

Un acercamiento a la detección de patrones frecuentes espacio-temporales para la velocidad y dirección del viento en el departamento de Nariño

GIEE
Universidad de Nariño
San Juan de Pasto, Colombia
Email:

GIIWW
Universidad de Nariño
San Juan de Pasto, Colombia
Email:

Abstract—En este artículo se presentan los componentes de una investigación orientada a analizar los datos de series de tiempo que provee Vaisala Inc, series de viento recopiladas desde el año 1980 hasta 2015 en el departamento de Nariño (Colombia) con el objetivo de descubrir patrones secuenciales en el comportamiento de la velocidad y dirección del viento. Se busca poder predecir con suficiente anticipación anomalías en el incremento o disminución de estas características para contribuir en la identificación de las ubicaciones más idóneas para la implementación de una futura solución eólica como alternativa energética.

Keywords—wind, time series

I. INTRODUCCIÓN

Diferentes alternativas energéticas han sido ampliamente exploradas en la literatura. En particular, la energía eólica ha registrado una importante acogida, en especial en Europa donde en las últimas décadas se puede evidenciar un crecimiento promedio anual de hasta el 30 % [1]. Sin embargo, la implementación de una solución eólica no resulta trivial. La primera etapa a afrontar es la ubicación de potenciales lugares donde las condiciones de viento resulten apropiadas. En este sentido, las principales características a monitorear son las velocidad y dirección del viento y se espera poder ubicar geográficamente aquellos lugares donde estas características registren elevados promedios durante un periodo considerable de tiempo. El proyecto general “Análisis de oportunidades energéticas con fuentes alternativas para el departamento de Nariño” tiene como uno de sus objetivos cumplir con este propósito.

A través de un proveedor externo se tiene acceso a series de tiempo de los últimos diez años con información sobre la ubicación de la muestra (latitud y longitud) y la correspondiente velocidad y orientación del viento. Con esta base de datos resultaría sencillo generar mapas del potencial eólico para el departamento. Sin embargo, también resulta de gran importancia para el proyecto conducir investigación de alto nivel orientada al análisis de los datos generados y disponibles. Esta investigación busca analizar las fuentes de datos disponibles en procura de la extracción de patrones secuenciales en el comportamiento de la dirección y velocidad del viento en aquellos lugares con potencial eólico. Dichos patrones servirán de soporte

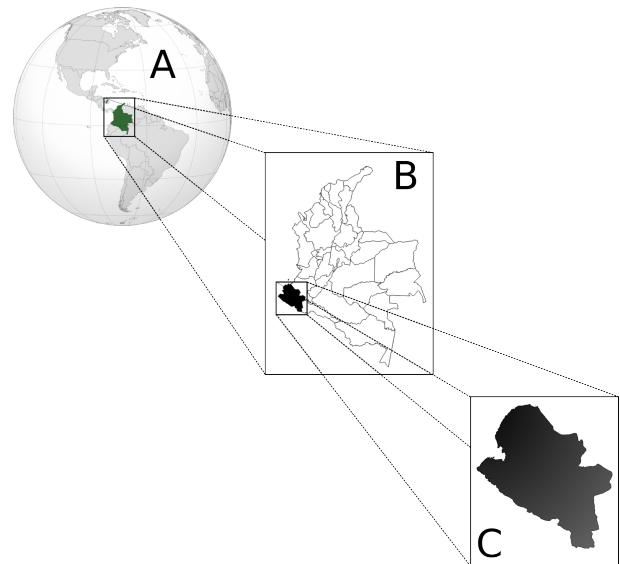


Figura 1. Localización area de estudio

para la detección de anomalías y para la toma de decisiones a la hora de la implementación de una solución final.

El área de estudio de esta investigación fue el departamento de Nariño (Colombia) el cual está ubicado en el extremo sur occidental de Colombia, en la frontera con Ecuador con una extensión aproximada de 33.268 km, una población de 1,702 millones según el censo de 2013, su ubicación está en latitud 00° 31' 08" y 02° 41' 08" Norte, Longitud 76° 51' 19" y 79° 01' 34" Oeste.

II. TRABAJOS RELACIONADOS

Diferentes estudios han explorado la construcción de mapas eólicos a partir de muestras tomadas en terreno. Por ejemplo, [2] utilizan técnicas de auto-correlación espacio-temporal, para estimar la potencia generada por turbinas en Irlanda a partir de pocos datos de entrada. Adicionales técnicas de interpolación espacial fueron utilizadas por [3] para generar mapas de viento de alta resolución en el Reino Unido. El objetivo perseguido por la investigación era comparar y evaluar diversos métodos de interpolación para seleccionar el más adecuado. En los Países Bajos, [4]

utilizan un modelo llamado de bicapa para estimar la velocidad del viento a partir de 31 estaciones meteorológicas.

Sin embargo, es importante resaltar que las características de velocidad y dirección del viento no son los únicos criterios a tener en cuenta a la hora de escoger las mejores ubicaciones. La evaluación multi-criterio (MCE) ha tenido una amplia acogida a la hora de evaluar características físicas junto con otros atributos como aspectos económicos y sociales.

[5] utilizan técnicas MCE para determinar posibles ubicaciones de turbinas en el norte de California evaluando componentes físicos, ambientales y humanos. [6] categoriza diferentes aspectos de acuerdo al potencial eólico y solar e identifica áreas susceptibles a la instalación de turbinas y paneles solares en Colorado.

Entre los aspectos evaluados se encuentran la distancia a carreteras y líneas de transmisión eléctrica, coberturas del terreno, densidad de población y áreas protegidas por la ley. [7] explora nuevos algoritmos de análisis de datos para modelar la ubicación de turbinas en Iowa apoyándose en un sistema espacial multi-criterio de soporte a la toma de decisiones. Nuevas técnicas de análisis de datos, comúnmente conocidas como minería de datos, han demostrado muy buenos resultados a la hora de modelar fenómenos atmosféricos.

Por ejemplo, [8] utiliza algoritmos para detectar patrones secuenciales en series de tiempo de viento desde estaciones en los Países Bajos para detectar anomalías en el flujo, velocidad y dirección del viento.

III. MEDOTLOGÍA

A. Obtención de datos

La primera parte de la investigación consistió en la identificación y exploración de las fuentes de datos disponibles. Se utilizó las bases de datos eólicas de un proveedor externo (Vaisala Inc) y se trabajó en una estrategia de muestreo significativa de 480 series de tiempo regularmente espaciados en el departamento de Nariño, con la cual se espera construir los mapas de potencial eólico necesarios a 50 y 120 metros de altura.

Las series de tiempo son conjuntos de reáñalisis de datos MERRA (Modern Era Retrospective-analysis for Research and Applications) la cual tiene una mayor resolución espacial y temporal (por hora), cada punto contiene atributos como: latitud, longitud, timestamp, presión, temperatura, velocidad y dirección del viento

B. Procesamiento de datos

Las series de tiempo fueron procesadas para ser subidas en una base de datos, la figura 2 muestra el modelo entidad relación con las siguientes tablas:

Tabla stations: esta tabla contiene las 480 estaciones con latitud y longitud, de las series de tiempo entregadas por Vaisala Inc.

Tabla timeseries50 y timeseries120: se guardan las series de tiempo desde el año 1980 a 2014 a 50 y 120 metros de altura respectivamente.

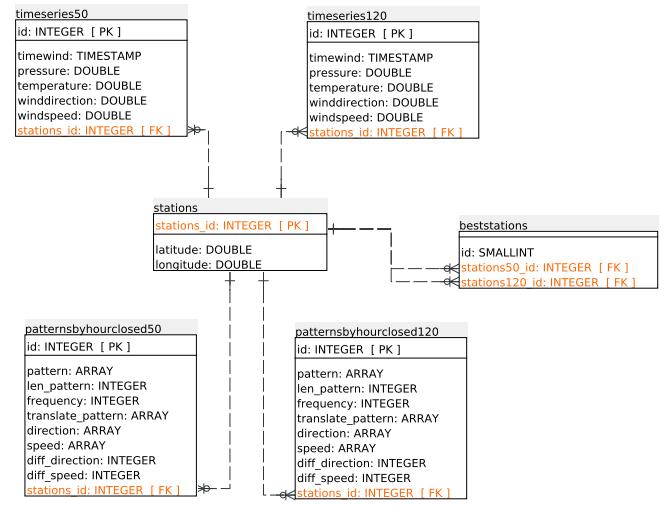


Figura 2. Modelo entidad relación wind

Tabla beststations: en esta tabla se guarda las estaciones que con un análisis muestra las mejores 50 estaciones.

Tabla patternbyhourclosed50 y patternbyhourclosed120: en esta tabla se guardaran los patrones detectados en el análisis de asociación que se mostrará mas adelante con las series de tiempo a 50 y 120 metros de altura.

C. Construcción mapas de velocidad de viento

Para la construcción de mapas de biomasa se utilizó el método Kriging que provee una solución al problema de la estimación basada en un modelo continuo de variación espacial estocástica, el objetivo de Kriging es el de estimar el valor de una variable aleatoria, Z, en uno o más puntos no muestreados o sobre grandes bloques.

El método Kriging recibe como entrada datos de la muestra, y una malla dependiendo de la resolución que se quiera obtener, por ello los datos de muestra se obtuvieron al descomponer la dirección del viento en sus componentes rectangulares multiplicarlas por la velocidad y luego agruparlos en cada punto por mes, año y uno general.; y la malla se construyó con puntos regulares espaciados cada 450 metros.

En la figura 3, figura 4, figura 5 se muestra los mapas obtenidos por meses, años y general entre el año 2000 a 2014 respectivamente.

D. Detección de patrones

Para la detección de patrones se utilizó el algoritmo lcm_seq propuesto por [9], este algoritmo es una variación de LCM para minería LCM_seq is a variation of LCM for minería de secuencias. Este algoritmo encuentra todos los patrones de secuencias que aparecen con frecuencia en una base de datos dada, cada transacción es considerada una secuencia. El algoritmo sigue el esquema llamado lapso de prefijo, pero las estructuras de datos y método de procesamiento se basan en LCM.

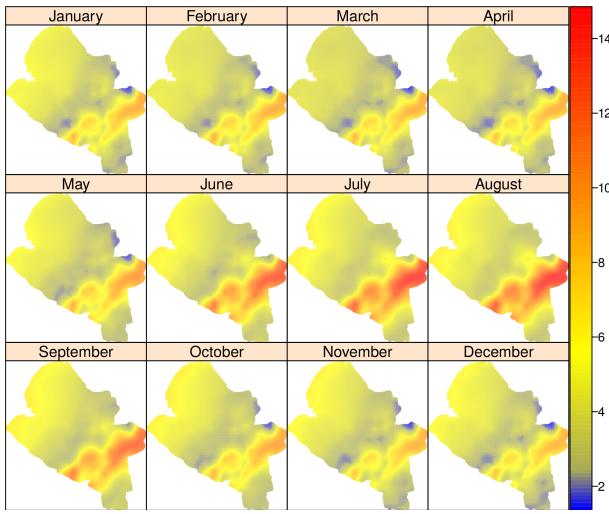


Figura 3. Velocidad del viento por meses a 50m

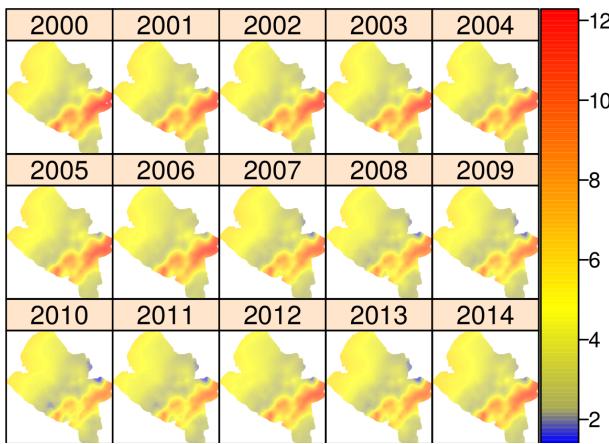


Figura 4. Velocidad del viento por años a 50m

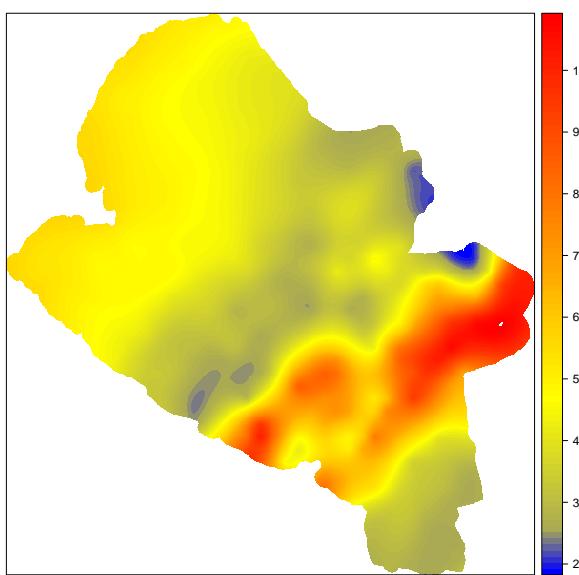


Figura 5. Velocidad del viento general a 50m

Tabla I. IDs DE VELOCIDAD Y DIRECCIÓN DE VIENTO PARA CADA GRUPO

ID_Speed	Speed(m/s)	ID_Direction	Degree
1	0 - 0.3	1	337.5 - 22.5
2	0.3 - 1.6	2	22.5 - 67.5
3	1.6 - 3.4	3	67.5 - 112.5
4	3.4 - 5.5	4	112.5 - 157.5
5	5.5 - 8.0	5	157.5 - 202.5
6	8 - 10.8	6	202.5 - 247.5
7	10.8 - 13.9	7	247.5 - 292.5
8	13.9 - 17.2	8	292.5 - 337.5
9	17.2 - 20.8		
10	20.8 - 24.4		
11	24.4 - 28.5		
12	28.5 - 32.7		
13	32.7 - inf		

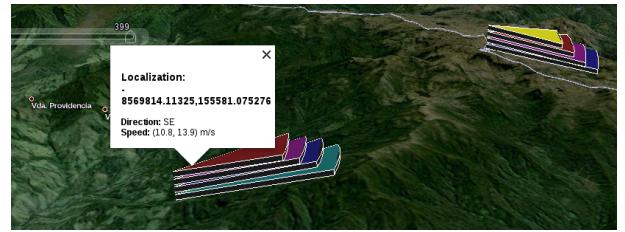


Figura 6. Patrón con diferentes velocidades a 50 metros

Para aplicar este algoritmo se construyó un script el cual convierte los datos de las series de tiempo al formato de LCM_seq, para esto se categorizó la dirección del viento en 8 bins, y la velocidad del viento en 13 bins, como lo muestra la tabla I, al categorizar según el valor de entrada se concatena el ID de dirección y de velocidad del viento.

Los resultados de esta ejecución exigirán una etapa de interpretación y discusión de los resultados donde se deberán seleccionar o implementar diferentes técnicas de visualización. Frecuentemente, al aplicar técnicas de minería de datos se produce una gran cantidad de resultados donde las técnicas de visualización y filtrado resultan fundamentales para el la compresión del conocimiento generado.

E. Visualización de patrones

Para la visualización de patrones se tuvo en cuenta cuatro atributos de los patrones encontrados, como lo es la frecuencia, longitud del patrón, diferente velocidad y diferente dirección. Para esta visualización se construyeron archivos KML(Keyhole markup language) el cual es un lenguaje marcado basado en XML para representar datos geográficos en tres dimensiones, los KML's se que se construyeron fueron con las 50 estaciones de mayor potencial, en la figura 6 se puede mirar el patrón que tiene diferentes de velocidades a 50 metros, los KML's que se construyeron se los puede descargar del repositorio del proyecto ¹.

AGRADECIMIENTOS

Universidad de Nariño, Universidad de los Andes y Sistema de Regalías.

REFERENCIAS

- [1] EWEA, "Wind industry factsheets," European Wind Energy Association (EWEA), Technical report, 2004. [Online]. Available: <http://www.ewea.org>

¹<https://github.com/poldrosky/alternar>

- [2] J. Haslett and A. E. Raftery, "Space-time modelling with long-memory dependence: Assessing ireland's wind power resource," *Applied Statistics*, vol. 38, no. 1, p. 1, 1989. [Online]. Available: <http://www.jstor.org/discover/10.2307/2347679?uid=3737808&uid=2134&uid=2478620387&uid=2&uid=70&uid=3&uid=2478620377&uid=60&sid=21103443692791>
- [3] W. Luo, M. C. Taylor, and S. R. Parker, "A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from england and wales," *International Journal of Climatology*, vol. 28, no. 7, pp. 947–959, jun 2008. [Online]. Available: <http://doi.wiley.com/10.1002/joc.1583>
- [4] A. Stepek and I. Wijnant, "Interpolating wind speed normals from the sparse dutch network to a high resolution grid using local roughness from land use maps," Royal Netherlands Meteorological Institute, De Bilt, Utrecht, Netherlands., Tech. Rep. TR-321, jun 2011. [Online]. Available: <http://www.knmi.nl/knmi-library/knmipubTR/TR321.pdf>
- [5] L. C. Rodman and R. K. Meentemeyer, "A geographic analysis of wind turbine placement in northern california," *Energy Policy*, vol. 34, no. 15, pp. 2137–2149, oct 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0301421505000893>
- [6] J. R. Janke, "Multicriteria GIS modeling of wind and solar farms in colorado," *Renewable Energy*, vol. 35, no. 10, pp. 2228–2234, oct 2010. [Online]. Available: <http://linkinghub.elsevier.com/retrieve/pii/S096014811000131X>
- [7] A. N. Petrov and J. M. Wessling, "Utilization of machine-learning algorithms for wind turbine site suitability modeling in iowa, USA," *Wind Energy*, 2014. [Online]. Available: <http://onlinelibrary.wiley.com/doi/10.1002/we.1723/abstract>
- [8] N. Yusof, R. Zurita-Milla, M.-J. Kraak, and B. Retsios, "Mining frequent spatio-temporal patterns in wind speed and direction," in *Connecting a Digital Europe Through Location and Place*, ser. Lecture Notes in Geoinformation and Cartography, J. Huerta, S. Schade, and C. Granell, Eds. Springer International Publishing, jan 2014, pp. 143–161. [Online]. Available: http://link.springer.com/chapter/10.1007/978-3-319-03611-3_9
- [9] T. Uno. lcm_seq readme. [Online]. Available: http://research.nii.ac.jp/~uno/code/lcm_seq.html