kasib hatla, “and Physics Global estimation of burned area using MODIS active fire observations,” no. 1, pp. 957–974, 2006.

[195] “Moderate Resolution Imaging Spectroradiometer (MODIS) | Earthdata.”

- [196] N. Vaiman, *Comparación de índices climáticos y espectrales en la estimación de rendimiento de maíz y soja a nivel departamental en Entre Ríos Facultad de Ciencias Exactas y Naturales Facultad de Agronomía Maestría en Meteorología Agrícola Comparación de índices clim.* PhD thesis, Universidad de Buenos Aires, 2018.
- [197] A. Daham, D. Han, M. Rico-Ramirez, and A. Marsh, “Analysis of NVDI variability in response to precipitation and air temperature in different regions of Iraq, using MODIS vegetation indices,” *Environmental Earth Sciences*, vol. 77, pp. 1–24, may 2018.
- [198] “MODIS Web.”
- [199] P. O. Pav, E. Chuvieco, and S. Catedr, “QUEMADA CON IMÁGENES DEL SENSOR ENVISAT-MERIS EN LA CUENCA,” 2009.
- [200] “Live Satellite Viewing. Browse real-time Earth observation..” (Citado en 129.) [201] “MODIS Web.”
- [202] “EarthExplorer.”
- [203] S. Blumberg, “Explore Earth Your Way with Worldview,” 2018.
- [204] J. H. Zhang, F. M. Yao, C. Liu, L. M. Yang, and V. K. Boken, “Detection, emission estimation and risk prediction of forest fires in China using satellite sensors and simulation models in the past three decades-An overview,” *International Journal of Environmental Research and Public Health*, vol. 8, no. 8, pp. 3156–3178, 2011.
- [205] “GDAL — GDAL documentation.”
- [206] F. Warmerdam, “The Geospatial Data Abstraction Library,” in *Open Source Approaches in Spatial Data Handling*, pp. 87–104, Berlin, Heidelberg: Springer Berlin Heidelberg, sep 2008.
- [207] OSGeo, “The Open Source Geospatial Foundation - OSGeo.”
- [208] “GitHub - nasa-gibs/worldview: Interactive interface for browsing global, full resolution satellite imagery.”
- [209] F. Warmerdam and E. Rouault, “GDAL — GDAL documentation,” 2019.
- [210] “Una nevada extinguió el fuego en Córdoba - LA NACION.”
- [211] “Informe Anual de Incendios Forestales del Año 2018,” 2008.
- [212] “Mendocino Complex Archives - Wildfire Today.”
- [213] “Mendocino fire has become California’s largest ever - The Washington Post.”

- [214] “Mendocino fire has become California’s largest ever - The Washington Post.”
- [215] “River Fire Archives - Wildfire Today.”
- [216] C. P. Walteneus Dargie, *Fundamentals of Wireless Sensor Networks: Theory and Practice*. Wiley, 2010. (Citado en 146.)
- [217] M. Á. Serna, A. Bermúdez, and R. Casado, “Modelado de incendios forestales con WSNs mediante múltiples envolventes,” in *Actas de las XXIII Jornadas de Paralelismo*, 2012.
- [218] C. M. Ramya, M. Shanmugaraj, and R. Prabakaran, “Study on ZigBee technology,” in *ICECT 2011 - 2011 3rd International Conference on Electronics Computer Technology*, vol. 6, pp. 297–301, 2011.
- [219] A. Koubâa, M. Alves, and E. Tovar, “IEEE 802.15.4 for Wireless Sensor Networks: A Technical Overview,” *Architecture*, no. July, 2005.
- [220] I. Aad and C. Castelluccia, “Differentiation mechanisms for IEEE 802.11,” in *Proceedings - IEEE INFOCOM*, vol. 1, pp. 209–218, 2001.
- [221] J. M. Marais, R. Malekian, and A. M. Abu-Mahfouz, “LoRa and LoRaWAN testbeds: A review,” in *2017 IEEE AFRICON: Science, Technology and Innovation for Africa, AFRICON 2017*, pp. 1496–1501, Institute of Electrical and Electronics Engineers Inc., nov 2017.
- [222] Pleva GmbH, “DHT11 Humidity & Temperature sensor,” *Melliand Textilberichte*, vol. 76, no. 12, p. 1112, 1995.
- [223] S. pandey, G. Kumar, and Shailesh, “Forest fire detection using GSM and MQ7 (gas sensor),” *International journal of embedded systems and applications*, 2011.
- [224] M. Camilleri and M. Montebello, “Optimising the meta-optimiser in machi ne learning problems,” in *ACM International Conference Proceeding Series*, vol. Part F128357, (New York, New York, USA), pp. 15–22, Association for Computing Machinery, feb 2017.



Esta red se constituyó formalmente en noviembre de 1996 y actualmente 51 universidades argentinas son miembros activos. Sus objetivos son:

“Coordinar actividades académicas relacionadas con el perfeccionamiento docente, la actualización curricular y la utilización de recursos compartidos en el apoyo al desarrollo de las carreras de Ciencia de la Computación y/o Informática en Argentina”.

“Establecer un marco de colaboración para el desarrollo de las actividades de posgrado en Ciencia de la Computación y/o Informática de modo de optimizar la asignación y el aprovechamiento de recursos”.

La presente tesis doctoral aborda la problemática de la incertidumbre existente en todo sistema de predicción, focalizando en el desarrollo de métodos de reducción de incertidumbre aplicados a la predicción de fenómenos naturales. Debido a que estos fenómenos suelen causar gran impacto, los métodos propuestos han sido desarrollados utilizando capacidades de alto rendimiento. El primer método desarrollado en esta tesis (ESS-IM), comenzó con el objetivo de lograr una mejora a una metodología previamente desarrollada denominada ESS (Sistema Estadístico Evolutivo). Se trabajó en el incremento del paralelismo de la metaheurística interna, incorporando una arquitectura basada en modelo de islas bajo un esquema de migración. Este desarrollo logró incrementar la capacidad de búsqueda de la metaheurística interna, impactando de forma directa en un incremento en la calidad de predicción del método. El método HESS-IM, se implementó de forma heterogénea (a nivel de hardware), logrando que los resultados obtenidos incrementen la calidad de predicción y eficiencia del método.