

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INFORMÁTICA
VALPARAÍSO - CHILE



**PARTICLE SWARM OPTIMIZATION PARA EL AJUSTE DE
MODELOS PROBABILÍSTICOS A DATOS DEL VIENTO EN
VALPARAÍSO**

ALONSO JAVIER SANDOVAL ACEVEDO

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL EN INFORMÁTICA

PROFESOR GUÍA : MARÍA CRISTINA RIFF
PROFESOR CORREFERENTE : ELIZABETH MONTERO URETA

Noviembre 2016

Resumen

La velocidad y la dirección del viento son factores esenciales en el estudio de este fenómeno meteorológico. Diversos trabajos realizados en las áreas de energía eólica, medio ambiente y desastres naturales proponen métodos para modelar estas características del viento. La distribución de Weibull es ampliamente utilizada para modelar la velocidad, mientras que la mezcla de distribuciones de von Mises es sugerida para la dirección. Comúnmente, para el ajuste de los parámetros de estas distribuciones a ciertos datos recolectados, se utilizan métodos numéricos, sin embargo, estos pueden presentar deficiencias en cuanto a tiempos de ejecución o calidad de solución. En esta memoria, se implementa la meta-heurística *Particle Swarm Optimization* para el ajuste de los parámetros de las distribuciones mencionadas a los datos del viento de la comuna de Valparaíso. Esta técnica provee de buenas soluciones en cortos tiempos de ejecución, siendo adaptable para el ajuste de ambas distribuciones. Además, se propone una estrategia para el control de la convergencia prematura del PSO, mejorando el trabajo realizado a la fecha.

Abstact

The speed and direction of the wind are essential factors in the study of this meteorological phenomenon. Many works performed in the areas of wind energy, environment and natural disasters propose methods to model these wind properties. The Weibull distribution is widely used to model velocity, while the mixture of von Mises distribution is suggested for direction. Commonly, numerical methods are used for the adjustment of the parameters of these distributions to certain data collected, however, these can present deficiencies in terms of execution times or quality solution. In this report, the meta-heuristic Particle Swarm Optimization is implemented for the adjustment of the parameters of the mentioned distributions to the data of the wind of the commune of Valparaiso. This technique provides good solutions in short execution times, being adaptable for the adjustment of both distributions. Furthermore, a strategy is proposed for the control of the premature convergence of the PSO algorithm, improving the work done to date.

Índice general

1. Introducción	1
1.1. Identificación del problema	1
1.2. Objetivos	3
1.2.1. Objetivos específicos	3
1.3. Estructura del documento	3
2. Estado del arte	5
2.1. Particle Swarm Optimization	5
2.2. Velocidad del viento	7
2.2.1. Distribución de Weibull	7
2.2.2. Métodos numéricos	8
2.2.3. Particle Swarm Optimization	11
2.3. Dirección del viento	12
2.4. Resumen del capítulo	16
3. PSO para la velocidad del viento en Valparaíso	17
3.1. Modelo Matemático	17
3.2. Estructura del PSO	17
3.2.1. Representación	18
3.2.2. Consideración de los parámetros	19
3.2.3. Descripción del algoritmo	20
3.3. Resultados	20
3.3.1. Experimentos	20
3.3.2. Análisis de los resultados	23
3.4. Conclusiones del capítulo	30
4. PSO para la dirección del viento en Valparaíso	31
4.1. Modelo Matemático	31
4.2. Estructura del PSO	33
4.2.1. Representación	34
4.2.2. Consideración de los parámetros	34
4.2.3. Descripción del algoritmo	36
4.3. Resultados	38
4.3.1. Experimentos	38
4.3.2. Análisis de los resultados	40

4.4. Conclusiones del capítulo	46
5. Aplicaciones	51
5.1. Esquema de uso del algoritmo	51
5.2. Uso de los resultados	52
5.2.1. Energía eólica	52
5.2.2. Propagación de incendios	54
5.2.3. Propagación de pesticidas	55
5.2.4. Análisis atmosférico	56
5.2.5. Resumen del capítulo	57
6. Conclusiones	58
Bibliography	60

Índice de figuras

2.1. Movimiento de una partícula	6
2.2. Función de distribución de probabilidad de Weibull	8
2.3. Distribución de Weibull con histograma - Maracanaú	12
2.4. Modelo de ajuste FVMV para suroeste y noreste en la estación Mersing .	14
2.5. Ajuste dirección del viento, aeropuerto St. John. Creado por Heckenberge- rova et al [5].	15
3.1. Ejemplo colección de datos Enero Valparaíso 2015	22
3.2. Superficie de datos viento de Valparaíso. Fuente: Elaboración Propia. . .	24
3.3. Ajuste con PSO a datos del viento de Valparaíso. Fuente: Elaboración Propia.	25
3.4. Ajuste con PSO a datos Valparaíso 2015, 2014 y 2013, baja y buena calidad. Fuente: Elaboración Propia.	26
3.5. Ajuste con PSO a registros del viento en Valparaíso (Con todos los datos). Fuente: Elaboración Propia.	27
3.6. Ajuste con PSO a datos Valparaíso 2015, por rango de meses. Fuente: Elaboración Propia.	28
3.7. Ajuste con PSO a datos (cifras puras) Valparaíso 2015, 2014, 2013	29
4.1. Pruebas iniciales. Fuente: Elaboración propia.	40
4.2. Graficos de ajustes anuales. Fuente: Elaboración propia.	41
4.3. Gráficos de ajuste de MVM por meses. Fuente: Elaboración propia. . . .	42
4.4. Gráficos de ajuste de MVM por meses. Fuente: Elaboración propia. . . .	43
4.5. Gráficos de ajuste por meses en coordenadas polares. Fuente: Elaboración propia.	47
4.6. Gráficos de ajuste por meses en coordenadas polares. Fuente: Elaboración propia.	48
4.7. Comparación de variaciones en el PSO. Fuente: Elaboración propia. . . .	49
4.8. Comparación de variaciones en el PSO. Fuente: Elaboración propia. . . .	50
5.1. Esquema de uso del algoritmo	53
5.2. Ejemplo de zonas a evaluar en la velocidad del viento. Fuente: Elaboración propia.	54
5.3. Ejemplo emplazamiento turbinas eólicas. Fuente: Jorge Mírez [23]. . . .	55
5.4. Principales procesos que interactúan en el comportamiento de los pesticidas después de su aplicación. Fuente: Bedos [26].	56

Índice de tablas

3.1. Tabla de pruebas, PSO velocidad del viento 29

4.1. Tabla de pruebas, PSO dirección del viento 44

Capítulo 1

Introducción

1.1. Identificación del problema

El viento es uno de los fenómenos meteorológicos más comunes de las zonas costeras de Valparaíso. Su presencia incide en distintos aspectos del medio ambiente, como las condiciones del clima, la sensación térmica, la propagación de los pesticidas en plantaciones, algunos desastres naturales como los incendios, entre otros. Esto atrae el interés de investigadores dedicados a estudiar este fenómeno, con el fin de poder controlar las variables que condicionan su comportamiento, de manera de poder predecir los fenómenos subyacentes a este.

Entre los diversos efectos del viento, se encuentra su capacidad de mover objetos, lo cual ha permitido al hombre aprovechar esta propiedad para convertir la energía cinética del viento en energía eléctrica. Dicha fuente de energía es conocida como energía eólica, ampliamente calificada como renovable y poco contaminante.

Acorde a esto, en los últimos años, tanto el gobierno de Chile como la ciudadanía, han mostrado un creciente interés en el uso de las fuentes de energías renovables y con poco impacto ambiental, por lo que distintos proyectos en la materia han sido llevados a cabo, desde estudios de factibilidad y recopilación de datos hasta el emplazamiento de las primeras centrales de fuentes de energía limpia. El año 2014, el ministerio de energía publicó un extenso reporte acerca de la situación actual del país en materias de energías renovables, en donde se pueden ver distintas proyecciones y estimaciones de implementación de posibles plantas de generación a lo largo del país. [1]

Cualquiera sea el motivo de estudio de las características del viento (en particular el tema energético), existen dos variables fundamentales a considerar: su velocidad y su dirección. Ejemplo de ello es el estudio realizado sobre el potencial eléctrico en la provincia de Bushehr, Irán por Dabbaghiyan et al. [2], o el trabajo de Chang en Taiwan [3].

Generalmente, se requiere previo al análisis de estas dos características recopilar mediciones del viento de la zona a estudiar. Actualmente estos datos están disponibles y son obtenidos por diversos centros meteorológicos a lo largo del país. Para esta memoria los datos fueron proporcionados por el servicio meteorológico de la Armada de Chile, que cuenta con el equipo necesario para registrar el comportamiento del viento a lo largo de las distintas épocas del año y en diferentes zonas de Chile. Los datos con los que se trabajó son de la comuna de Valparaíso entre los años 2013, 2014 y 2015.

El objetivo es proponer modelos basados en los datos históricos del viento con el fin de obtener la distribución de su velocidad y su dirección para, por ejemplo, poder evaluar el potencial eléctrico de cierta zona o prevenir la propagación de incendios.

En la literatura, son ampliamente aceptadas la distribución de *Weibull* para modelar el conjunto de datos de velocidad y la distribución de von Mises para el modelo de la dirección del viento. Ambas distribuciones probabilistas requieren de la determinación de parámetros para que el modelo se ajuste a los datos obtenidos. La elección del método para poder encontrar los parámetros de ajustes definirá la calidad de los modelos, por ello, diversas técnicas han sido presentadas en la literatura, las cuales tienen mayor o menor precisión, dependiendo de las características de los datos. Comúnmente, se utilizan métodos numéricos para la determinación de estos parámetros, sin embargo, estudios recientes han abordado nuevas estrategias utilizando métodos heurísticos, en particular la meta-heurística *Particle Swarm Optimization* (en adelante también referido como PSO), con la cual se ha logrado mejorar la calidad del modelo para la velocidad del viento, como se expone en el caso de estudio en el noreste de Brasil por Carneiro et. al. [4]. Por otra parte, para la dirección del viento también se ha propuesto obtener los parámetros de ajuste a través del uso de PSO, como se explica en Heckenbergerova et al. [5]. La ventaja del uso de PSO es que es un algoritmo de uso general, de fácil implementación, cuyo tiempo de ejecución es mucho menor que los métodos numéricos tradicionales.

Por lo anterior, el trabajo a realizar en esta memoria se centra en usar PSO para encontrar los parámetros de ajuste de la distribución de *Weibull* con datos de la zona de la comuna de Valparaíso y abordar la estrategia propuesta por Heckenbergerova et al. [5] para el ajuste

del modelo de dirección del viento.

1.2. Objetivos

Aplicar métodos actuales de optimización basados en la meta-heurística *Particle Swarm Optimization* para el ajuste de modelos de dirección y velocidad del viento a los datos recopilados del viento en Valparaíso, con el fin de presentar resultados que permitan inferir información precisa acerca de las condiciones de la región para la generación de energía eólica, prevención de la propagación de incendios, entre otras potenciales aplicaciones.

1.2.1. Objetivos específicos

1. Implementar un algoritmo basado en *Particle Swarm Optimization* para **encontrar** los parámetros de un modelo probabilístico que se ajusten a los datos de la velocidad del viento en Valparaíso.
2. Implementar un algoritmo basado en *Particle Swarm Optimization* para **optimizar** los parámetros de un modelo probabilístico que se ajusten a los datos de la dirección del viento en Valparaíso.
3. Evaluar los modelos sobre los datos del viento para validar la propuesta realizada.

1.3. Estructura del documento

La estructura del documento tiene la siguiente forma:

1. El **Capítulo 2** resume el trabajo actual encontrado en la literatura acerca de la meta-heurística *Particle Swarm Optimization* y las técnicas existentes para hallar los parámetros de ajuste de las distribuciones de densidad de probabilidad de Weibull y la *mixture of von Mises distribution*.
2. El **Capítulo 3** explica la implementación del PSO para el ajuste de la función de densidad de probabilidad de Weibull a los datos de velocidad del viento.

3. El **Capítulo 4** explica la implementación del PSO para el ajuste de la función de densidad de probabilidad de von Mises (*mixture of von Mises distribution*) a los datos de dirección del viento.
4. El **Capítulo 5** expone algunas de las potenciales aplicaciones de los algoritmos propuestos para el ajuste de modelos en datos de velocidad y dirección del viento.
5. El **Capítulo 6** muestra las conclusiones generales obtenidas en esta memoria.

Capítulo 2

Estado del arte

En este capítulo se presenta una revisión de la literatura sobre la meta-heurística *Particle Swarm Optimization*, con el fin de exponer el desarrollo actual de la técnica. Posteriormente, se revisarán los trabajos en donde se hayan utilizado las distribuciones de Weibull y von Mises para el acople de datos del viento, haciendo énfasis en aquellos que utilicen PSO para encontrar los parámetros de ajuste.

2.1. Particle Swarm Optimization

Como se introduce en el artículo de Kaveh [6], el algoritmo *Particle Swarm Optimization* es una meta-heurística inspirada en las observaciones de la naturaleza acerca del comportamiento social de poblaciones de enjambres. Esta abstracción está basada en la interacción en grupo de seres vivos. Por ejemplo, las gaviotas suelen volar en conjuntos cerca del mar, en búsqueda de zonas donde hayan alimento (peces), moviéndose individualmente, pero a la vez, siendo influenciadas por el grupo. El método simula la conducta de los individuos a través de partículas que se mueven dentro de un espacio (rango de posibles valores conocido como espacio de búsqueda), siendo éstas afectadas por factores individuales (conocimiento propio del lugar donde se está) y factores colectivos (conocimiento del enjambre del mejor lugar encontrado por el grupo), los cuales dirigen el movimiento de éstos a zonas escogidas por una función objetivo (*fitness function*).

Para cada partícula su vector posición \vec{x} representa una solución candidata, la cual varía dentro del espacio de búsqueda a velocidad \vec{v} . Después de varias iteraciones, el enjambre o conjunto de partículas se irá concentrando en aquellas zonas donde las posiciones obtengan

mejores puntajes al ser evaluadas por la función objetivo.

El modelo clásico presentado por Kennedy y Eberhart [7], describe la variación de la velocidad y de la posición de las partículas como se presenta a continuación:

$$v_{i,j}^{k+1} = v_{i,j}^k + c_1 r_1 (x_{best_i}^k - x_{i,j}^k) + c_2 r_2 (x_{gbest}^k - x_{i,j}^k) \quad (2.1)$$

$$x_{i,j}^{k+1} = x_{i,j}^k + v_{i,j}^{k+1} \quad (2.2)$$

Como se explica en Kaveh [6] $x_{i,j}^k$ y $v_{i,j}^k$ son la j -ésima componente de la posición y la velocidad de la partícula i respectivamente en la iteración o tiempo k , r_1 y r_2 son número aleatorios uniformes en el rango 0 a 1, x_{best_i} y x_{gbest} representan las mejores soluciones alcanzadas por la partícula i y por el vecindario respectivamente, c_1 y c_2 son parámetros que representan la confianza en la solución individual de la partícula (parámetro cognitivo) y la incidencia del aspecto colectivo o solución global (parámetro social), respectivamente. Un esquema de la interacción de estos componentes se aprecia en la figura 2.1

El modelo clásico presentado tiene ciertas deficiencias, en particular la forma en que se

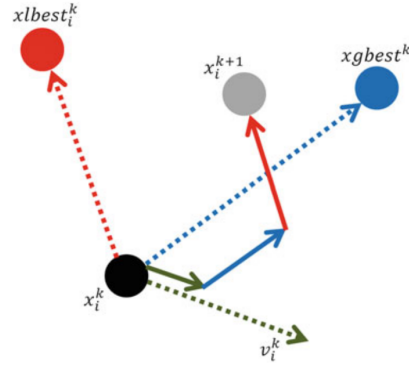


Figura 2.1: Movimiento de una partícula

Creado por Kaveh [6].

actualiza la velocidad no considera una ponderación para el aporte de la velocidad previa en una partícula. Para corregir esto, simplemente se añade un factor que escala la velocidad previa. Como se explica en Kaveh [6], este peso que se añade a la velocidad previa permite tener un control más apropiado respecto al comportamiento del enjambre. En términos extremos, si la velocidad previa se elimina (peso nulo), las partículas quedan atrapadas en la región local de partida, pero si se le da demasiado peso, las partículas son reacias a converger a las buenas posiciones conocidas.

Es por esto que la forma del PSO actual tiene un parámetro w que representa la incidencia

de la velocidad previa en la velocidad actual de la partícula (parámetro llamado factor de inercia). Nuevamente, se tiene que la partícula actualiza su velocidad de la siguiente forma:

$$v_{i,j}^{k+1} = wv_{i,j}^k + c_1r_1(xbest_{i,j}^k - x_{i,j}^k) + c_2r_2(xgbest_j^k - x_{i,j}^k) \quad (2.3)$$

$$x_{i,j}^{k+1} = x_{i,j}^k + v_{i,j}^{k+1} \quad (2.4)$$

Donde w es el factor conocido como “inercia” de la partícula, y regula la incidencia de la velocidad previa en la actual, tal como podría esperarse de una partícula en aceleración.

A modo de complemento, en el trabajo inicialmente citado, se presenta una revisión completa del estado del arte del método *Particle Swarm Optimization* en términos de diseño o arquitectura del algoritmo, donde se exponen las distintas modificaciones y alternativas existentes en la literatura que pretenden mejorar aspectos como:

1. Configuración de parámetros (inercia, cognitivo, social, aleatorios).
2. Problemas asociados a la convergencia prematura.
3. Estructura de algoritmo o topologías que modifican la comunicación entre partículas (o la incidencia de las soluciones globales y particulares).
4. Sesgos en la búsqueda por la forma de la región o por la interacción de las partículas (operadores de combinación como el promedio, que tienden a centrar la búsqueda en determinada región).
5. Algoritmos híbridos con PSO.
6. Versión discreta del PSO.

2.2. Velocidad del viento

2.2.1. Distribución de Weibull

Dado un conjunto de datos de velocidad obtenidos de la medición del viento, se puede crear un histograma que represente la frecuencia de éstos. A partir de esto, es posible ajustar un modelo probabilístico (distribución de densidad de probabilidad) que explique el comportamiento de las velocidades del viento medido. Dicho modelo comúnmente se

basa en la distribución de Weibull, la cual es ampliamente aceptada por la comunidad dedicada al estudio meteorológico, tal y como se menciona en el trabajo de Carneiro et al. [4], Kongnam et al. [8], Dabbaghiyan et al. [2], Fadare [9], Weisser [10] y Chang [3]. En el trabajo realizado por Carneiro et al. [4], se describe la distribución de Weibull como:

$$f_{weibull}(v) = \frac{k}{c} \cdot \left(\frac{v}{c}\right)^{k-1} \cdot e^{-\left(\frac{v}{c}\right)^k} \quad (2.5)$$

Donde k y c son los parámetros de ajuste que representan la forma y la escala (o amplitud) de la distribución respectivamente, y v es el valor de la velocidad del viento a la que el modelo asociará una determinada densidad. Un ejemplo de cómo varía la forma de esta distribución se aprecia en la figura 2.2, en donde se ven distintas curvas de Weibull, con diferentes parámetros k manteniendo c constante.

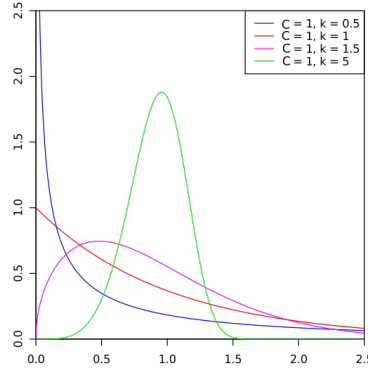


Figura 2.2: Función de distribución de probabilidad de Weibull
Adaptación propia desde [11].

2.2.2. Métodos numéricos

Tradicionalmente, se utilizan métodos numéricos para estimar los parámetros de la distribución de Weibull. En el artículo de Chang [12], se realiza una comparación de seis métodos numéricos comúnmente utilizados para la obtención de k y c . A continuación, se describen brevemente estos métodos:

1. **The Moment:** Se basa en la iteración numérica de las siguientes dos ecuaciones:

$$\bar{v} = c\Gamma(1 + \frac{1}{k}) \quad (2.6)$$

$$\sigma = c[\Gamma(1 + \frac{2}{k}) - \Gamma^2(1 + \frac{1}{k})]^{\frac{1}{2}} \quad (2.7)$$

Donde \bar{v} es el promedio y σ la desviación estándar de los datos de velocidad del viento.

2. **Empirical:** Considerado un caso especial del método del momento. Los parámetros son calculados de la siguiente forma:

$$k = (\frac{\sigma}{\bar{v}})^{-1,086} \quad (2.8)$$

$$c = \frac{\bar{v}}{\Gamma(1 + \frac{1}{k})} \quad (2.9)$$

3. **Graphical:** Se ajustan rectas a los datos de velocidad del viento usando mínimos cuadrados. Con una doble transformación logarítmica, la función de distribución acumulativa queda:

$$\ln\{-\ln[1 - F(v)]\} = k \ln(v) - k \ln(c) \quad (2.10)$$

Realizando un gráfico para $\ln(v)$ en vez de $\ln(-\ln(1 - F(v)))$, la pendiente de la recta que se ajusta mejor a los pares de datos es el parámetro de la forma de la distribución de Weibull. El parámetro de escala se obtiene por la intersección con la coordenada y .

4. **Maximum likelihood:** En este método, son necesarias muchas iteraciones. Los parámetros de Weibull están dado por:

$$k = [\frac{\sum_{i=1}^n v_i^k \ln(v_i)}{\sum_{i=1}^n v_i^k} - \frac{\sum_{i=1}^n \ln(v_i)}{n}]^{(-1)} \quad (2.11)$$

$$c = (\frac{1}{n} \sum_{i=1}^n v_i^k)^{\frac{1}{k}} \quad (2.12)$$

Donde v_i es la velocidad del viento en el paso i y n es el número de puntos de datos distintos de cero.

5. **Modified maximum likelihood:** Este método es utilizado si es que se tiene disponible datos de velocidad del viento en una distribución de frecuencias. Los parámetros de Weibull son calculados como:

$$k = \left[\frac{\sum_{i=1}^n v_i^k \ln(v_i) f(v_i)}{\sum_{i=1}^n v_i^k f(v_i)} - \frac{\sum_{i=1}^n \ln(v_i) f(v_i)}{f(v \geq 0)} \right]^{-1} \quad (2.13)$$

$$c = \left[\frac{1}{f(v \geq 0)} \sum_{i=1}^n v_i^k f(v_i) \right]^{1/k} \quad (2.14)$$

Donde v_i es la velocidad del viento central al intervalo i , n es el número de intervalos, $f(v_i)$ es la frecuencia de la velocidad del viento dentro del intervalo i y $f(v \geq 0)$ la probabilidad de que la velocidad del viento sea mayor o igual a cero.

6. **Energy pattern factor method:** El factor del patrón de energía es definido como:

$$E_{pf} = \frac{\bar{v}^3}{\bar{v}^3} \quad (2.15)$$

Donde \bar{v}^3 es el promedio de las velocidades del viento cúbicas. Los parámetros de Weibull pueden ser calculados como:

$$k = 1 + \frac{3,69}{E_{pf}^2} \quad (2.16)$$

$$c = \frac{\bar{v}}{\Gamma(1 + \frac{1}{k})} \quad (2.17)$$

Estos métodos fueron comparados a través de pruebas de desempeño, usando una simulación basada en el método de Montecarlo. El análisis de los datos del viento fue desarrollado bajo criterios tales como el test Kolmogorov-Smirnov, *parameter error*, *root mean square error* y el error de energía del viento. De ello, bajo distintas condiciones, ciertos métodos se comportan mejor que otros al momento de ajustar la distribución de Weibull a los datos de prueba.

En la búsqueda de nuevas alternativas surge una propuesta para mejorar el ajuste de la función de distribución de probabilidad a los datos, a través del uso de la meta-heurística *Particle Swarm Optimization*, propuesta que mejora la calidad de los resultados en comparación con los métodos numéricos presentados.

2.2.3. Particle Swarm Optimization

En Carneiro et al. [4], se realiza un caso de estudio de las características del viento en las zonas costeras de Parnaíba y Maracanaú, y en una zona interior, Petrolina, en Brasil. Allí se explica la necesidad de obtener un modelo para el comportamiento estocástico del viento, de manera de poder evaluar el potencial energético de aquellas regiones. Tal y como se señaló anteriormente, el modelo utilizado para este propósito es la distribución de Weibull. Para poder utilizar dicha distribución, es necesario ajustar los parámetros del modelo a los datos recolectados.

En el estudio mencionado, se propone un *Particle Swarm Optimization* para encontrar los parámetros k y c de la distribución de Weibull y a su vez corroborar que la calidad de la solución encontrada por el PSO, comparada con los métodos numéricos tradicionales, es mejor.

Así, se define la función de aptitud para el PSO como:

$$\epsilon(v_i) = \frac{1}{2} \sum_{i=0}^n (f_{real}(v_i) - f_{weibull}(v_i))^2 \quad (2.18)$$

Donde ϵ , es el error cuadrático a minimizar entre los valores del histograma de los datos recolectados y la función de distribución de Weibull.

El PSO utilizado es el modelo clásico presentado en la sección anterior, considerando los parámetros w , $c1$ y $c2$, sin embargo, para controlar la convergencia del método, se establece que estos parámetros varíen durante la ejecución del algoritmo dentro de un rango definido ($w \in \{0,4, 0,9\}$, $c1$ y $c2 \in \{0, 2,5\}$), aumentando los factores de inercia y cognitivo, y disminuyendo el factor social a medida que se avanza en las iteraciones. Esto se traduce en que se explotan las soluciones iniciales, para posteriormente, en las iteraciones finales explorar en una vecindad local al óptimo encontrado hasta el momento.

Finalmente, para evaluar los resultados de la propuesta, se compara el PSO con cinco de los seis métodos numéricos mencionados en la sección 2.2.2 utilizados para la estimación de los parámetros de Weibull: *Moment Method* (M), *Energy Method* (E), *Energy Pattern Factor Method* (EPF), *Energy Equivalent Method* (EE) y *Maximum Likelihood* (ML). Además, para medir la calidad de cada método, se utilizan tres *test* estadísticos: *correlation* (r), *relative bias* (RB) y *root mean square error* (RMSE).

Los resultados que se exponen en el trabajo citado demuestran que mediante la meta-heurística *Particle Swarm Optimization* se pueden obtener los mejores parámetros de ajuste para la distribución de Weibull. Un ejemplo de esto es expuesto en la figura 2.3.

En Kongnam et al. [8], la meta-heurística *Particle Swarm Optimization* es utilizada para el

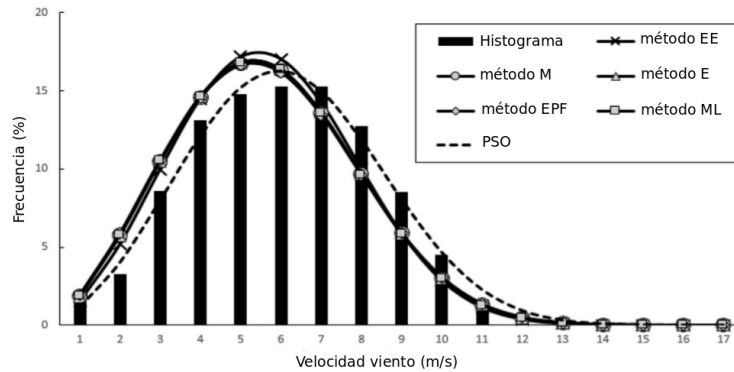


Figura 2.3: Distribución de Weibull con histograma - Maracanaú
Creado por Carneiro et al. [4].

problema del control de la velocidad de las turbinas de viento en una torre eólica, con el fin de maximizar la generación de energía. En dicho trabajo, se utiliza la distribución de Weibull para el modelado de la velocidad del viento. La construcción del PSO es llevada a cabo considerando el problema de la convergencia prematura, por lo que se desarrollan funciones que hacen variar estos parámetros a lo largo de la ejecución, de manera de evitar que el algoritmo se estanque en las iteraciones iniciales.

2.3. Dirección del viento

En el trabajo acerca del modelado del comportamiento de la dirección del viento en Malasia [13], se propone que para identificar la(s) dirección(es) dominante del viento se utilice la función de densidad de probabilidad *finite von Mises-Fisher* (FVMF) ajustada a las mediciones obtenidas. Estos datos acerca del viento fueron obtenidos desde cinco estaciones meteorológicas ubicadas en distintas zonas en la península de Malasia.

La FVMF, o la función de densidad de probabilidad de von Mises como también se le

llamará más adelante, de forma genérica, está definida de la siguiente forma:

$$f(x; \mu_h, k_h) = \sum_{h=1}^H (w_h) \frac{k_h^{\frac{d}{2}-1}}{(2\pi)^{\frac{d}{2}} I_{\frac{d}{2}-1}(k)} e^{(k_h \mu_h^T x)} \quad (2.19)$$

Donde $x = [\cos(\theta_i), \sin(\theta_i)]$, $\frac{k_h^{\frac{d}{2}-1}}{(2\pi)^{\frac{d}{2}} I_{\frac{d}{2}-1}(k)}$ es una constante de normalización, d es la dimensión del vector aleatorio x ($d = 2$, para este caso), μ_h es el parámetro de dirección predominante (análogo a la media μ en la distribución normal), k_h es el parámetro de concentración (análogo al recíproco de la dispersión σ^2), estos dos últimos para cada $h = 1, 2, \dots, H$ componente del FVFM y w_h es el parámetro de mezcla o peso de las funciones de von Mises (*mixture parameter*).

Además, el parámetro de mezcla del FVMF está sujeto a la siguiente restricción:

$$0 \leq w_h \leq 1 \text{ y } \sum_{h=1}^H w_h = 1 \text{ para } (h = 1, 2, \dots, H) \quad (2.20)$$

Para estimar los parámetros del FVMF, se sugiere utilizar el método *expectation maximization*, debido a que los métodos regulares son incapaces de manejar la complejidad del modelo, consideraciones que se mencionan en el trabajo de Banerjee et al [14].

Por último, los resultados de este trabajo muestran que FVMF provee un razonable ajuste a diferentes conjunto de datos, obteniendo un modelo que explica más del 90 % de la variación de los datos, en este caso, obtenidos de estaciones ubicadas en la península de Malasia. En la figura 2.4 se aprecia el ajuste del modelo a los datos, tanto la comparación con el histograma, como en su versión circular.

En el trabajo de Heckenbergerova et al. [5], se utiliza una estrategia diferente a la anteriormente mencionada, basados en *Particle Swarm Optimization*. Proponen una forma distinta para encontrar un modelo de ajuste, utilizando la distribución estadística *finite mixture of circular normal von Mises* (MvM), o simplemente *mixture of von Mises distribution*, similar a la mencionada previamente.

En este caso, se define la *simple von Mises distribution* (SvM) como:

$$f(\theta; \mu, k) = \frac{1}{2\pi I_0(k)} e^{k \cos(\theta - \mu)} \quad (2.21)$$

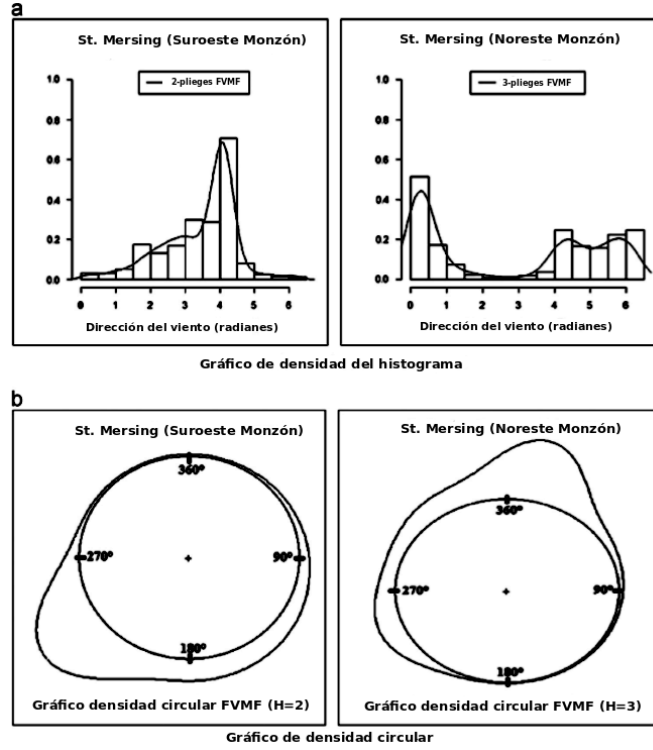


Figura 2.4: Modelo de ajuste FVMV para suroeste y noreste en la estación Mersing
Creado por [13].

Donde $k \geq 0$, $0 \leq \mu \leq 2\pi$, $0 \leq \theta \leq 2\pi$ y $I_0(k)$ representa la versión modificada de la función de Bessel de primera clase y orden cero:

$$I_0(k) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} e^{k \cos(\theta)} d\theta = \sum_{k=0}^{\infty} \frac{1}{(k!)^2} \left(\frac{k}{2}\right)^{2k} \quad (2.22)$$

Para $k = 0$, la distribución SvM se vuelve uniforme alrededor de un círculo con todas las direcciones equi-probables. Cuando una colección de datos tiene más de una dirección predominante, es necesario utilizar una mezcla (*mixture*) de distribuciones. Así, la función de densidad de probabilidad *finite mixture of simple von Mises* (MvM-pdf) queda como:

$$\phi(\theta; v) = \sum_{j=1}^k w_j \cdot f_j(\theta; \mu_j, k_j) \quad (2.23)$$

Donde k es el número de funciones de la mezcla, j es el índice de una particular SvM-pdf con parámetros μ_j y k_j , θ es una variable angular ($0 \leq \theta \leq 2\pi$), y v es un vector parámetro

de la forma:

$$v = (\mu, k, w) = (\mu_1, \dots, \mu_k, k_1, \dots, k_k, w_1, \dots, w_k) \quad (2.24)$$

Para lograr el objetivo, se obtiene en primer lugar una aproximación numérica de los parámetros del MvM a partir de los datos recolectados de la dirección del viento, estrategia nombrada como estimación analítica en el trabajo de Heckenbergerova et al. [5]. Luego optimiza estos parámetros mediante el uso de la meta-heurística *Particle Swarm Optimization*, con parámetros fijos, en donde la solución está representada por una codificación del vector \vec{v} mencionado anteriormente (Sección 2.24).

Como test estadístico, es utilizado el *Pearson's chi-squared goodness-off-fit*. Los resultados muestran la mejora que se logra a la estimación inicial, comparando estos resultados con otra propuesta similar la cual se expone en un trabajo previo de los mismos autores [15], en donde se utiliza la misma estrategia pero reemplazando el PSO con algoritmos genéticos. Sin embargo, estos resultados no consiguen pasar el test estadístico impuesto por ellos mismos, por lo que existe trabajo futuro a realizar para mejorar la propuesta y lograr la precisión deseada.

Los resultados obtenidos para los datos recolectados en el aeropuerto de St John localizado en Newfoundland, Canadá, son apreciables en la figura 2.5.

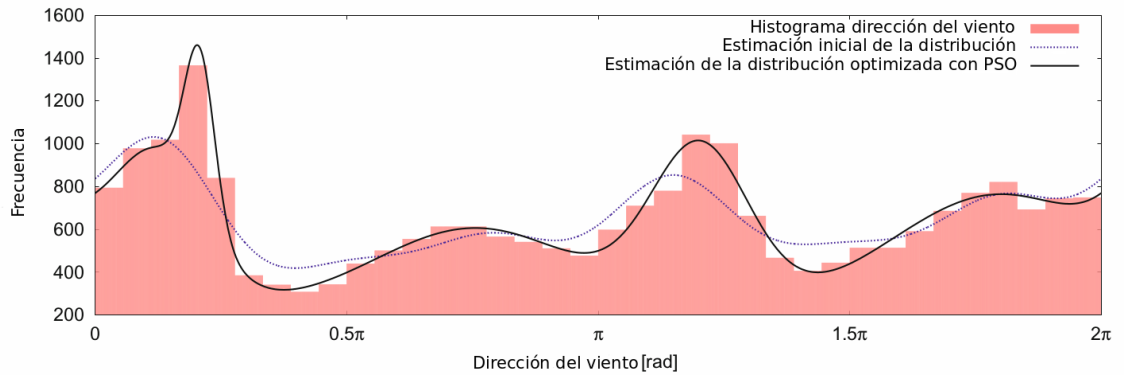


Figura 2.5: Ajuste dirección del viento, aeropuerto St. John.
Creado por Heckenbergerova et al [5].

Otro trabajo interesante de revisar es el caso de estudio realizado por Carta et al. [16], en donde se utiliza la *finite mixture of von Mises distribution* para representar la velocidad direccional de viento. El ajuste es realizado mediante el método de mínimos cuadrados, y resuelto con el algoritmo Levenberg-Marquardt [17] el cual requiere de un punto de

partida o solución inicial. Para ello, utilizan el mismo método que se propone en el trabajo de Heckenbergerova et al. [15] el cual se utiliza en este trabajo y que es explicado en la siguiente sección 4.1. La calidad de la solución es medida a través del test R^2 , y dentro de las conclusiones que se obtienen, es que más de 6 mezclas (*mixtures*) de von Mises no incrementan considerablemente el valor de R^2 .

El método es aplicado a datos recolectados desde dos estaciones del clima en el archipiélago de las islas Canarias. Una vez más, se confirma que la *finite mixture of von Mises distribution* es aplicable a distintas regiones con una o más direcciones predominantes.

2.4. Resumen del capítulo

Para modelar la velocidad y la dirección del viento, son comúnmente utilizadas las distribuciones de Weibull y von Mises respectivamente. Diversas técnicas han sido propuestas para ajustar los parámetros de las distribuciones de densidad de probabilidad, con el objetivo de representar lo mejor posible el conjunto de datos del viento. Normalmente se utilizan métodos numéricos, sin embargo, otras técnicas han sido exploradas con el fin de mejorar tiempos de ejecución y calidad de solución. Entre ellas, la meta-heurística *Particle Swarm Optimization* ha demostrado conseguir buenos resultados, en diversos conjuntos de datos provenientes de distintas zonas del mundo.

El modelo de la dirección el viento es un tema menos tratado que el de la velocidad. Por lo mismo, el ajuste realizado mediante el PSO aún presenta ciertas deficiencias, como las expuestas en las conclusiones de Heckenbergerova et al. [5], las cuales revelan que sus resultados no obtienen los valores esperados para el test propuesto.

Capítulo 3

PSO para la velocidad del viento en Valparaíso

En este capítulo se presenta la implementación de la técnica *Particle Swarm Optimization* basada principalmente en el trabajo de Carneiro et al. [4]. En principio se definen y explican los principales conceptos para entender la solución propuesta, para luego mostrar los resultados obtenidos al buscar los parámetros de ajuste de la distribución de Weibull.

3.1. Modelo Matemático

Como se mencionó anteriormente, para encontrar los parámetros de la distribución de Weibull que se ajusten a los datos de prueba se utilizará la meta-heurística *Particle Swarm Optimization*. La función de distribución de Weibull está definida en la ecuación 2.5. La función objetivo se describe con la fórmula 2.18 y es aquella con la que se busca minimizar el error cuadrático entre la frecuencia real de los datos y la estimada por la distribución de Weibull. Los parámetros a encontrar k y c deben ser ≥ 0 . Para encontrar estos parámetros se construyó el PSO que se detalla a continuación.

3.2. Estructura del PSO

En esta sección se detallará cada una de los componentes del algoritmo, destacando algunas consideraciones importantes para el correcto desempeño de este método.

3.2.1. Representación

Cada vector posición de las partículas del enjambre representa una solución candidata la cual varía dentro de cierto espacio de búsqueda definido por los límites de las componentes. Así, para el caso de los parámetros de la distribución de Weibull, la posición de las partículas está representada por los parámetros k y c quedando de la forma:

$$x = (k, c) \quad (3.1)$$

Para ambos parámetros, se establecen los límites entre $0 \leq (k, c) \leq 20$, criterio que se basa en el trabajo de Carneiro et al. [4]. De esta forma, las partículas se moverán dentro de ese rango, manteniéndose en los lugares que minimizan la función objetivo, la cual representa el error de la predicción de Weibull versus los datos reales.

Así, para cada partícula se define una estructura que posee las siguientes propiedades:

1. Posición: vector de largo dos de números flotantes, los cuales representan la ubicación de la partícula dentro del espacio de búsqueda y sus componentes a los parámetros k y c .
2. Velocidad: vector de números flotantes que representan el cambio de valor de cada componente de la posición de la partícula en determinada iteración. Se actualiza en base a la ecuación de velocidad del PSO:

$$v_{i,j}^{k+1} = wv_{i,j}^k + c_1r_1(xbest_{i,j}^k - x_{i,j}^k) + c_2r_2(xgbest_j^k - x_{i,j}^k) \quad (3.2)$$

En donde $v_{i,j}^{k+1}$ es la velocidad de la partícula en el instante $k + 1$, w , c_1 y c_2 son los parámetros de inercia, cognitivo y social respectivamente, r_1 y r_2 son números aleatorios entre $[0, 1]$, $xbest_{i,j}^k$ y $xgbest_j^k$ son la mejor solución de la partícula y la mejor solución del vecindario respectivamente en el instante k , $x_{i,j}^k$ representa la posición de la partícula, o la solución, en el instante k .

3. Mejor resultado personal: vector flotante que guarda la mejor posición conseguida por la partícula durante las iteraciones transcurridas.

Mientras que el enjambre, siendo esencialmente una estructura que posee referencia a todas las partículas, queda representado de la siguiente forma:

1. Partículas: Arreglo de referencias a las estructuras de partículas creadas.
2. Mejor posición global: De todos los mejores resultados de cada partícula, se almacena la mejor posición de todas. La que persiste al final del ciclo de iteraciones, es la solución final.

3.2.2. Consideración de los parámetros

Los parámetros w de inercia, c_1 cognitivo y c_2 social del PSO determinan con que peso inciden en el camino que siguen las partículas del enjambre, la velocidad anterior, la solución particular y la solución global respectivamente.

La inercia w regula qué magnitud de la velocidad anterior se mantiene en la actual, así la partícula se comporta similar a un cuerpo desacelerando. En Kaveh [6] se explica que en términos extremos, si la velocidad previa se elimina (peso nulo), las partículas no pueden salir del círculo local relativo a las soluciones iniciales, lo que sería equivalente a un procedimiento de búsqueda local. Por el contrario, si se le da mucho peso, las partículas tendería a huir de las buenas posiciones conocidas.

El parámetro cognitivo c_1 controla la influencia de la mejor solución encontrada por la partícula misma en la velocidad actual. Si se le da un peso mayor a este valor, la partícula tenderá a explorar las zonas locales a la posición desde donde partió.

El parámetro social c_2 controla la influencia de la mejor solución global encontrada por el enjambre en la velocidad actual. Si se le da un peso mayor a este valor, la partícula tenderá a explotar las zonas locales a la posición de las solución global.

La convergencia prematura es un comportamiento presente en el *Particle Swarm Optimization*, como se explica nuevamente en Kaveh [6]. Esto se debe a que normalmente, las partículas se concentran en una “nube” local a la mejor solución global una vez que se ha avanzado ciertas iteraciones, lo que por una lado hace que el algoritmo sea más rápido que otras técnicas afines, como los algoritmos evolutivos, pero por otro impide que, a partir de cierto punto, se siga mejorando la solución. Por tanto, hay un límite óptimo para las iteraciones, ya que en cierto momento las partículas se quedarán atrapadas en un óptimo local.

A modo de evitar la convergencia prematura, varias técnicas han sido propuestas, las cuales se resumen en el trabajo de Kaveh [6]. Para esta memoria, se utilizará la recomendación de

Chang [3] para la variación de parámetros del enjambre:

$$w(j) = (1 - \frac{j}{iter_{max}})^\alpha (w_{max} - w_{min}) + w_{min} \quad (3.3)$$

$$c_1(j) = (1 - \frac{j}{iter_{max}})^\beta (c_{1max} - c_{1min}) + c_{1min} \quad (3.4)$$

$$c_2(j) = (1 - \frac{j}{iter_{max}})^\gamma (c_{2min} - c_{2max}) + c_{2max} \quad (3.5)$$

Donde $w(j)$, $c_1(j)$, $c_2(j)$, son los parámetros de inercia, cognitivo y social en el instante j . Los valores para los rangos de los parámetros son $w_{max} = 0,9$ y $w_{min} = 0,4$, c_{1max} , c_{2max} y c_{1min} , c_{2min} , 2.5 y 0 respectivamente. Los parámetros α , β , γ son definidos como 0.5, 1.5 y 1.0 respectivamente. La variable $iter_{max}$ es el máximo número de iteraciones. Esta modificación provocará que las partículas “frenen” su convergencia al óptimo local, dado que en cada iteración se da menos peso a la solución global y mayor peso al óptimo de cada partícula y a la velocidad anterior.

3.2.3. Descripción del algoritmo

La lógica del algoritmo PSO (ver Algoritmo 1) se basa en mover las partículas dentro del rango definido para los componentes de la solución hasta que todas las partículas se concentren en alguna zona que represente una buena solución al problema, no necesariamente el óptimo. Lo importante en cada iteración es actualizar o mover el enjambre, revisar y guardar las mejores soluciones y actualizar los parámetros de inercia, cognitivo y social que definen las velocidades.

3.3. Resultados

En esta sección se presentan los resultados obtenidos en los experimentos realizados para el ajuste de la distribución de Weibull a diferentes conjuntos de datos del viento en Valparaíso.

3.3.1. Experimentos

Los experimentos fueron probados hasta en un máximo de 1000 iteraciones y 50 partículas, (A excepción del experimento donde se consideraron todos los promedios

Algorithm 1 PSO para el ajuste de los parámetros de la distribución de Weibull

Require: Datos de frecuencias de velocidades del viento

Ensure: Valores para los parámetros k y c

enjambre = inicializar($w, c1, c2$)

for $i = 1$ to $Iter_{max}$ **do**

for Each partícula en enjambre **do**

 actualizarVelocidadPartícula(partícula)

 actualizarPosiciónPartícula(partícula)

 revisarLímitesPosición(partícula)

 guardarMejorResultadoPartícula(partícula)

end for

 guardarMejorResultadoGlobal(enjambre)

 actualizarParámetros(enjambre)

end for

retornarMejorResultadoGlobal(enjambre)

diarios, 2013, 2014 y 2015, en el cual, se utilizaron 200 partículas). Los parámetros de w , $c1$ y $c2$ fueron definidos tal y como explica en el modelo matemático, en la sección 3.1.

Los experimentos fueron realizados con datos del viento obtenidos por la Armada de Chile para la región de Valparaíso en los años 2013, 2014 y 2015. Estos fueron tratados mediante *scripts* desarrollados en *Python* para obtener las frecuencias de las distintas velocidades del viento registradas a lo largo del año. Los datos se organizaban de la siguiente forma: Por cada año, se tiene una tabla en un archivo excel de cada mes, en donde se registra por cada fila los resultados de la medición de cada día. Las mediciones son registradas en un intervalo de tres horas, es decir, se tienen registros diarios para las 3:00, 6:00, 9:00, 12:00, 15:00, 18:00, 21:00 y 00:00 horas. Un ejemplo es la tabla mostrada en la Figura 3.1.

El ajuste de la distribución de Weibull a estos datos de velocidad del viento se hizo considerando las siguientes configuraciones para el cálculo del histograma de frecuencias:

1. **Todos los años.** Se considera el promedio diario de velocidad del viento como dato unitario para el cálculo de las frecuencias, considerando todos los días en el intervalo de Enero del 2013 hasta Diciembre del 2015.
2. **Anual.** Se considera el promedio diario de velocidad del viento como dato unitario para el cálculo de las frecuencias en un lapso anual (2013, 2014 y 2015).
3. **Por temporada.** Se considera el promedio diario de velocidad del viento como dato unitario para el cálculo de las frecuencias en un lapso de tres meses (Enero - Marzo ;

PROMEDIOS MENSUALES	12	05
MÁXIMA INTENSIDAD MENSUAL	22	

Obtenido desde el Instituto Meteorológico de la Armada de Chile.

Abril - Junio; Julio - Septiembre; Octubre - Diciembre).

4. **Datos brutos.** Se considera cada medición realizada (8 por día) como dato unitario, en un lapso de un año.

Una vez obtenido los datos de frecuencias, se procede a aplicar el algoritmo *Particle Swarm Optimization* donde se obtienen los parámetros de ajuste k y c . De esta manera, se evalúa la calidad del modelo generado (distribución de densidad de probabilidad de Weibull), para las distintas configuraciones mediante gráficos y los siguientes test estadísticos (utilizados también en el trabajo de Carneiro et al. [4]):

- ### 1. Root Mean Square Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (X_i - Y_i)^2}{N}} \quad (3.6)$$

2. Correlation

$$r = \frac{\sum_{i=1}^N (X_i - X_{med}) \cdot (Y_i - Y_{med})}{\sqrt{\sum_{i=1}^N (X_i - X_{med})^2} \cdot \sqrt{\sum_{i=1}^N (Y_i - Y_{med})^2}} \quad (3.7)$$

3. Relative Bias

$$RB = \frac{X_{med} - Y_{med}}{Y_{med}} \quad (3.8)$$

Donde N es el número de datos, Y_i la frecuencia de dichos datos, X_i la frecuencia entregada por la distribución de Weibull, X_{med} la media de X_i e Y_{med} la media de Y_i .

Las pruebas fueron realizadas en un computador con sistema operativo Ubuntu 16.04 64-bit, 3.8 GB de memoria y procesador doble núcleo Intel Pentium 2.60 GHz.

3.3.2. Análisis de los resultados

Visualización de los datos

Los gráficos 3.2(a), 3.2(b) y 3.2(c) muestran la distribución de datos de velocidad del viento en Valparaíso a lo largo de los meses del año y las horas del día, lo cual permite visualizar la naturaleza de la intensidad del viento de forma cualitativa. Por ejemplo, se logra apreciar que las máximas velocidades son obtenidas en los meses finales de primavera y comienzos de verano.

Las superficies que se exponen no representan la distribución continua de los datos. Estos fueron graficados de forma discreta y posteriormente la herramienta utilizada para generar la superficie unió los puntos mediante rectas.

Experimento 1, datos anuales y promedios diarios

Las figuras 3.3(a), 3.3(b) y 3.3(c) muestran el ajuste de la distribución de Weibull a los histogramas de datos del viento (promedios diarios), con los parámetros k y c que se muestran en las primeras tres filas de la tabla 3.1 determinados por el PSO. El ajuste tiene buena forma, lo cual es corroborado por los datos estadísticos obtenidos con los test

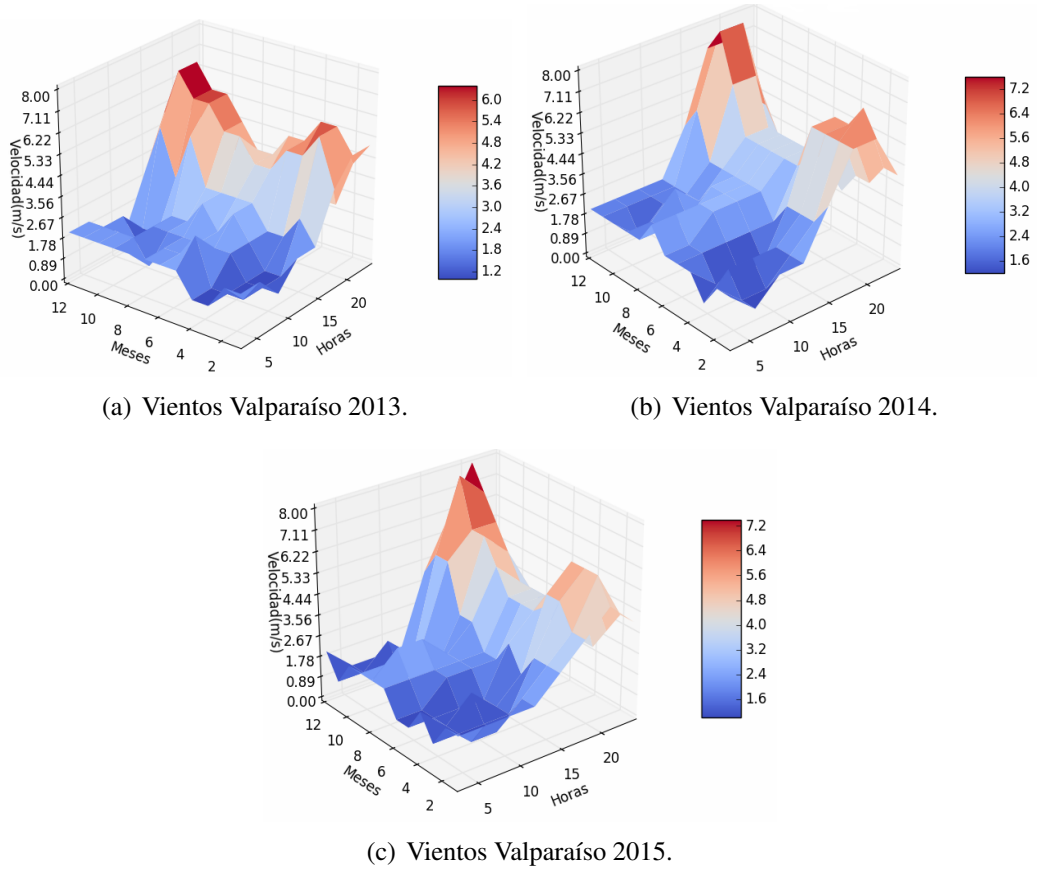
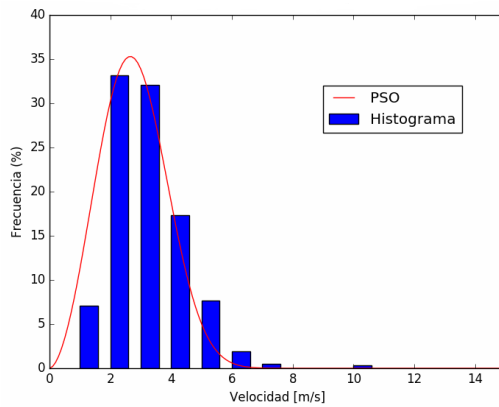


Figura 3.2: Superficie de datos viento de Valparaíso.
Fuente: Elaboración Propia.

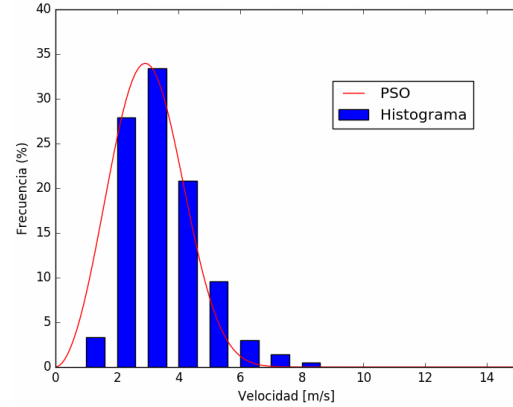
previamente mencionados (RMSE, r , RB), expuestos en la tabla 3.1 en las filas 1, 2 y 3. Si se compara con la precisión conseguida en el trabajo de Carneiro et al. [4], se aprecia que el ajuste conseguido es levemente más impreciso, sobre todo en lo relativo al test *relative bias* (RB) el cual es una medida de distancia de la frecuencia estimada con la de los datos. Esto podría deberse a la naturaleza de los datos trabajados.

Experimento 2, datos de tres años y promedios diarios

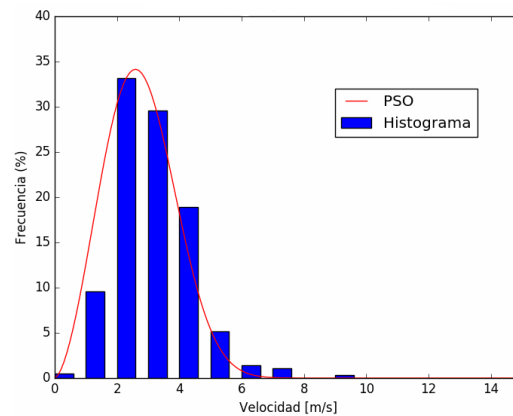
En este experimento se realizó el ajuste considerando los promedios diarios y un intervalo de tres años consecutivos. El gráfico 3.4(b), muestra el resultado del ajuste con PSO y la configuración estándar de los demás experimentos, es decir, 100 iteraciones y 50 partículas. En este gráfico se aprecia que el ajuste no es bueno, a pesar de las cifras en la tabla 3.1, fila 4: PSO (50p), dado que oscila bastante alrededor de las barras del histograma,



(a) PSO Valparaíso 2013.



(b) PSO Valparaíso 2014.



(c) PSO Valparaíso 2015.

Figura 3.3: Ajuste con PSO a datos del viento de Valparaíso.
Fuente: Elaboración Propia.

por lo que se repite el experimento aumentando el número de partículas a 200 obteniendo el gráfico 3.4(a), con el cual se obtiene un ajuste más adecuado, además de mejorar los resultados de los test estadísticos (tabla 3.1, fila 5: PSO (200p)).

Experimento 3, ajuste a datos anuales con resultados del experimento 2

Los gráficos 3.5(a), 3.5(b) y 3.7 son ajustes de Weibull con los parámetros obtenidos en el experimento anterior. Es decir, la idea es evaluar el modelo general de los tres años versus el histograma de datos de cada año en particular. El ajuste desde los resultados estadísticos (tabla 3.1), es levemente menos preciso que el modelo ajustado a cada año en particular, pero sigue siendo aceptable como posible opción a considerar.

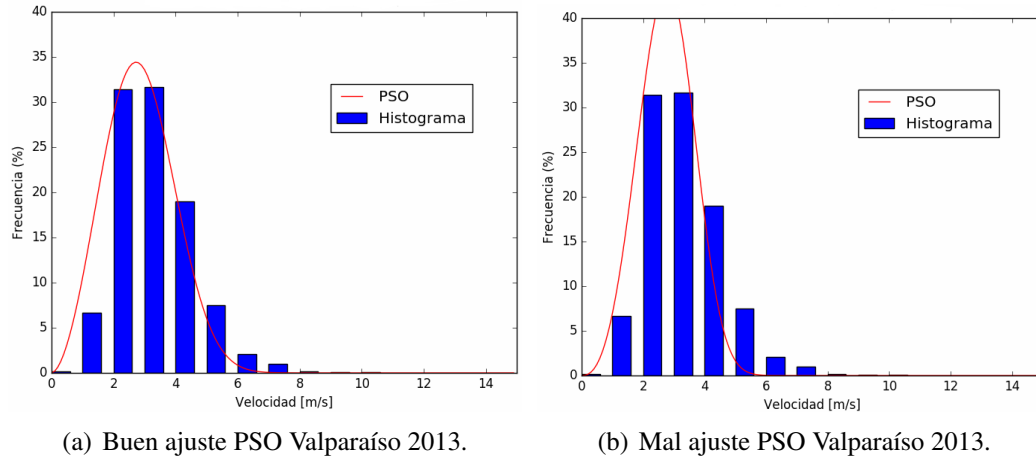


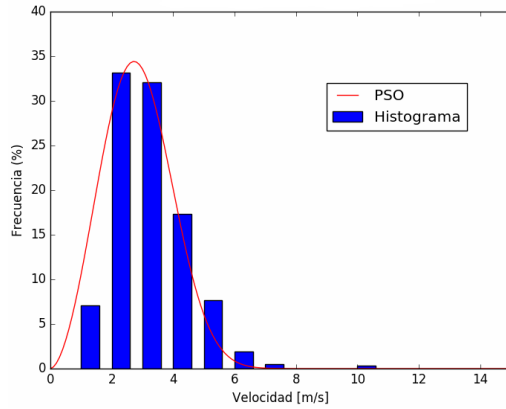
Figura 3.4: Ajuste con PSO a datos Valparaíso 2015, 2014 y 2013, baja y buena calidad.
Fuente: Elaboración Propia.

Experimento 4, ajuste a datos de tres meses y promedios diarios

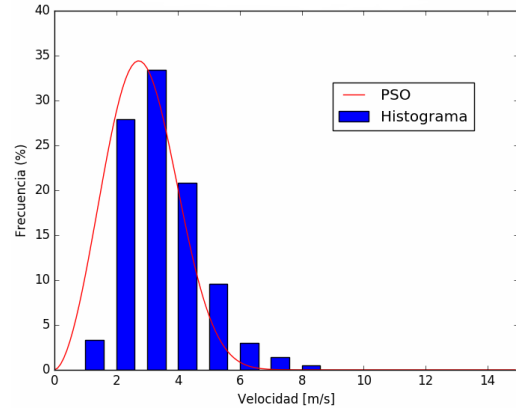
Es posible que se requiera un análisis más acotado, por ello los gráficos 3.6(a), 3.6(b), 3.6(c), 3.6(d), muestran un ajuste considerando un lapso de 3 meses para el año 2015, con el que se demuestra que es posible definir cualquier intervalo (manteniendo como unidad de dato el promedio diario de velocidad del viento) y obtener un ajuste adecuado de los datos mediante la distribución de Weibull.

Experimento 4, ajuste a datos año 2015 y datos brutos

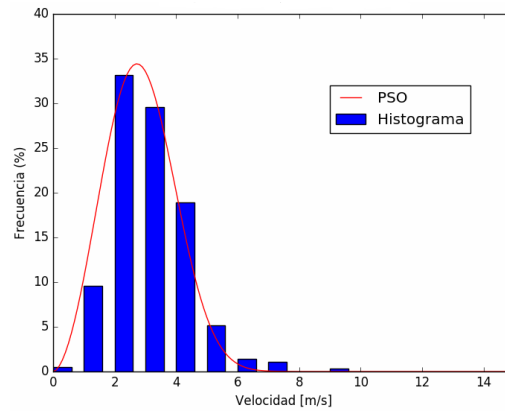
La razón de por qué se utiliza el promedio diario de los datos del viento para ajustar Weibull y no las mediciones puras (las mediciones tomadas cada 3 horas diariamente) es expuesta en el gráfico 3.7. La distribución de Weibull no se ajusta a una distribución de datos con más de un máximo, por lo que de requerirse un modelo para este caso se debe buscar otra distribución o modificar la distribución de Weibull. Es por ello que normalmente se utiliza un promedio de los datos, como se realiza en el trabajo de Farade [9].



(a) Velocidad viento Valparaíso 2013.



(b) Velocidad viento Valparaíso 2014.



(c) Velocidad viento Valparaíso 2015.

Figura 3.5: Ajuste con PSO a registros del viento en Valparaíso (Con todos los datos).

Fuente: Elaboración Propia.

Resumen de los experimentos

En esta sección se exponen los resultados obtenidos de los experimentos realizados. En este caso la aplicación de la propuesta realizada en Carneiro et al. [4], obtuvo los resultados esperados en cuanto a la calidad del ajuste de la distribución de Weibull a los datos del viento de Valparaíso.

La tabla se organiza como se explica a continuación:

1. **Método:** Método utilizado para el ajuste de los parámetros de Weibull.
2. **Periodo:** Periodo de tiempo considerado para los datos en el experimento.
3. **k:** Parámetro k de la distribución de Weibull.

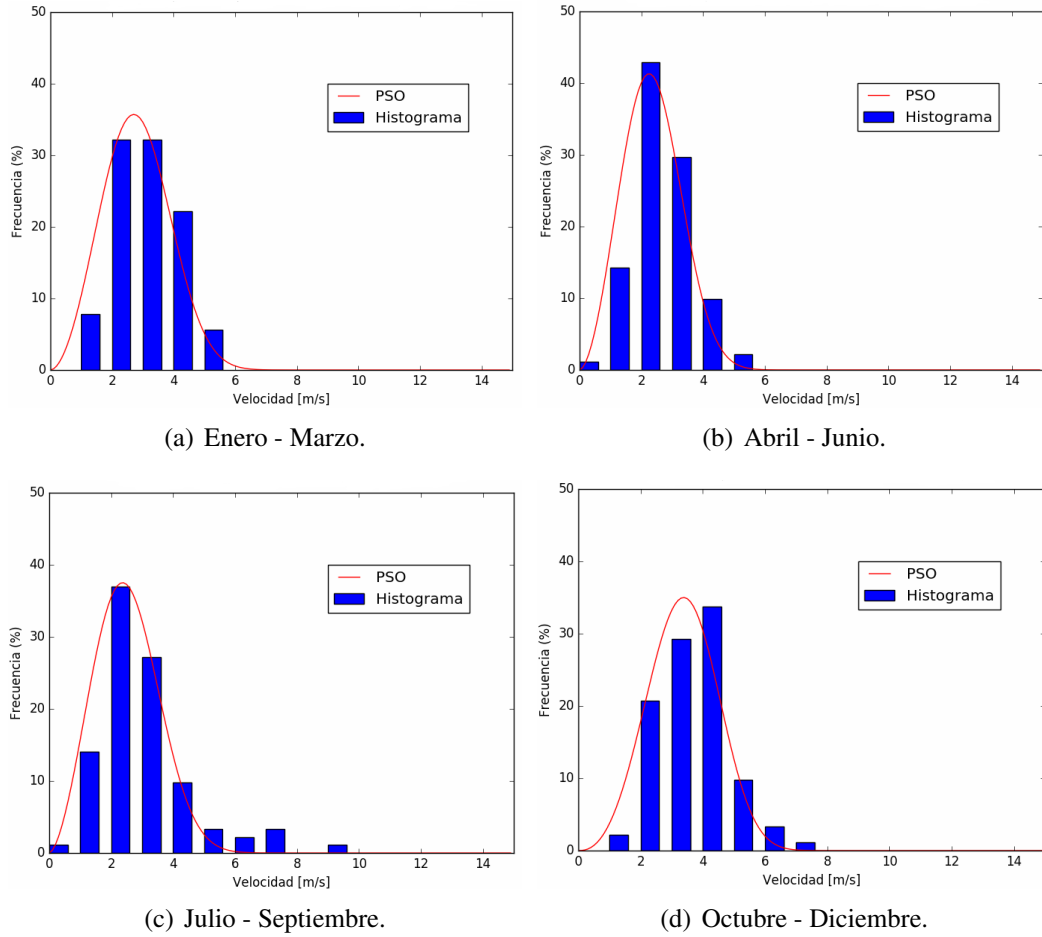


Figura 3.6: Ajuste con PSO a datos Valparaíso 2015, por rango de meses.

Fuente: Elaboración Propia.

4. **c**: Parámetro c de la distribución de Weibull.
5. **RMSE**: Test estadístico conocido como *root mean square error*.
6. **r**: Test estadístico conocido como *correlation*.
7. **RB**: Test estadístico conocido como *relative bias*.
8. **Tiempo**: Tiempo total de ejecución del algoritmo.

Los tiempos omitidos en la Tabla 3.1, en las filas 6, 7 y 8 indican que no se calcularon los parámetros de la función de densidad, sino que se utilizaron los obtenidos previamente, en el experimento de la fila 5. El número de iteraciones del PSO no se considera debido a que

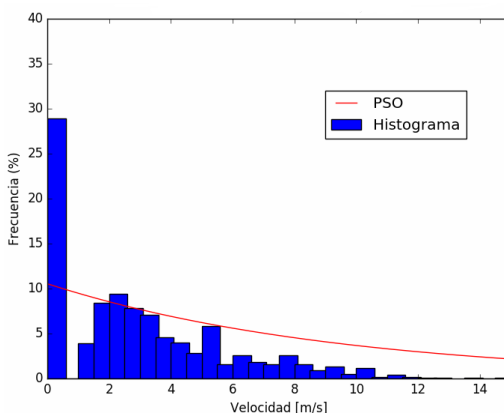


Figura 3.7: Ajuste con PSO a datos (cifras puras) Valparaíso 2015, 2014, 2013

Fuente: Elaboración Propia.

sólo se itera en la fase de desarrollo del algoritmo. Cuando este está completo, basta con una ejecución para tener el resultado definitivo del experimento.

Tabla 3.1: Tabla de pruebas, PSO velocidad del viento

#	Método	Periodo	k	c	RMSE	r	RB	Tiempo
1	PSO	2013	2.78	3.12	$2.26 \cdot 10^{-2}$	0.984	$1.98 \cdot 10^{-3}$	1.74s
2	PSO	2014	2.91	3.37	$2.32 \cdot 10^{-2}$	0.982	$7.54 \cdot 10^{-4}$	1.63s
3	PSO	2015	2.65	3.10	$1.64 \cdot 10^{-2}$	0.992	$3.02 \cdot 10^{-3}$	1.59s
4	PSO (50p)	2015-14-13	3.47	3.07	$3.61 \cdot 10^{-2}$	0.975	$4.11 \cdot 10^{-3}$	2.00s
5	PSO (200p)	2015-14-13	2.78	3.20	$1.61 \cdot 10^{-2}$	0.994	$1.91 \cdot 10^{-3}$	7.78s
6	PSO (200p y todos los datos)	2013	2.78	3.20	$2.41 \cdot 10^{-2}$	0.981	$1.92 \cdot 10^{-3}$	-
7	PSO (200p y todos los datos)	2014	2.78	3.20	$3.01 \cdot 10^{-2}$	0.970	$8.89 \cdot 10^{-6}$	-
8	PSO (200p y todos los datos)	2015	2.78	3.20	$2.02 \cdot 10^{-2}$	0.986	$1.92 \cdot 10^{-3}$	-
9	PSO	Ene-Mar	2.85	3.15	$2.31 \cdot 10^{-2}$	0.982	$6.41 \cdot 10^{-3}$	1.29s
10	PSO	Abr-Jun	2.76	2.65	$2.04 \cdot 10^{-2}$	0.993	$3.03 \cdot 10^{-3}$	1.34s
11	PSO	Jul-Sep	2.66	2.83	$2.51 \cdot 10^{-2}$	0.985	$4.43 \cdot 10^{-3}$	1.19s
12	PSO	Oct-Dic	3.40	3.75	$2.60 \cdot 10^{-2}$	0.978	$7.16 \cdot 10^{-4}$	1.30s
13	PSO (datos brutos)	2015	1.00	9.49	$4.51 \cdot 10^{-2}$	0.751	$6.7 \cdot 10^{-1}$	1.65s

3.4. Conclusiones del capítulo

En este capítulo se implementó la propuesta realizada en Carneiro et al. [4] para ajustar los parámetros de la distribución de Weibull a los datos de velocidad del viento en Valparaíso. Los resultados confirman que el método propuesto permite obtener soluciones de buena calidad, bajo el punto de vista de los test estadísticos aplicados (RMSE, r y RB), y los gráficos generados. Sin embargo, se debe tener en consideración que la distribución de Weibull representa el promedio diario de velocidades medidas del viento, detalle que no siempre se menciona en los trabajos que utilizan esta distribución, pero que se explican en algunos como en Fadare [9]. Este modelo es útil para aplicaciones como, la estimación del potencial eléctrico en determinada región.

Si se desea otra representación para las velocidades del viento, que permita representar velocidades predominantes, o una descripción más directa de los datos, como los representados en la Figura 3.7, se debe buscar otra distribución o una versión modificada de Weibull.

Capítulo 4

PSO para la dirección del viento en Valparaíso

4.1. Modelo Matemático

Como se comenta anteriormente, la distribución de densidad de probabilidad que se utilizará para describir la distribución de datos de dirección del viento es la *finite mixtures of von mises distribution* descrita en 2.23, la cual consiste básicamente en una combinación lineal de la *simple von Mises distribution* descrita en 2.21.

De forma preliminar, los datos se ordenan en un histograma de densidad con el cual se obtiene un esqueleto de la distribución de densidad de probabilidad. Posteriormente se requieren encontrar los parámetros de ajuste μ_j , k_j y w_j para cada j -ésima *simple von Mises distribution*. La forma en que se realiza esto último en este trabajo está basado en el documento de Carta et al. [16] y se describe a continuación.

Para la construcción del histograma se divide el rango de datos que va de 0 a 2π en T clases con frecuencia O_i la cual representa la suma de las observaciones en el rango de la clase T . Posteriormente se definen k sectores del mismo largo desde las T clases, relacionados al número de direcciones de viento predominantes (o con mayor frecuencia). Esto define el número de funciones de von Mises a utilizar. La estimación de k se realiza mediante la observación del histograma de los datos, observando la cantidad de direcciones con altas frecuencias (también puede ser considerado un parámetro de ajuste, que requiera de pruebas empíricas para encontrar el valor adecuado). Además, se sigue la observación empírica en el trabajo de Carta et al. [16], en donde se concluye que valores superiores a 6 *mixtures* no

mejoran considerablemente la calidad del ajuste.

Para la aproximación inicial de los parámetros de la *mixture of von mises distribution* se utiliza una estimación numérica, utilizada por Heckenbergerova et al. [5] [15] y Carta et al. [16], basada en los datos recolectados acerca de la dirección del viento.

Sea $j \in \{1 \dots k\}$ el subíndice del sector representado por la j -ésima función de von Mises. La dirección del viento predominante μ_j se estima de la siguiente forma:

$$\mu_j = \begin{cases} \arctan(\frac{s_j}{c_j}) & s_j \geq 0, c_j > 0 \\ \frac{\pi}{2} & s_j > 0, c_j = 0 \\ \pi + \arctan(\frac{s_j}{c_j}) & c_j < 0 \\ \pi & s_j > 0, c_j = -1 \\ 2\pi + \arctan(\frac{s_j}{c_j}) & s_j < 0, c_j > 0 \\ 3\frac{\pi}{2} & s_j < 0, c_j = 0 \end{cases} \quad (4.1)$$

En donde s_j y c_j representan el seno y coseno promedio del sector j .

Tradicionalmente, se estima el parámetro de concentración k_j con la ecuación:

$$\frac{I_1(k_j)}{I_0(k_j)} = \sqrt{s_j^2 + c_j^2} \quad (4.2)$$

Donde $I_1(k_j)$ es la función modificada de Bessel de primera clase y orden 1. Como se explica en Banerjee et al. [14], debido a la falta de una solución analítica a la ecuación 4.2, no es posible estimar directamente los valores de k . Se podrían utilizar métodos para ecuaciones no lineales, pero para datos de altas dimensiones, problemas de desbordamiento (*overflow*) o inestabilidad numérica se vuelven concurrentes. Por tanto, se utiliza la propuesta realizada en el trabajo de Heckenbergerova et al. [5] con lo cual el parámetro k_j puede ser aproximado por:

$$|k_j| = \{23,29041409 - 16,8617370 \sqrt[4]{s_j^2 + c_j^2}\} \quad (4.3)$$

También existe otra forma similar a esta aproximación, utilizada por Heckenbergerova et al. en [15] y por Carta et al. [16], la cual consiste en la siguiente fórmula:

$$k_j = \{23,29041409 - 16,8617370\sqrt{s_j^2 + c_j^2} - 17,4749884\exp(-(s_j^2 + c_j^2))\}^{-1} \quad (4.4)$$

Los pesos iniciales w_j son aproximados como:

$$w_j = \frac{\sum_{i=J_l}^{J_u} O_i}{\sum_{i=1}^T O_i} \quad (4.5)$$

Donde J_l y J_u son los índices de los bordes del sector j .

La función objetivo para el PSO es el test estadístico χ^2 descrito en Heckenbergerova et al. [5] como sigue a continuación:

$$\chi^2 = \sum_{i=1}^T \frac{(O_i - np_i)^2}{np_i} \quad (4.6)$$

Donde T es el número de clases de frecuencia definido para construir el histograma, n es la suma de las frecuencias observadas O_i y p_i es la probabilidad teórica de cada clase de frecuencia predicha por el modelo ajustado.

Para el cálculo del p_i se utiliza:

$$p_i = \int_{l_i}^{u_i} f(x)dx \quad (4.7)$$

Donde u_i y l_i son los bordes de la i -ésima clase de frecuencia.

La forma de la solución a encontrar es descrita en 2.24. Esta es restringida por la condición para los pesos de la *mixture von Mises distribution*, la cual obliga a que se deba cumplir que la suma de los pesos sea igual a 1, como se describe en la sección 2.20.

4.2. Estructura del PSO

Es esta sección se detallará cada una de los componentes del algoritmo para el ajuste de los parámetros de la distribución de von Mises para la dirección del viento. A diferencia del algoritmo anterior, en este se incluye una estimación inicial de la solución y una propuesta al manejo de parámetros del PSO.

4.2.1. Representación

La representación del PSO es similar al utilizado para el ajuste de la distribución de datos de velocidad del viento. Las partículas y el enjambre están representados por las ecuaciones descritas en 3.2.1 y 3.2.1 respectivamente.

La solución para el PSO que mejora la estimación inicial de los parámetros para la *mixture von Mises distribution* está representado por un vector v en el cual se encuentran los valores para todos los parámetros de cada *simple von Mises distribution*. Estos valores están codificados para que el algoritmo se mueva en el rango desde 0 a 1.

El vector solución tiene la forma:

$$v = (\overbrace{v_1, \dots, v_k}^{\mu}, \overbrace{v_{k+1}, \dots, v_{2k}}^k, \overbrace{v_{2k+1}, \dots, v_n}^w). \quad (4.8)$$

El parámetro μ está representado en el rango $i \in \{1, \dots, k\}$ y para ser decodificado debe ser escalado por 2π .

El parámetro k está representado en el rango $i \in \{k + 1, \dots, 2k\}$ y para ser decodificado debe ser escalado por $[0, 700]$.

El parámetro w_j está representado en el rango $i \in \{2k + 1, \dots, n\}$ cuyos valores van en el rango $[0, 1]$.

4.2.2. Consideración de los parámetros

En el trabajo de Heckenbergerova et al. [5] los parámetros del PSO son fijados de la siguiente forma.

1. Para el factor de inercia w el valor es 0.89.
2. Para el factor cognitivo c_1 el valor es 0.5.
3. Para el factor social c_2 el valor es 0.7.

Sin embargo, como se concluye en su trabajo y de acuerdo a los resultados de la tabla 4.1, estos valores no generan resultados satisfactorios, es decir, las soluciones encontradas bajo esos parámetros, al ser evaluadas en la función objetivo, no satisfacen los requerimientos para el test realizado por ellos mismos. Por esto, en esta memoria se propone una estrategia distinta para el control de los parámetros del PSO.

Anteriormente se menciona la estrategia sugerida en el trabajo de Chang [3] para la variación de parámetros del PSO, descrita en 3.3, con el fin de evitar una convergencia prematura en el ajuste de la distribución de Weibull. Esta estrategia podría ser considerada para intentar mejorar los resultados conseguidos por Heckenbergerova et al. [5], sin embargo, la sugerencia presenta inconvenientes por que depende del número de iteraciones máximo que se defina para el algoritmo.

Al realizarse diversas pruebas con cantidad de iteraciones distintas, se puede evidenciar que el hecho de que el número de iteraciones máximo sean un valor que afecta la variación de parámetros, tiene como consecuencia que el algoritmo tome diferentes caminos hacia la solución final, por ende, un aumento en el número de iteraciones no constituiría necesariamente una mejora en calidad de la solución. De cierta forma, el número de iteraciones máximo se transforma en un factor pseudo-aleatorio para el PSO.

Por lo tanto, debido a que los parámetros fijos no consiguen buenos resultados y la variación de parámetros propuesta por Chang [3] presenta inconvenientes estructurales para el PSO, se propone el siguiente método para el ajuste de parámetros.

Sean $w(j)$, $c_1(j)$ y $c_2(j)$ los parámetros de inercia, cognitivo y social del PSO en el instante j respectivamente:

$$w(j) = w(j) + (w_{max} - w(j)) * F \quad (4.9)$$

$$c_1(j) = c_1(j) + (c_{1max} - c_1(j)) * F \quad (4.10)$$

$$c_2(j) = c_2(j) + (c_{2min} - c_2(j)) * F \quad (4.11)$$

Donde F es un factor de avance dentro de un rango definido para los parámetros el cual está en función de una valor de la función objetivo esperado y el valor actual de la solución. Es decir:

$$F = \frac{\text{Valor esperado en la FO}}{\text{Valor actual en la FO}} \quad (4.12)$$

Si el factor F sobrepasa el valor de 1 (caso en el que la solución actual es mejor que la esperada considerando la minimización de la función objetivo), se deja de cambiar los

parámetros. Esto se resume como:

$$F = \begin{cases} \frac{P}{p_j} & P \leq p_j \\ 0 & P > p_j \end{cases} \quad (4.13)$$

Donde P es un valor esperado para la función objetivo (definido como parámetro) y p_j es el valor en la función objetivo de la mejor solución hasta el momento. La actualización de parámetros se realiza solamente cuando se mejora la solución global actual. La variación tiene la misma lógica que la propuesta de Chang [3], es decir, abolir el aporte de la solución global hacia el final de las iteraciones.

En síntesis, el cambio de los parámetros se produce en proporción a la distancia del valor actual con respecto a un valor final y a razón de un valor esperado para la función objetivo sobre el valor actual de la mejor solución.

Con esto se logra mejorar el rendimiento del algoritmo alcanzando mejores soluciones en tiempos menores de ejecución como se puede observar en la sección de resultados 4.3.

4.2.3. Descripción del algoritmo

La estructura del algoritmo definida para el PSO consiste en dos fases. La primera, una aproximación basada en la estimación numérica de los parámetros requeridos para la *mixture of von Mises distribution* a través de operaciones simples con los datos recolectados, y la segunda, una mejora de la solución inicial obtenida en la fase anterior mediante el uso de la meta-heurística *Particle Swarm Optimization*.

El método para la aproximación inicial de la solución se describe en el Algoritmo 2. En donde la estimación de los parámetros se realiza como se describe en la ecuación 4.1 para los μ_j , la ecuación 4.2 para los k_j y la ecuación 4.5 para los pesos w_j . Una vez obtenida la aproximación inicial se procede a mejorar esta mediante el uso del PSO descrito en el Algoritmo 3. Para la inicialización de las partículas, se realizaron pequeñas perturbaciones a la solución inicial tal y como se sugiere en Heckenbergerova et al. [5]. Esto evita que la solución escape a zonas que tengan un buen resultado en la función objetivo, pero cuya forma escape a la del histograma. Debido a que la función objetivo definida en 4.6 mide las diferencias de frecuencias entre los datos reales y los teóricos, es decir, las áreas de las barras del histograma de densidad versus el área bajo la curva de la distribución de probabilidad en algún intervalo, más de una forma de la curva podría parecer una buena

Algorithm 2 Aproximación inicial de los parámetros de la *mixture von Mises distribution*

Require: Datos de frecuencias de la dirección del viento.

Require: K , Cantidad de *simple von Mises distribution*.

Require: T , clases de frecuencias.

Require: D , Total de datos.

Ensure: Valores para los parámetros μ_j , k_j y w_j , para cada $j \in \{1, \dots, k\}$.

```
sol = inicializarVectorSolución(3*K)
for  $j = 0$  to  $K$  do
  datos $j$  = datosEnRango( $j * D / K$ )
  s $j$  = obtenerSenoPromedio(datos $j$ )
  c $j$  = obtenerCosenoPromedio(datos $j$ )
  u $j$  = obtenerDirecciónPredominante(s $j$ , c $j$ )
  k $j$  = obtenerConcentración(s $j$ , c $j$ )
  w $j$  = obtenerPeso( $j * (T / K)$ ,  $(j + 1) * (T / K)$ )
  addToSolution(sol, u $j$ , k $j$ , w $j$ )
end for
retornarSoluciónInicial(sol).
```

Algorithm 3 PSO para la mejora de la aproximación de los parámetros de la *mixture von Mises distribution*

Require: Datos de la dirección del viento.

Require: Solución inicial para el ajuste de la *mixture von Mises distribution*.

Ensure: Solución inicial mejorada.

```
enjambre = inicializar(w, c1, c2)
mejorSol
for  $i = 1$  to  $Iter_{max}$  do
  for Each partículas en enjambre do
    actualizarVelocidadPartícula(partícula)
    actualizarPosiciónPartícula(partícula)
    revisarLímitesPosición(partícula)
    guardarMejorResultadoPartícula(partícula)
  end for
  mejorSol = guardarMejorResultadoGlobal(enjambre)
  if Se mejora solución global then
    actualizarParámetros(enjambre, mejorSol)
  end if
end for
retornarMejorResultadoGlobal(enjambre).
```

solución como se vé en el gráfico de mal ajuste 3.4(b). Por ello, la idea es mantener la forma inicial encontrada, mejorándola sin deformarla. Así, las perturbaciones iniciales a los valores de las posiciones de las partículas eran del orden de 10^{-3} .

Los parámetros del PSO varían de acuerdo a lo definido anteriormente en las ecuaciones

4.9, 4.10 y 4.11.

La forma en que se cuidaron las condiciones de borde consistieron en limitar el avance de las partículas a los bordes 0 y 1 manteniéndolos en dichos valores si es que se excedían a ellos.

Para cuidar la restricción de pesos se normalizaran los valores determinados en cada iteración, es decir, se suman todos los valores w_j y se ponderan dichos valores por el recíproco de la suma obtenida.

Debido a que la función objetivo implica determinar la frecuencia teórica, es necesario determinar la probabilidad de cierto rango de direcciones mediante el cálculo del área bajo la curva de la distribución de densidad de probabilidad, para luego ser multiplicada por la suma del total de datos y así obtener el valor requerido. Por ende, para el cálculo de la integral se utilizaron sumas de Riemann con una partición conveniente al desempeño del algoritmo y la precisión requerida.

Finalmente, la solución obtenida es decodificada tal y como se explica en la sección 4.2.1.

4.3. Resultados

En esta sección se presentan los experimentos realizados y los resultados obtenidos al realizar el ajuste de los parámetros de la *mixture of von Mises distribution* a los datos de dirección del viento mediante el uso del PSO.

4.3.1. Experimentos

Similar a los descrito en la sección 3.3, los datos de dirección del viento son tratados para rescatar las mediciones pertinentes al trabajo aquí expuesto. Estos se encuentran inicialmente en un formato como el que se puede apreciar en la Figura 3.1.

Nuevamente, las pruebas fueron realizadas en un computador con sistema operativo Ubuntu 16.04 64-bit, 3.8 GB de memoria y procesador doble núcleo Intel Pentium 2.60 GHz.

Para evaluar la calidad de la solución, se utilizó el test *Chi square goodness fit* [18], con lo cual se evalúa que tan bien representa el modelo propuesto a los datos medidos. Para ello, la hipótesis nula H_0 afirma que los datos de dirección del viento se distribuyen según la función de densidad *mixture of von Mises distribution* y la hipótesis alternativa H_1 niega

dicha afirmación. Se rechaza H_0 si el valor de la función objetivo del PSO para la solución final encontrada excede el valor crítico de χ^2 para un nivel de significancia de $\alpha = 0,05$ y 13 grados de libertad, es decir **22.362**, valor que puede encontrarse en la tabla de la distribución χ^2 [19]. Los grados de libertad son definidos a partir de la cantidad de clases de frecuencia definidas para el estudio, en este caso, se dividió el rango de valores de $[0, 2\pi]$ en 14 tramos iguales, por lo que quedan $(n - 1)$ grados de libertad, 13 en este caso.

Los experimentos consistieron en el ajuste de varios subconjuntos de datos provenientes de las mediciones obtenidas para la dirección del viento en los años 2013, 2014 y 2015. Así, se prueba la utilidad de la propuesta realizada independiente del rango de tiempo a modelar. Es importante considerar que no se incluyen los días en los que no hubo viento por la evidente imposibilidad de registrar la dirección. Los subconjuntos definidos fueron los siguientes:

1. **Anual:** Se consideran los datos de todo el año elegido.
2. **Meses acumulados:** Se considera una agrupación mensual pero reuniendo los datos de tres años consecutivos (2013, 2014, 2015). Es decir, para el mes de Enero, se ajusta el modelo a los datos de Enero-2013, Enero-2014 y Enero-2015 en conjunto.
3. **Meses:** Se escogen algunos meses para ser comparados consigo mismos durante los tres años escogidos. Por ejemplo, Enero-2013, Enero-2014 y Enero-2015 por separado.

Para el funcionamiento del PSO, se estableció un límite de 50500 iteraciones, se utilizaron 100 partículas, y se usó como criterio de parada si es que el valor en la función objetivo de la mejor solución encontrada al momento era menor a 22.362 (criterio basado en la estrategia de cumplir el test *Chi square goodness fit*). El mismo valor se estableció como valor esperado para la función objetivo al utilizarse la variación de parámetros definida en las ecuaciones 4.9, 4.10 y 4.11. La cantidad de *mixture of simple von Mises distribution* utilizadas fue de 7, siguiendo la recomendación de Carta et al. [16] y algunas pruebas iniciales.

Por último, se comparan los resultados obtenidos por el PSO propuesto en el trabajo de Heckenbergerova et al. [5] con la propuesta realizada en esta memoria para la variación de parámetros del PSO definidas en las ecuaciones 4.9, 4.10 y 4.11.

4.3.2. Análisis de los resultados

Experimento 1, pruebas iniciales

En la Figura 4.1(a) se puede observar un ajuste bastante distorsionado respecto al histograma de datos. Esto se debe a que la función objetivo del PSO evalúa la diferencia entre las frecuencias experimentales y las obtenidas teóricamente, por lo tanto, diferentes curvas pueden tener igual magnitud del área bajo la curva y por ende, la misma probabilidad con la que se obtiene la frecuencia teórica.

Por esto, es importante que el PSO busque mejorar la solución inicial en una vecindad cercana a esta, de manera de obtener una evolución como la que se aprecia en la Figura 4.1(b). Para lograr esto, el algoritmo debe inicializar las partículas en la posición de la solución inicial encontrada más una pequeña perturbación.

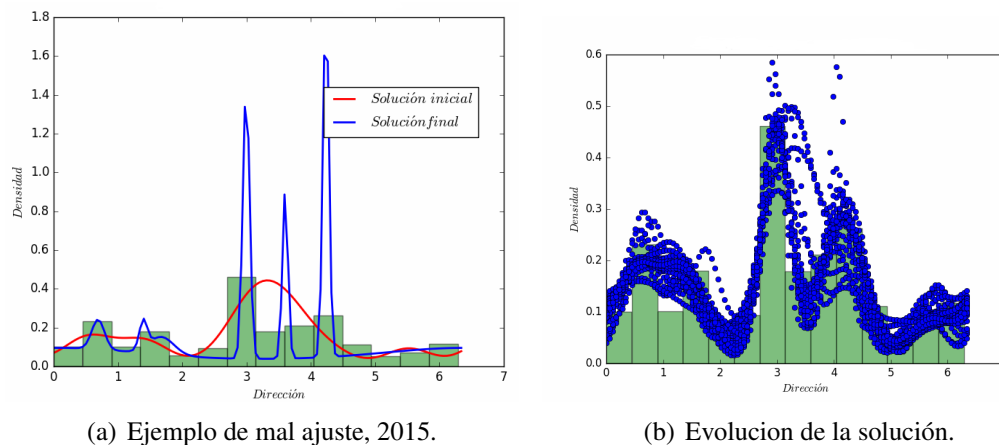


Figura 4.1: Pruebas iniciales.
Fuente: Elaboración propia.

Experimento 2, Ajustes anuales

En este experimento se ajustan los vientos anualmente, en los años de los cuales se obtuvieron datos para este estudio, resultado la Figura 4.2(c) para el 2013, la Figura 4.2(b) para el 2014 y la Figura 4.2(a) para el 2015. Tal como se realiza en otros estudios [5] [13], el ajuste de los datos de dirección del viento en un formato de largo plazo, permite observar el comportamiento global de los vientos pudiendo evaluar la norma o generalidad en los datos registrados. En las Figuras referenciadas previamente se observa la evolución de la solución desde la aproximación inicial hasta la mejorada por el PSO.

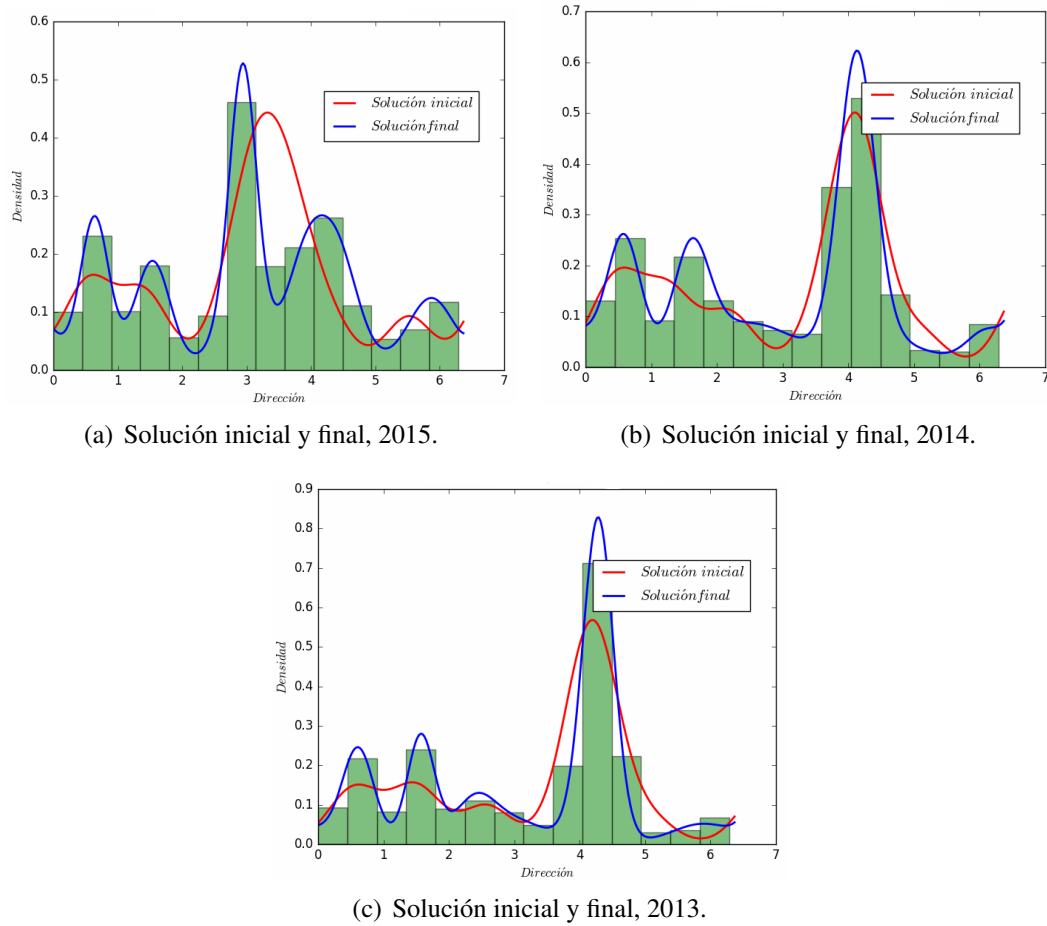
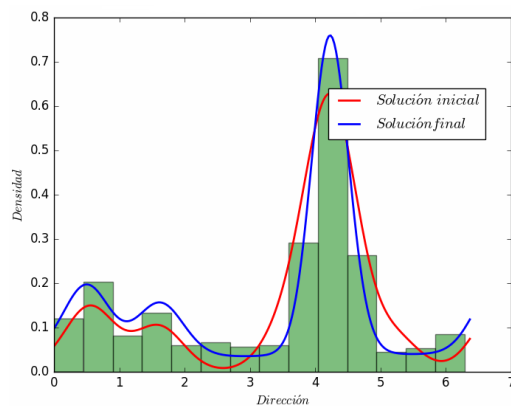


Figura 4.2: Graficos de ajustes anuales.
Fuente: Elaboración propia.

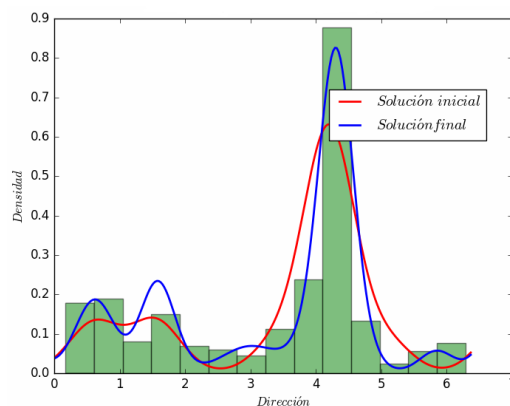
Experimento 3, Ajustes de meses acumulados

Similar al ejercicio anterior, los gráficos agrupados en 4.3 y 4.4, permiten visualizar el ajuste de los datos del viento por meses. En este caso se agruparon los datos de los tres años escogidos por cada mes, es decir, la Figura 4.3(a), por ejemplo, contiene los datos del mes de enero de los años 2013, 2014 y 2015.

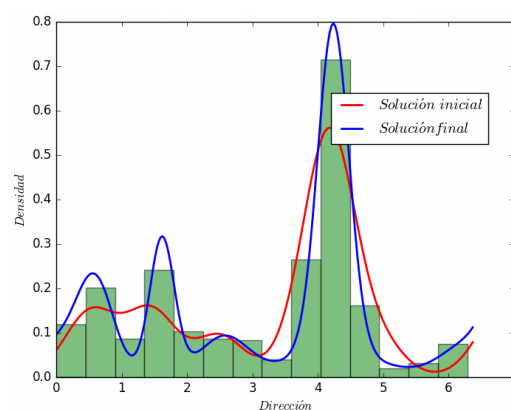
Se puede observar que algunos gráficos muestran un mejor ajuste que otros, a pesar del correcto valor obtenido en la función objetivo como se expresa posteriormente en la tabla 4.1. Esto se debe a que probablemente la solución se escapa más de los deseado de la solución inicial.



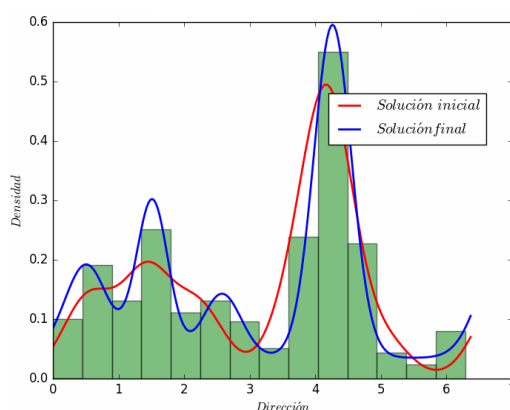
(a) Solución inicial y final, Enero.



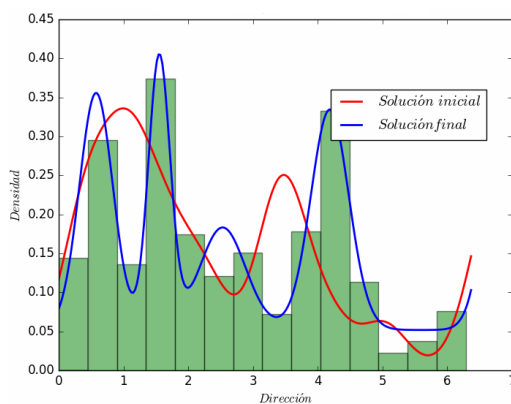
(b) Solución inicial y final, Febrero.



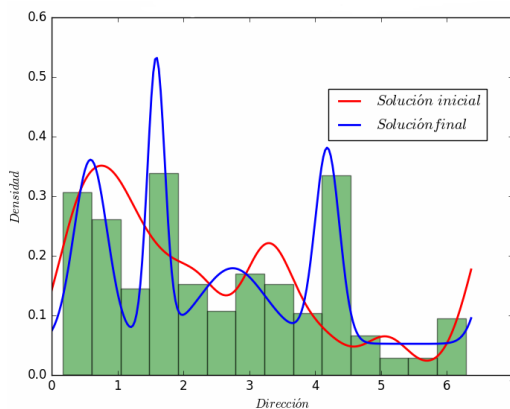
(c) Solución inicial y final, Marzo.



(d) Solución inicial y final, Abril.



(e) Solución inicial y final, Mayo.

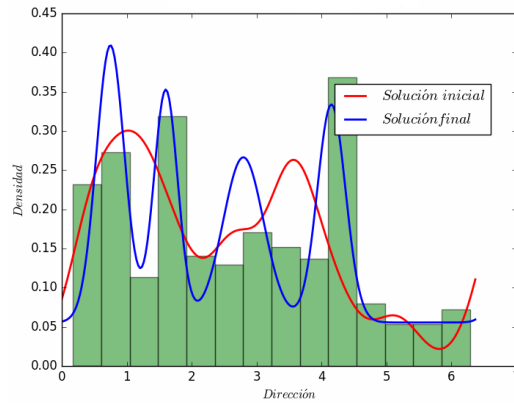


(f) Solución inicial y final, Junio.

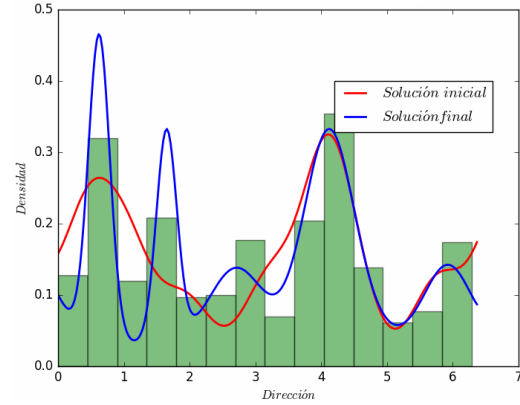
Figura 4.3: Gráficos de ajuste de MVM por meses.
Fuente: Elaboración propia.

Experimento 4, Ajuste por meses, visualización en coordenadas polares

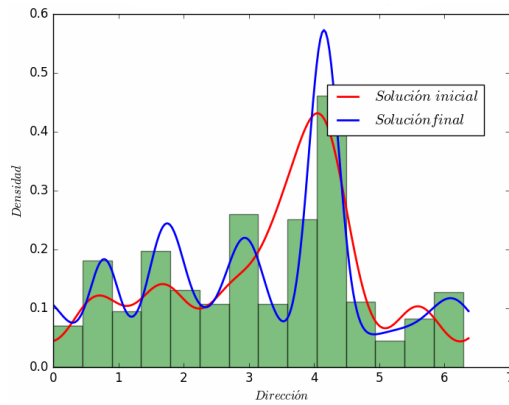
En las Figuras agrupadas 4.5 y 4.6 se aprecia una visualización más interesante. Allí se puede ver de forma más intuitiva las direcciones dominantes del viento, relativas al sistema de referencia utilizado en meteorología conocido comúnmente como la rosa de los



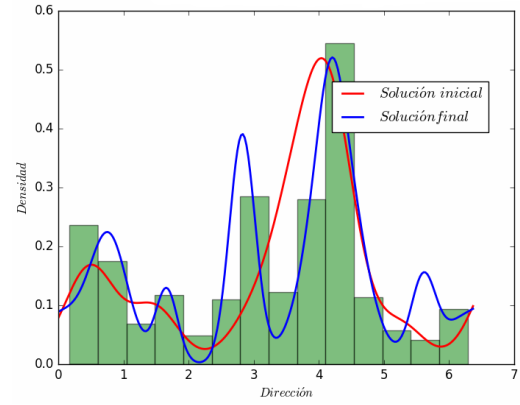
(a) Solución inicial y final, Julio.



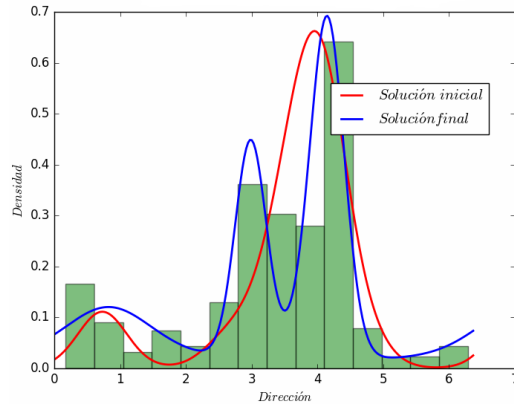
(b) Solución inicial y final, Agosto.



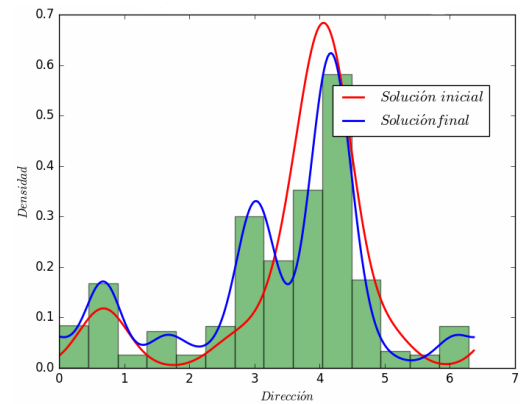
(c) Solución inicial y final, Septiembre.



(d) Solución inicial y final, Octubre.



(e) Solución inicial y final, Noviembre.



(f) Solución inicial y final, Diciembre.

Figura 4.4: Gráficos de ajuste de MVM por meses.

Fuente: Elaboración propia.

vientos [20]. En el mes de Enero, Figuras 4.5(a) 4.5(b) 4.5(c), los vientos parecen provenir principalmente desde el suroeste, mientras que en Mayo, Figuras 4.5(d) 4.5(e) 4.5(f), el origen del viento es más variable, proviniendo desde el noreste y el sur principalmente.

En Septiembre, Figuras 4.6(a) 4.6(b) 4.6(c), vuelve a haber una fuente de los vientos predominante el suroeste.

Resumen de los experimentos

Tabla 4.1: Tabla de pruebas, PSO dirección del viento

Periodo de tiempo	PSI	PSFV	PSFF	Tiempo V	Tiempo F	Iteraciones V	Iteraciones F	Cantidad datos
2015	461.83	-	251.37	-	38m46.1s	-	50500	2076
2015	461.83	21.58	234.97	8m5.7s	47m47.1s	9005	50500	2076
2014	301.12	19.86	56.57	20m37.8s	46m23.2s	29494	50500	2499
2013	424.24	20.76	69.53	11m41.1s	48m23.7s	16778	50500	2346
Enero	144.42	20.89	22.18	0m3.3s	25m37.2s	86	27276	633
Febrero	130.39	19.45	22.27	0m12.2s	0m07.1s	288	134	567
Marzo	104.54	18.61	22.31	0m3.1s	0m13.3s	49	228	564
Abril	74.97	21.13	22.25	0m12.5s	17m58.4s	192	18544	559
Mayo	243.65	21.65	81.75	1m6.9s	42m59.6s	1237	50000	589
Junio	286.30	21.01	188.91	1m42.9s	40m52.5s	1989	50000	554
Julio	177.33	22.06	31.39	0m17.8s	45m48.3s	315	50000	604
Agosto	79.57	21.94	22.17	0m12.9s	40m26.3s	202	41204	578
Septiembre	92.77	21.31	22.31	0m1.7s	0m1.7s	27	166	541
Octubre	188.14	21.81	37.50	0m45.3s	42m56.4s	825	50000	564
Noviembre	363.90	26.25	88.16	42m21.4s	42m53.6s	50500	50000	582
Diciembre	335.28	22.30	37.81	9m56.1s	43m52.9s	11640	50000	586
Enero 2015	42.94	16.34	18.77	0m4.1s	0m0.1s	61	10	220
Enero 2014	68.10	20.68	21.61	0m1.8s	0m0.7s	29	126	206
Enero 2013	194.98	19.31	24.86	0m0.4s	46m29.2s	6	50000	207
Mayo 2015	60.84	17.99	26.69	0m2.3s	41m53.8s	39	50000	173
Mayo 2014	142.85	20.94	20.48	0m4.8s	0m0.3s	82	50	213
Mayo 2013	135.68	20.88	22.34	0m5.3s	0m1.1s	97	2089	203
Septiembre 2015	77.41	21.29	25.54	0m8.4s	41m48.7s	137	50000	140
Septiembre 2014	19.56	18.24	15.86	0m0.1s	0m0.1s	1	1	200
Septiembre 2013	68.90	16.61	21.89	0m0.3s	0m0.9s	5	150	201

La tabla 4.1 muestra un resumen de los resultados obtenidos al ajustar la distribución de probabilidad *von Mises distribution* en diferentes subconjuntos de tiempo de los datos colectados en los años 2013, 2014 y 2015. La tabla está organizada como sigue a continuación. La estrategia **Variable** hace referencia a la propuesta realizada en esta memoria para la variación de parámetros del PSO explicada en la sección 4.2.2, mientras que la estrategia **Fija** es la seguida en el trabajo de Heckenbergerova et al. [5], la cual consta de la definición de parámetros fijos para el PSO.

1. *Periodo de tiempo*: Rango de tiempo donde se consideraron los datos para el ajuste.
2. *PSI*: Puntaje solución inicial, es decir, el valor obtenido al evaluar la solución inicial en la función objetivo.
3. *PSFV*: Puntaje solución final *variable*, se refiere a la solución obtenida por el PSO al utilizar la estrategia **Variable**.

4. *PSFF*: Puntaje solución final *fija*, se refiere a la solución obtenida por el PSO al utilizar la estrategia **Fija**.
5. *Tiempo V*: Tiempo empleado por el PSO con la estrategia **Variable**.
6. *Tiempo F*: Tiempo empleado por el PSO con la estrategia **Fija**.
7. *Iteraciones V*: Iteraciones empleadas hasta llegar a la mejor solución por el PSO con la estrategia **Variable**.
8. *Iteraciones F*: Iteraciones empleadas hasta llegar a la mejor solución por el PSO con la estrategia **Fija**.
9. *Cantidad datos*: Cantidad de datos utilizados en el ajuste.

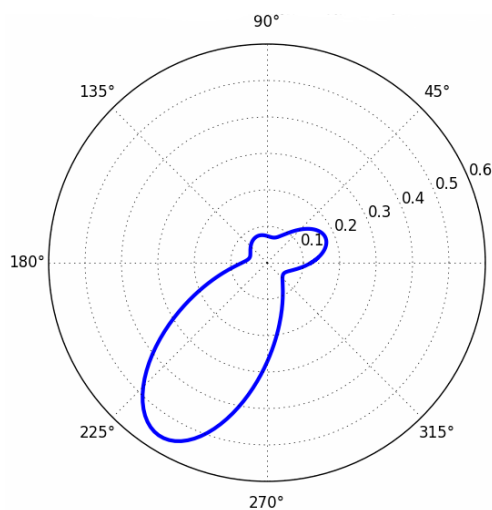
Los gráficos en 4.7 y 4.8 muestran la comparación de los rendimientos de los algoritmos con parámetros fijos y variables para el PSO según la propuesta explicada en la sección 4.2.2. Es claro que en la mayoría de los experimentos, la propuesta realizada tiene mejor rendimiento y en el peor de los casos se comporta igual o levemente peor que la forma tradicional de parámetros fijos utilizada en el trabajo de Heckenbergerova et al. [5].

4.4. Conclusiones del capítulo

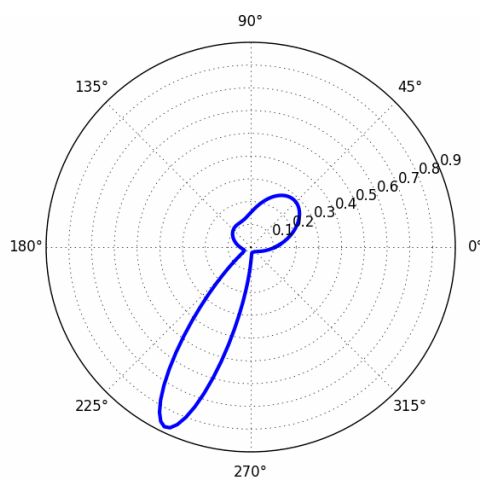
En este capítulo se ajustaron los parámetros de la función *mixture of von Mises distribution* a datos de dirección del viento en Valparaíso. Basado en el trabajo de Heckenbergerova et al. [5], se implementó el *Particle Swarm Optimization* con una modificación propuesta en esta memoria para mejorar los resultados del trabajo citado.

Como se puede ver en la tabla 4.1, la modificación propuesta y explicada en las ecuaciones 4.9, 4.10 y 4.11 para el control de los parámetros del PSO, permite obtener mejores tiempos de ejecución y soluciones de mejor calidad que los valores utilizados en el trabajo de Heckenbergerova et al. [5]. Esto se debe al problema de convergencia prematura del PSO, el cual no fue considerado en dicho trabajo.

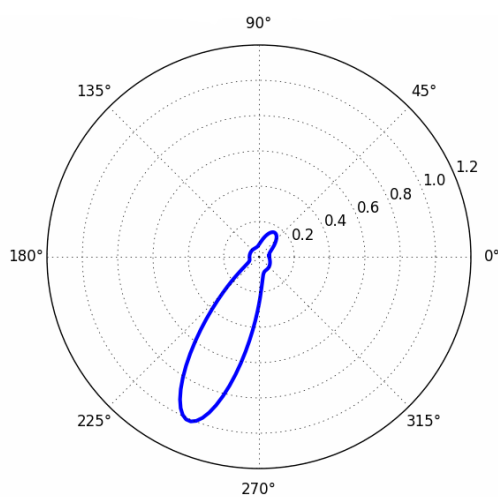
De las Figuras expuestas en 4.3 y 4.4 se puede observar que a pesar del buen puntaje conseguido por la solución, el gráfico difiere del histograma en varias pruebas más de lo deseado. Esto se debe a que la función objetivo definida en el PSO procura que las áreas bajo la curva de la función de densidad de probabilidad coincidan con las del histograma, sin embargo, más de una forma de la función de densidad puede cumplir dicho objetivo.



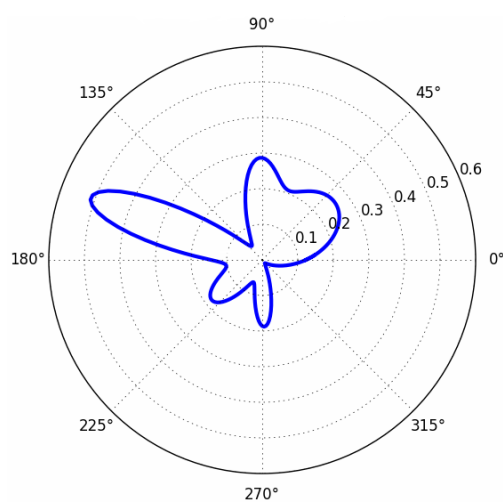
(a) Dirección enero 2015.



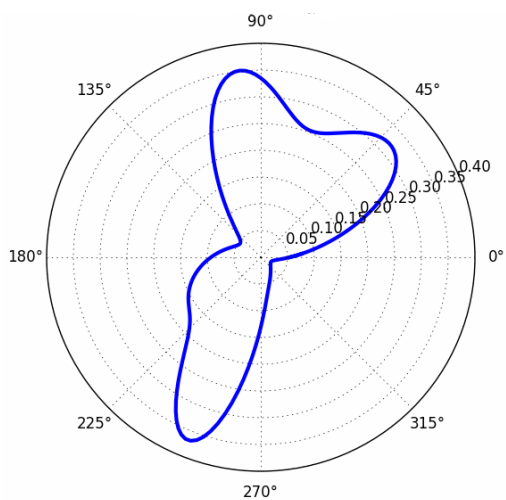
(b) Dirección enero 2014.



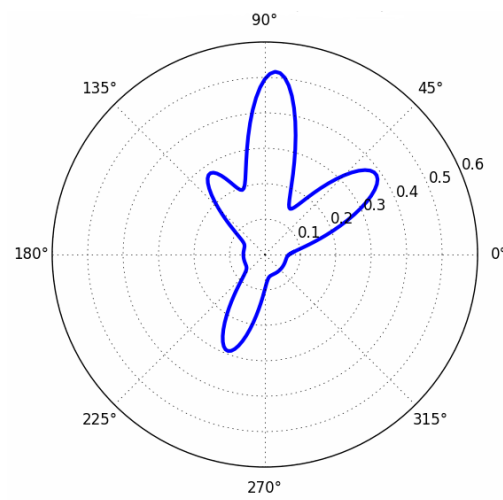
(c) Dirección enero 2013.



(d) Dirección mayo 2015.



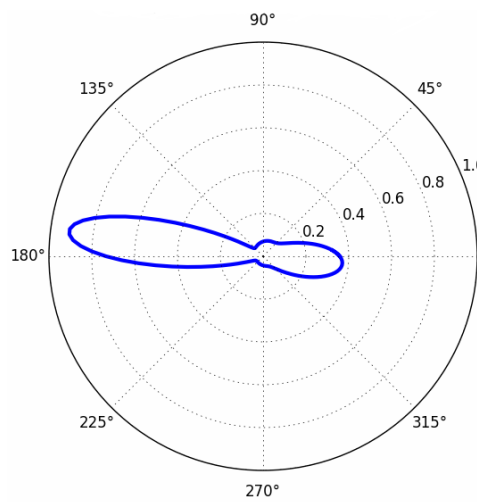
(e) Dirección mayo 2014.



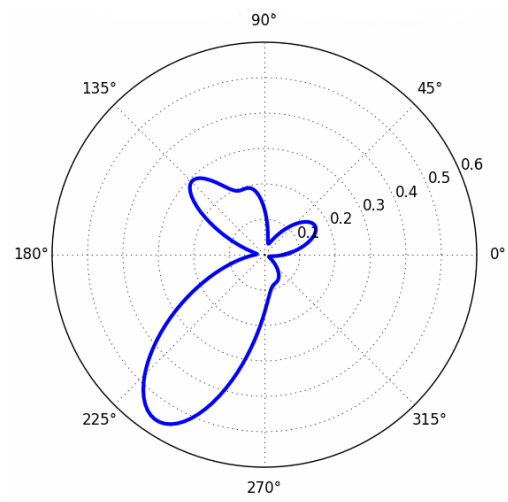
(f) Dirección mayo 2013.

Figura 4.5: Gráficos de ajuste por meses en coordenadas polares.

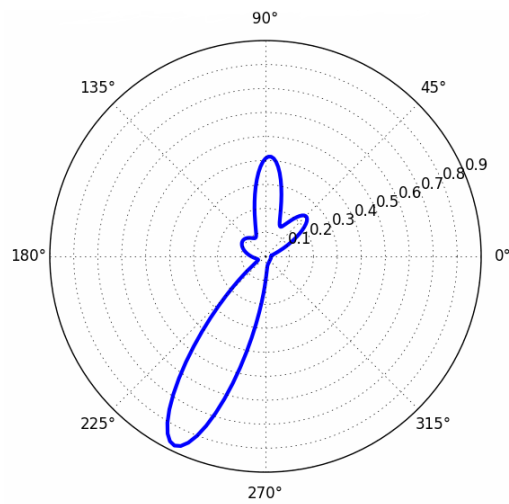
Fuente: Elaboración propia.



(a) Dirección septiembre 2015.



(b) Dirección septiembre 2014.



(c) Dirección septiembre 2013.

Figura 4.6: Gráficos de ajuste por meses en coordenadas polares.
Fuente: Elaboración propia.

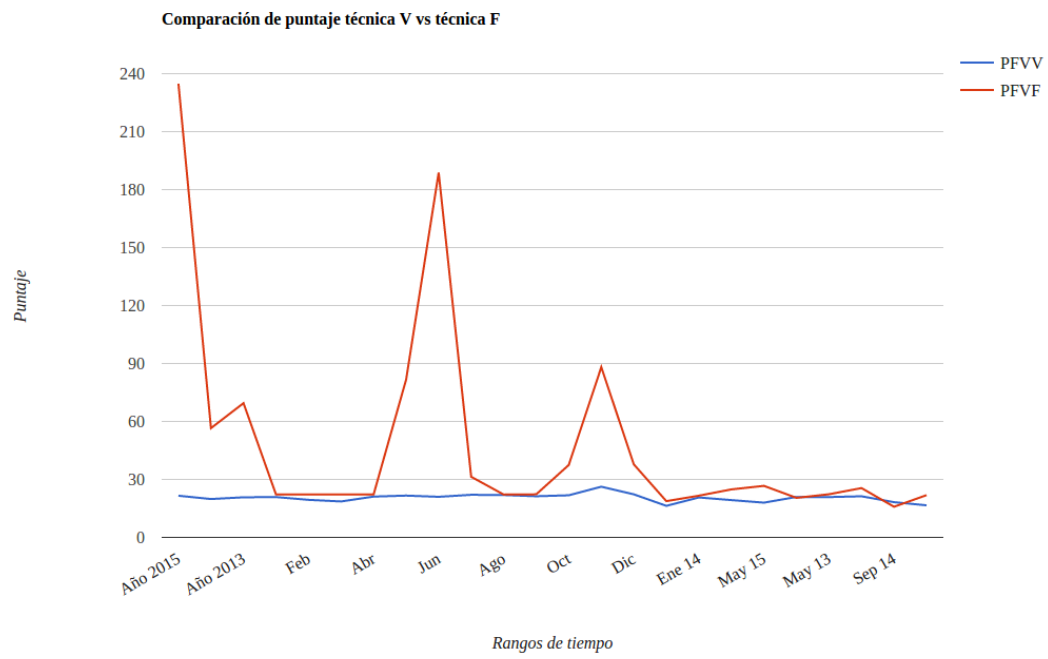
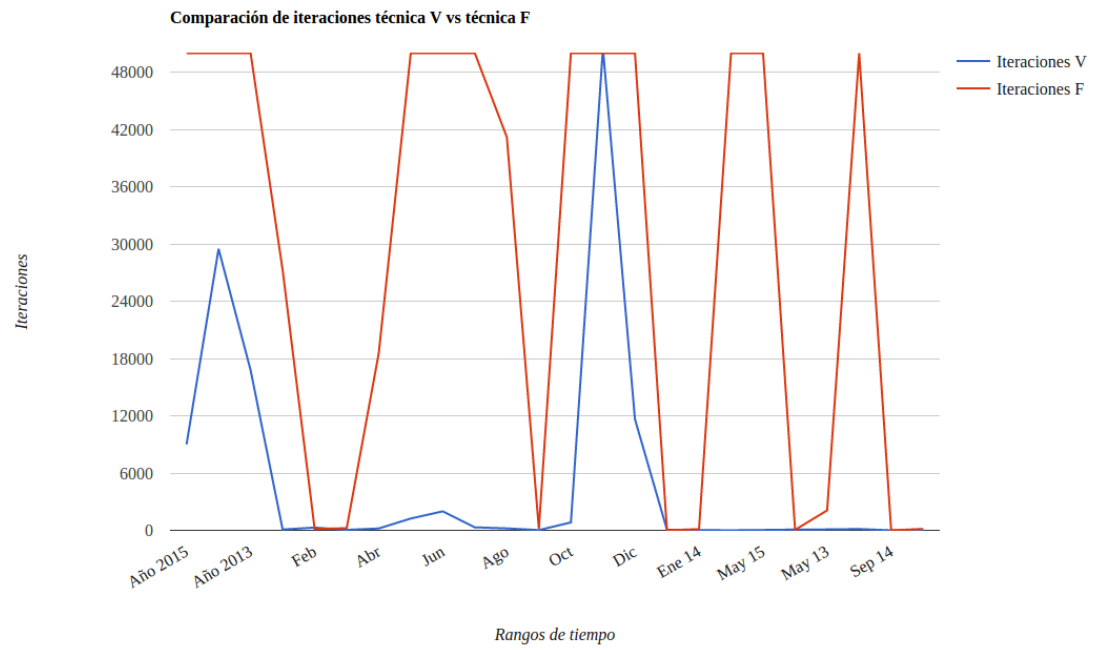


Figura 4.7: Comparación de variaciones en el PSO.
Fuente: Elaboración propia.

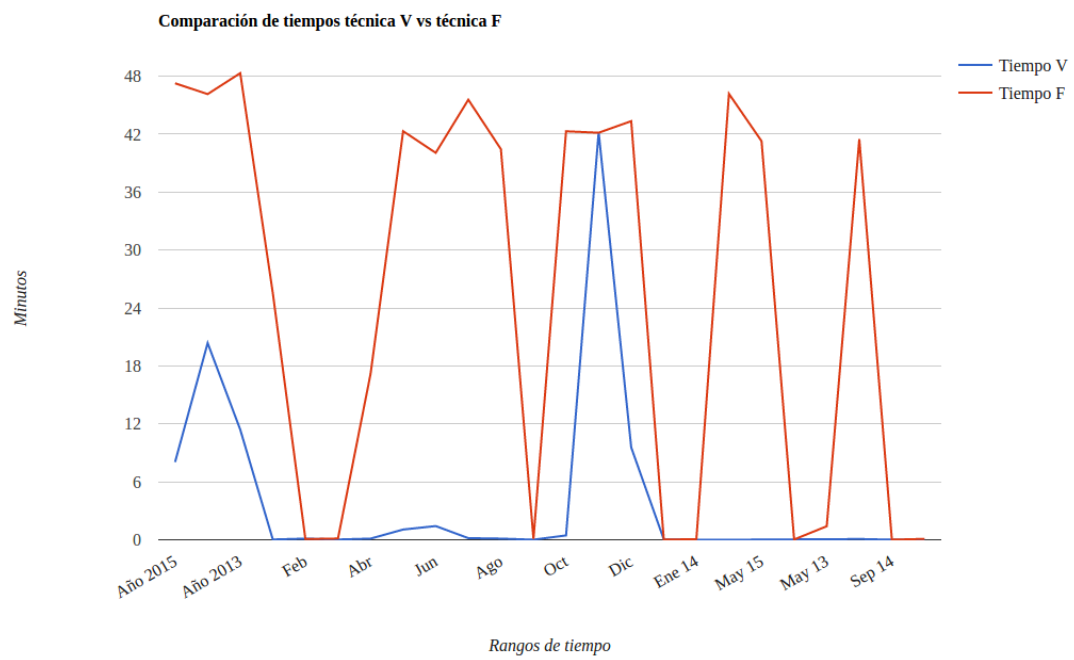


Figura 4.8: Comparación de variaciones en el PSO.
Fuente: Elaboración propia.

Capítulo 5

Aplicaciones

En esta sección se explicará a modo general el uso de la técnica propuesta y las potenciales aplicaciones que tienen los resultados obtenidos.

5.1. Esquema de uso del algoritmo

El esquema 5.1 resume el funcionamiento del algoritmo para el ajuste del modelo probabilístico a los datos del viento. A continuación se detallarán los pasos a seguir en el proceso de obtener el ajuste de parámetros de la distribución elegida.

1. Lo primero que se realiza es el formateo de los datos, es decir, la conversión desde un conjunto de mediciones al formato utilizado por el programa desarrollado, en este caso se utiliza el *comma-separated values* CSV. El *script* utilizado dependerá del formato externo del cual provengan los datos.
2. Una vez obtenidos los datos en el formato deseado, se procede a calcular las frecuencias de los registros obtenidos, para posteriormente ser comparadas con el modelo teórico. El número de frecuencias está determinado por la cantidad de subgrupos definidos que representan intervalos de datos.
3. Antes de comenzar con el ajuste, se calcula una solución inicial ya sea de forma aleatoria (como es el caso para el ajuste de la velocidad del viento) o mediante algún método conveniente (como la aproximación numérica para el caso del ajuste de la dirección del viento).

4. Definiendo la cantidad de partículas y el número de iteraciones, se ejecuta el algoritmo *Particle Swarm Optimization* para que mejore la estimación inicial de acuerdo a una función objetivo previamente definida.
5. Finalmente, una vez que el algoritmo termina, entrega los valores de los parámetros para la función de densidad de probabilidad (fdp) elegida. En esta memoria se utilizaron la distribución de Weibull y la *mixture of von Mises distribution*. Con ello, es posible elaborar un histograma y graficar la fdp para evaluar el ajuste obtenido.

En meteorología, la velocidad del viento se registra en nudos por segundo, pero en esta memoria se convirtieron los datos de acuerdo al sistema internacional SI, o sea, a metros por segundo. La conversión consta en multiplicar el valor de los nudos/segundo por 0.514444 para obtener los metros/segundo correspondientes, como se describe en el libro de unidades internacionales de medidas de la Universidad de Illinois [21].

El sistema de referencia utilizado para la interpretación de los datos de dirección del viento es el conocido como la rosa de los vientos [20]. Dicho sistema ubica el Norte en el grado 0, al Este en el grado 90, al Sur en el grado 180 y al Oeste en el grado 270, siendo la dirección determinada el origen desde donde proviene el viento. Ejemplo, si la dirección predominante del viento en cierto intervalo de tiempo es de 90 grados, entonces diremos que la corriente de viento proviene desde el este.

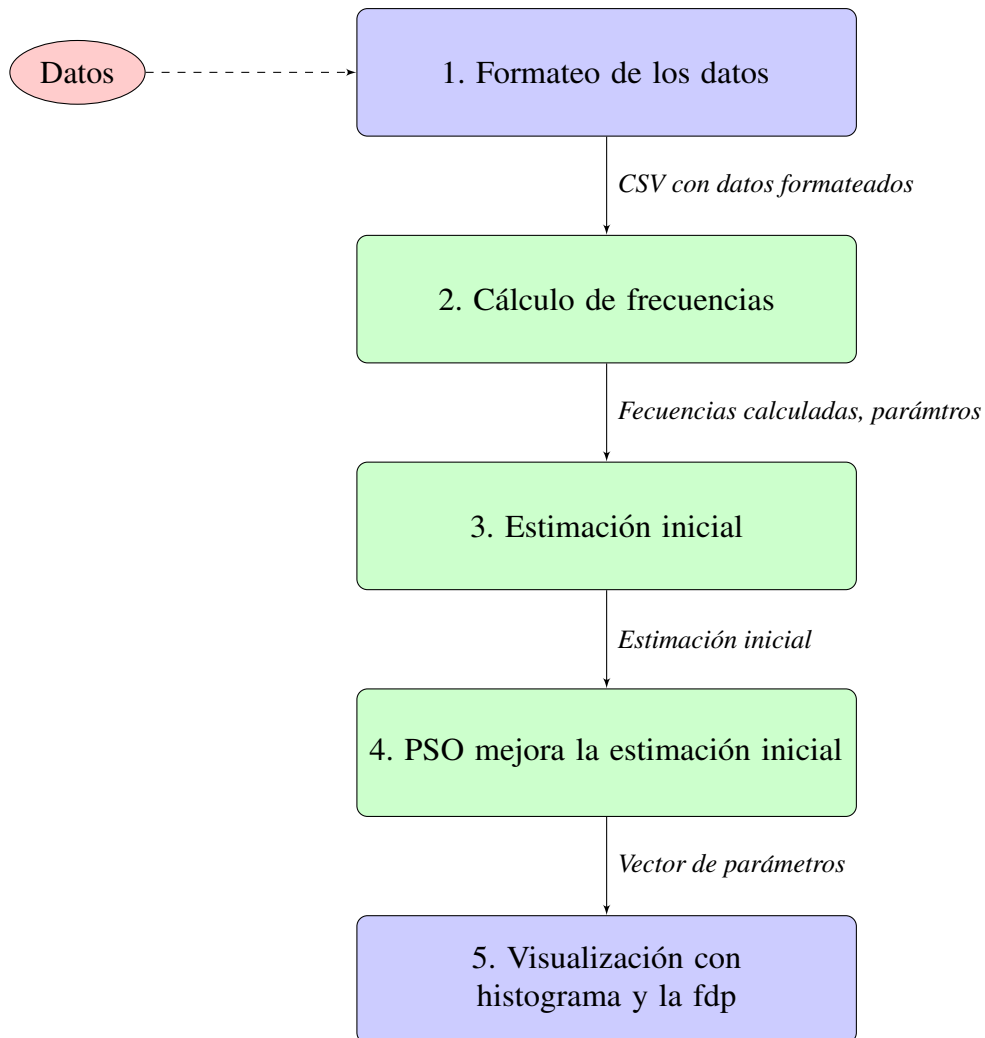
5.2. Uso de los resultados

En la siguiente sección se describirán algunas de las potenciales aplicaciones de los resultados que se obtienen al utilizar las técnicas propuestas en esta memoria.

5.2.1. Energía eólica

En el ámbito de la energía eólica, como se expone en Dabbaghiyan et al. [2], Farade [9], Weissner [10], Masseran et al. [13], Shata et al. [22] y Carneiro et al. [4], la determinación del potencial eléctrico del viento en una región es fundamental al momento de determinar la viabilidad de un proyecto de generación de energías renovables. Así mismo, saber cuáles son las direcciones del viento predominantes en una zona ayudará a determinar el mejor emplazamiento posible para una planta de generación de energía eólica.

Figura 5.1: Esquema de uso del algoritmo



El Dr. Samir Kouro, ingeniero civil electrónico de la Universidad Técnica Federico Santa María, explica desde la perspectiva del tema tratado en esta memoria los aspectos a considerar al momento de evaluar un proyecto de energía eólica:

1. La distribución de Weibull permite obtener información acerca de la magnitudes promedio alcanzadas por la velocidad del viento en un determinado rango de tiempo. De esta forma, se puede evaluar el potencial eléctrico de una región con el fin de determinar la factibilidad de un proyecto eólico. Así mismo, los valores extremos de la distribución de Weibull, en particular los valores máximos, ayudan a determinar las características de las turbinas a implementar, es decir, la resistencia de las torres que sostienen las hélices y la forma de las mismas, para que estas sean resistentes a las fuertes corrientes de viento.

La Figura 5.2, muestra las zonas importantes a considerar acerca del comportamiento de la velocidad del viento.

2. El modelo para la dirección del viento permite determinar las direcciones predominantes de manera de establecer los lugares en donde se ubicaran las turbinas de viento. Un mal emplazamiento de estas implica una disminución en el rendimiento de las turbinas y un desaprovechamiento de la energía obtenida desde el viento.

La figura 5.3 muestra un esquema de emplazamiento de turbinas para aprovechar la dirección del viento.

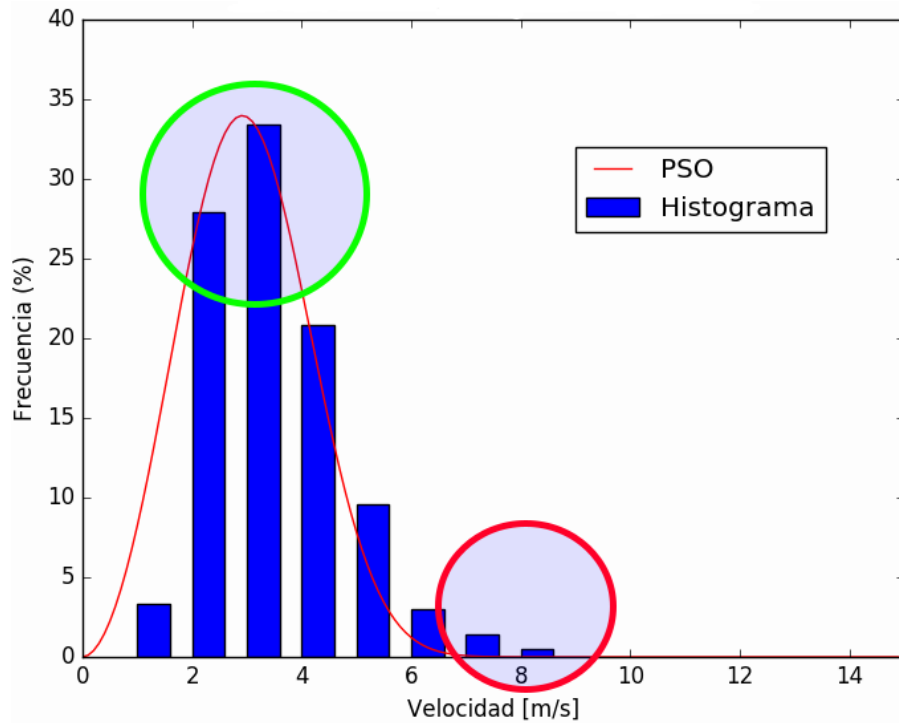


Figura 5.2: Ejemplo de zonas a evaluar en la velocidad del viento.

Fuente: Elaboración propia.

5.2.2. Propagación de incendios

Los incendios son una problemática que muchos lugares del mundo deben combatir debido a diversos factores que los provocan, ya sean de origen natural, debido a las condiciones climáticas, por errores humanos, entre otros. En particular, el grave incendio ocurrido en Valparaíso el mes de abril del año 2014 [24] que provocó severos daños

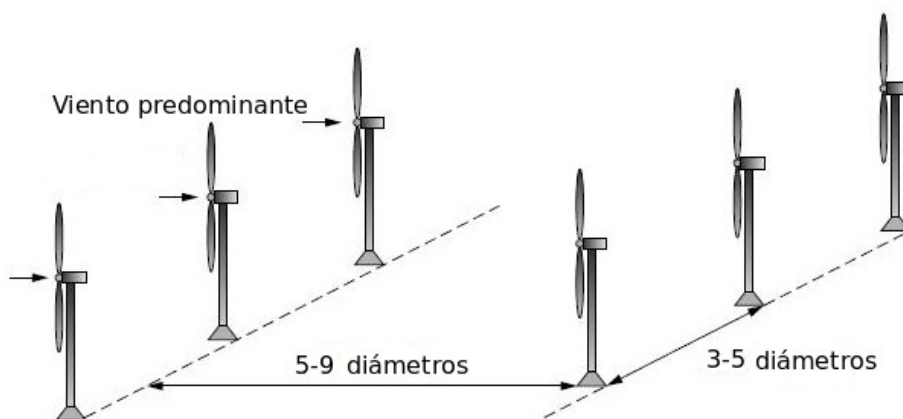


Figura 5.3: Ejemplo emplazamiento turbinas eólicas.
Fuente: Jorge Mírez [23].

a la ciudad, es una muestra de la influencia que tienen las condiciones del clima en la propagación de los incendios. Como se destaca en el artículo [24] las altas temperaturas y los fuertes vientos son un factor negativo que ayuda a los incendios a expandirse y revivir focos parcialmente extinguidos.

En el trabajo de Beer [25], se expone la relación entre el fuego y los vientos, explicando por ejemplo el cómo influye la velocidad de los vientos en la propagación de las llamas. Las cifras que se obtienen de los modelos obtenidos en esta memoria permiten realizar una estimación inicial del comportamiento del viento en, por ejemplo, las fechas recurrentes de incendios forestales, de forma tal que se pueda realizar un plan estratégico que permita abordar estos incidentes de manera más eficiente, disminuyendo el impacto en las zonas afectadas.

5.2.3. Propagación de pesticidas

En el área de la agricultura, la volatilización representa la mayor forma de disipación de pesticidas aplicadas a suelos o cultivos. Como se explica en el trabajo de Bedos et al. [26], generalmente la volatilización de los pesticidas (ver Figura 5.4) en un proceso que dura varias semanas, por lo que es necesario tener un control sobre las variables que inciden en este proceso. Entre los factores atmosféricos, la velocidad del viento tiene una directa incidencia en la razón de volatilización de los pesticidas.

La distribución de Weibull, puede proveer de información relativa a las velocidades de viento predominantes y así ayudar a planificar la aplicación de pesticidas en los cultivos.

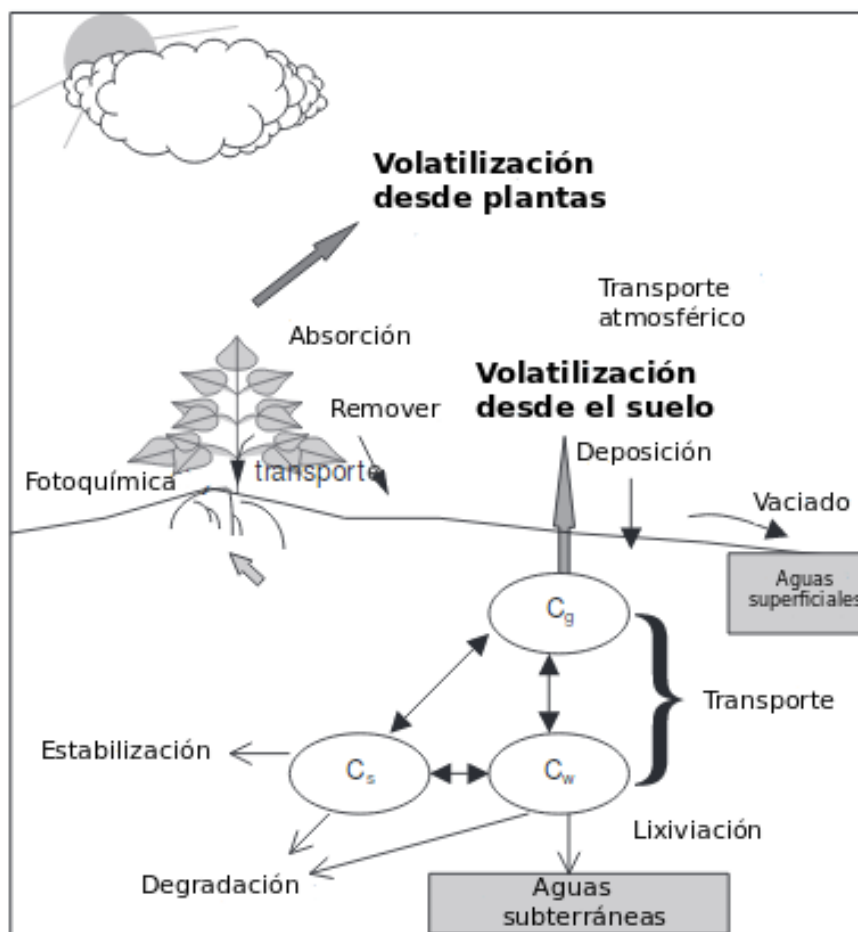


Figura 5.4: Principales procesos que interactúan en el comportamiento de los pesticidas después de su aplicación.
Fuente: Bedos [26].

5.2.4. Análisis atmosférico

En el estudio realizado por García-Portugués et al. [27], se propone un método para explorar la relación entre la dirección del viento y las concentraciones atmosféricas de SO_2 monitoreadas en una estación cercana a una planta de energía en la región de Galicia, España, con el fin de comparar la eficiencia de las medidas de precaución implementadas en ese país para la reducción de la contaminación. Las técnicas propuesta en esta memoria, pueden utilizarse para el modelamiento de la dirección del viento, de manera de aportar información relevante al comportamiento de elementos contaminantes en la atmósfera.

5.2.5. Resumen del capítulo

Los métodos propuestos en esta memoria entregan información teórica basada en datos históricos acerca de la velocidad y la dirección del viento. Dicha información puede ser utilizada principalmente para la evaluación de la factibilidad de distintos proyectos afines. En otros casos, permite realizar un estudio pertinente a las condiciones climáticas, en las que el viento es un factor determinante en el comportamiento del entorno.

Capítulo 6

Conclusiones

En esta memoria se implementaron dos algoritmos para el ajuste de parámetros de distribuciones de densidad de probabilidad a datos de velocidad y dirección del viento.

La distribución de Weibull, es una función utilizada ampliamente para el ajuste de datos de velocidad promedio del viento. Esto se debe principalmente, a que la mayoría de los estudios que utilizan esta distribución, están relacionadas con la evaluación del potencial eléctrico de cierta región. La diferencia entre esos estudios radica esencialmente en el método utilizado para encontrar los parámetros k y c de la distribución de Weibull.

Los resultados obtenidos en esta memoria, confirman que el *Particle Swarm Optimization*, permite encontrar los parámetros de ajuste para la función de Weibull, con un nivel de representación de los datos reales aceptable. Sin embargo, si se desea realizar otras representaciones de la velocidad del viento, que no sean los promedios diarios de las mediciones como base de los datos, se debe modificar la función de Weibull o buscar otra distribución.

La función *mixture of von Mises distribution* es ampliamente utilizada para representar los datos de dirección del viento. A diferencia de la distribución de Weibull, el ajuste de los parámetros de von Mises es un proceso más difícil, que requiere de un cálculo numérico complejo o una estimación inicial débil de la solución buscada para luego ser mejorada con alguna técnica. En esta memoria, se trabaja desde lo realizado por Heckenbergerova et al. [5], en donde no se consiguen los resultados esperados debido a que no se trató el problema de la convergencia prematura del *Particle Swarm Optimization*. Cuando no es suficiente con fijar los parámetros del PSO, se pueden aplicar estrategias como la propuesta en esta memoria en las ecuaciones 4.9, 4.10 y 4.11 para controlar los parámetros del algoritmo durante su ejecución.

Más allá de la distribución a utilizar y del contexto en el que se esté trabajando, la meta-heurística *Particle Swarm Optimization* es una buena herramienta para ajustar parámetros, dado que posee buenos tiempos de ejecución y se adapta fácilmente cambiando la función objetivo. Cuando el espacio de búsqueda es muy amplio, como en el caso del ajuste de la distribución de von Mises, es conveniente guiar la búsqueda a través de soluciones iniciales. Como trabajo futuro queda mejorar la forma de las curvas obtenidas para la dirección del viento, es decir, la similitud con el histograma de datos. Esto podría realizarse mejorando la función objetivo para que tenga en cuenta dicho aspecto. También queda pendiente evaluar otras estrategias para el control de parámetros del PSO, que permitan controlar aspectos como la convergencia prematura.

Bibliografía

- [1] C. O. Santana, M. Falvey, M. Ibarra, and M. García, “Energías renovables en Chile. el potencial eólico, solar e hidroeléctrico de Arica a Chiloé,” *Proyecto Estrategia de Expansión de las Energías Renovables en los Sistemas Eléctricos Interconectados*, pp. 17–54, 2014.
- [2] A. Dabbaghiyan, F. Fazelpour, M. D. Abnavim, and M. a. Rosen, “Evaluation of wind energy potential in province of Bushehr, Iran,” 2015.
- [3] T. P. Chang, “Wind energy assessment incorporating particle swarm optimization method,” 2010.
- [4] T. C. Carneiro, S. P. Melo, P. C. M. Carvalho, and A. P. d. S. Braga, “Particle swarm optimization method for estimation of Weibull parameters: A case study for the Brazilian northeast region,” 2015.
- [5] J. Heckenbergerova, P. Musilek, and P. Krömer, “Optimization of wind direction distribution parameters using particle swarm optimization,” 2015.
- [6] A. Kaveh, “Particle swarm optimization,” pp. 9–40, 2014. Chapter of Advances in Metaheuristic Algorithms for Optimal Design of Structures.
- [7] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proceedings of IEEE international conference on neural networks*, pp. 1942 – 1948, 1995.
- [8] C. Kongnam and S. Nuchprayoon, “A particle swarm optimization for wind energy control problem,” 2015.
- [9] D. Fadare, “A statistical analysis of wind energy potential in Ibadan, Nigeria, based on Weibull distribution function,” 2008.

- [10] D. Weisser, “A wind energy analysis if grenada: an estimation using the weibull density function,” 2002.
- [11] Calimo, “Weibull probability density function.” URL https://commons.wikimedia.org/wiki/File:Weibull_PDF.svg, 2010. [Revisado el 19/06/2016].
- [12] T. P. Chang, “Performance comparison of six numerical methods in estimating weibull parameters for wind energy application,” 2010.
- [13] N. Masseran and A. M. Razali, “Modeling the wind direction behaviors during the monsoon seasons in peninsular malaysia,” 2015.
- [14] A. Banerjee, I. S. Dhillon, J. Ghosh, and S. Sra, “Clustering on the unit hypersphere using von mises-fisher distributions,” *Journal of Machine Learning Research*, 2005.
- [15] J. Heckenbergerova, P. Musilek, J. Mejznar, and M. Vancura, “Estimation of wind direction distribution with genetic algorithms,” 2013.
- [16] J. A. Carta, C. Bueno, and P. Ramírez, “Statistical modelling of directional wind speeds using mixtures of von mises distribution: Case study,” 2007.
- [17] H. P. Gavin, “The levenberg-marquardt method for nonlinear least squares curve-fitting problems,” 2016.
- [18] Y. U. Department of statistics, “Chi-square goodness of fit test.” URL <https://www.stat.yale.edu/Courses/1997-98/101/chigf.htm>, 1998. [Revisado el 31/10/2016].
- [19] E. C. o. S. Department of statistics, “Chi-square distribution table.” URL <http://sites.stat.psu.edu/~mga/401/tables/Chi-square-table.pdf>, 2006. [Revisado el 31/10/2016].
- [20] E. L. Ramírez and S. F. M. Vivanco, “Análisis de parámetros metereológicos utilizando el modelo de la rosa de los vientos,” 2007.
- [21] E. A. Mechtly, “The international system of units. physical constants and conversion factors,” 1973.

- [22] A. A. Shata and H. Rolf, “Evaluation of wind energy potential and electricity generation on the coast of mediterranean sea in egypt,” 2005.
- [23] J. L. M. Tarrillo, “Distribución óptima de torres en un parque eólico.” URL https://jmirez.files.wordpress.com/2011/07/optima_distribucion_parque_eolico.jpg, 2011. [Revisado el 20/11/2016].
- [24] B. Mundo, “Chile: un incendio consume valparaíso.” URL http://www.bbc.com/mundo/ultimas_noticias/2014/04/140412_ultnot_chile_incendio_valparaiso_jgc, 2014. [Revisado el 20/11/2016].
- [25] T. Beer, “The interaction of wind and fire,” 1991.
- [26] C. Bedos, P. Cellier, R. Calvet, E. Barriuso, and B. Gabrielle, “Mass transfer of pesticides into the atmosphere by volatilization from soils and plants: overview,” 2001.
- [27] E. G. Portugués, R. M. Crujeiras, and W. G. Manteiga, “Exploring wind direction and dióxido de azufre concentration by circular-linear density estimation,” 2012.