

Short Term Demand Forecasting methodology for Power Decision Making Based on Markov Chain. Study Case - EEQ

Metodología basada en Cadenas de Markov para la Predicción de la Demanda y Toma de Decisiones en el corto plazo. Caso de Estudio: Empresa Eléctrica Quito

Roberto Sanchez¹Patricio Barrera¹

¹ Operador Nacional de Electricidad, CENACE, Quito, Ecuador
E-mail: rsanchez@cenace.org.ec; pbarrera@cenace.org.ec

Abstract

This investigation is focused on the prediction of the electrical demand in short time. For this purpose, the “demand profiles” and the real time signal of the electrical demand of the Empresa Eléctrica Quito S.A. are used in order to determine which profile is expected to happen during the day. In this sense, this study uses the Hidden Markov Model for forecasting the electrical demand in short time. This approach first applies a learning/training process using data from the Sistema de Información Validada Operativa (SIVO). Later, a discovery process of demand profiles is performed in order to determine the most expected profile to happen during the day. This approach establishes an “expected demand area” that shall be a reference for the definitive behavior of the electrical demand.

This methodology was applied over the EEQ system and evaluated during 30 days. The final tool successes 86% of the cases and the actual value of the electrical demand in real time is inside of the band of the expected demand area.

The purpose of this work is to build an application that assist operators of the National Interconnected System, NIS, to make the decisions in short time, optimizing the resources for generation.

Index terms— prediction of electrical demand, machine learning, hidden markov model, artificial intelligence.

Resumen

La investigación del presente trabajo está centrada en determinar el pronóstico de la demanda de potencia eléctrica en corto plazo. Para ello, se utilizó y se comparó los “perfiles de demanda” y la señal en tiempo real de la demanda eléctrica de la Empresa Eléctrica Quito S.A, EEQ, para llegar a determinar el perfil más esperado en el día. En este sentido, se utilizó el Modelo Oculto de Markov (Hidden Markov Model, HMM) para el pronóstico de la demanda en horizonte de tiempo de corto plazo. Para esto, primeramente se realizó un proceso de aprendizaje/entrenamiento al modelo con la base de datos Sistema de Información Validada Operativa, SIVO. Posteriormente, se realizó el proceso de descubrimiento de perfiles de demanda, que permitirá en pasos posteriores encontrar el perfil más esperado a ocurrir durante el día. La propuesta establece un “área de demanda esperada” que se convierte en una referencia que define el comportamiento de la demanda lo largo del día.

Se realizó una evaluación en un periodo de 30 días de la metodología aplicada al sistema de la EEQ, y se observó que la herramienta acierta en un 86% de los casos y el valor de demanda en tiempo real se encuentra dentro de la banda de demanda esperada.

El propósito de este trabajo es brindar una aplicación a los operadores del Sistema Nacional Interconectado, SNI, del Operador Nacional, CENACE, que permita tomar decisiones en el periodo de corto plazo optimizando los recursos de generadores existentes.

Palabras clave— Predicción de Demanda, Machine Learning, Modelo de Markov, artificial intelligence.

Recibido: 14-05-2018, Aprobado tras revisión: 24-07-2018

Forma sugerida de citación: Sanchez, R.; Barrera, P. (2018). “Metodología basada en Cadenas de Markov para la Predicción de La Demanda y Toma de Decisiones en el corto plazo. Caso de Estudio: Empresa Eléctrica Quito”. Revista Técnica “energía”. No. 15, Issue I, Pp. 44-50

ISSN On-line: 2602-8492 - ISSN Impreso: 1390-5074

© 2018 Operador Nacional de Electricidad, CENACE

1. INTRODUCCIÓN

Uno de las componentes esenciales para el desarrollo económico de un país es la electricidad, la disponibilidad de fuentes suficientes y confiables de energía a menor costo. En base a estos antecedentes, es indispensable poder determinar el comportamiento actual y futuro de la demanda eléctrica, considerando los posibles impactos producidos por variables ambientales, económicas, sociales, tecnológicas y políticas.

La exigencia en el pronóstico de la demanda eléctrica, permite a las empresas del sector eléctrico, planificar sus operaciones, identificar futuros requerimientos de infraestructura, coordinar el funcionamiento de los recursos de generación existentes, permitiendo optimizar los problemas operativos y económicos, que una mala predicción podría presentar.

En el presente trabajo se entrenaron y validaron, diferentes Modelos Ocultos de Markov (Hidden Markov Model, HMM) para el pronóstico de la demanda en horizonte de tiempo de corto plazo. Los modelos HMM han sido entrenados utilizando técnicas de aprendizaje automático no supervisado. Se desea en una siguiente entrega de este artículo, implementar modelos que utilicen técnicas de aprendizaje automático supervisado basándose en los resultados de este estudio.

Un HMM es un modelo que genera distribuciones probabilísticas, basándose en la información proveniente de la secuencia de muestras observadas. Debido a su capacidad de detectar secuencias temporales, este modelo es ampliamente utilizado para el descubrimiento de patrones temporales en aplicaciones como reconocimiento del habla, escritura y gestos, y además, reconocimiento del ADN humano [1, 2, 4].

Para el desarrollo de esta investigación, se utilizó los datos de demanda de la Empresa Eléctrica Quito S.A, EEQ, con resolución media horaria de los años 2014 hasta 2017 obtenidas de la base de datos Sistema de Información Operativa, SIVO, de la Subgerencia de Análisis de la Operación de CENACE (SAO).

Mediante este método se predice la demanda de la EEQ a corto plazo, en base a la demanda de tiempo real obtenida de un PI-Server. De esta manera, la propuesta se convierte en una herramienta de pronóstico de demanda diario para el análisis de soluciones en el despacho de energía.

2. METODOLOGÍA

2.1. Definiciones

Definiciones y conceptos propios utilizados para la descripción de la metodología:

Muestra de demanda media-horaria (O_i): Vector medio-horario que describe el comportamiento de la demanda durante un día (48 valores).

Perfiles de demanda (P_i): Grupo de muestras de demanda media-horaria, que tienen un comportamiento similar. Se encuentra definido por el vector promedio μ , acompañado del vector de desviación estándar σ , cada desviación estándar asociada a su valor promedio:

$$P_i = [(\mu_1, \mu_2, \dots, \mu_{48}), (\sigma_1, \sigma_2, \dots, \sigma_{48})] \quad (1)$$

Familia de perfiles de demanda: Una familia de perfiles asocia un conjunto de perfiles de demanda, donde los perfiles pertenecientes a dicha familia tiene un grado de similitud asociado.

Perfil de demanda esperado (P_e): Dada la variable en tiempo real de la demanda X con h valores desde las 00:00:00 del día. El perfil de demanda esperado, es aquel que se acerca lo más posible a la señal de tiempo real. Se lo determina a través de minimización del error por partes entre la colección de perfiles de demanda μ_i y la variable en tiempo real:

$$ID_{P_e} = \operatorname{argmin}_{i=1}^K [\sum_{j=1}^h (\mu_{ij} - X_j)^2] \quad (2)$$

Desviación estándar de la demanda esperada: Es el vector σ del perfil de demanda esperado.

Área de demanda esperada: Área definida por el perfil de demanda esperado, donde el límite superior e inferior se calculan usando la desviación estándar y su valor promedio:

$$L_{min} = \{\mu_i - \alpha \cdot \sigma_i \quad \forall i \in [1, 48]\} \quad (3)$$

$$L_{max} = \{\mu_i + \alpha \cdot \sigma_i \quad \forall i \in [1, 48]\} \quad (4)$$

Aprendizaje supervisado: Técnica del aprendizaje automático que mapea un vector de entrada con una salida deseada. El entrenamiento de los modelos asociados se realiza a través de muestras entrada-salida.

Aprendizaje no supervisado: Técnica del aprendizaje automático que infiere una función/modelo para describir estructuras ocultas de un conjunto de datos no etiquetado.

2.2. Metodología

La metodología empleada usa el aprendizaje automático no supervisado para el descubrimiento de patrones típicos y singulares de la demanda, luego usando técnicas de agrupamiento jerárquico, se determinan las familias de demanda existentes. Finalmente la señal en tiempo real es buscada dentro de una familia de perfiles de demanda, el perfil de demanda más acercado es usado para calcular el área de demanda esperada. En la Fig. 1 se ilustra la metodología.



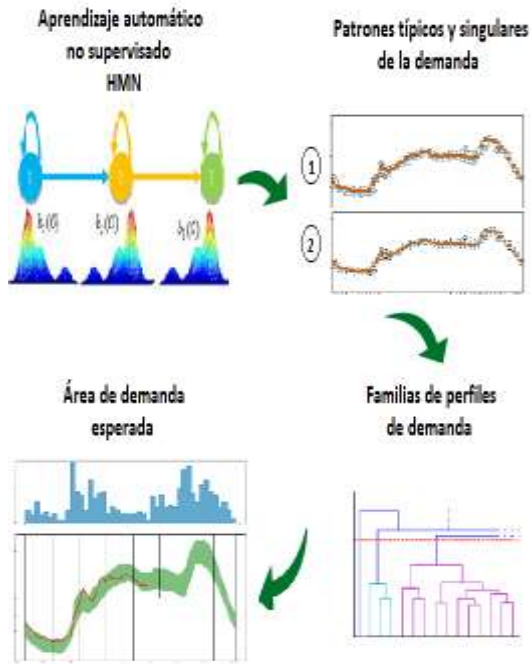


Figura 1: Metodología aplicada

2.2.1 Descubrimiento de perfiles de demanda

Para realizar el descubrimiento de perfiles de demanda, se realizó el entrenamiento y selección del mejor modelo oculto de Markov. El detalle del desarrollo de este modelo es extenso y se puede encontrar un excelente tutorial en [3], y literatura que describe sus detalles en [1],[4] y [7].

En los siguientes párrafos se incluye los conceptos más importantes de dicho modelo, con el objetivo de explicar la metodología seguida. Se Define el HMM siguiendo la nomenclatura descrita en [7]:

$\mathbf{O} = (\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_T)$ es la secuencia observada. Cada muestra (i.e. $\mathbf{O}_n \in [1, T]$) es una muestra de demanda media-horaria, donde T es el número total de muestras.

K es el número de estados ocultos del HMM.

$\mathbf{S} = (\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_T)$ es la secuencia de estados ocultos de Markov asociados a la secuencia observada \mathbf{O} .

$\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_K)$ es la probabilidad inicial del modelo donde $\pi_i = P(\mathbf{S}_{i=1} | \mathbf{O}_1)$ y $\sum_{i=1}^K \pi_i = 1$

$a_{i,j}$ es la probabilidad de transición desde el estado oculto \mathbf{S}_i hasta el estado oculto \mathbf{S}_j , denotado como: $a_{i,j} = P(\mathbf{S}_{t+1}=j | \mathbf{S}_t=i)$.

$\mathbf{A} = \{a_{ij} | i \in [1, K]; j \in [1, K]\}$ es la matriz de probabilidades de transición de los estados ocultos, donde $\sum_{j=1}^K a_{ij} = 1$.

$b_{k,t} | t \in [1, T]$ es la probabilidad de que una muestra \mathbf{O}_t haya sido emitida por el estado \mathbf{S}_k en el tiempo t.

$\mathbf{B} = \{b_{k,t} | k \in [1, K]; t \in [1, T]\}$ es la matriz de probabilidad de emisiones de muestras, donde $b_{k,t} = P(\mathbf{O}_t | \mathbf{S}_k)$. Típicamente se asume una distribución Gaussiana multivariante.

$\boldsymbol{\lambda} = \{\boldsymbol{\pi}, \mathbf{A}, \mathbf{B}\}$ son los vectores/matrices que especifican el modelo HMM.

Un modelo HMM realiza tres suposiciones referentes a la cadena de Markov oculta:

Suposición del horizonte limitado: la probabilidad de estar en un estado en tiempo t (\mathbf{S}_t) depende solamente del estado en el tiempo t-1 (\mathbf{S}_{t-1}). Se asume que el estado \mathbf{S}_{t-1} tiene suficiente información para predecir el futuro. Formalmente:

$$P(\mathbf{S}_t | \mathbf{S}_{t-1}, \mathbf{S}_{t-2}, \dots, \mathbf{S}_1) = P(\mathbf{S}_t | \mathbf{S}_{t-1}) \quad (5)$$

Suposición del proceso estacionario: la distribución condicional sobre un siguiente estado dado no cambia a pesar del tiempo. Es decir, se asume que las probabilidades de transición son independientes del tiempo real en el que se dio dicha transición:

$$P(\mathbf{S}_{t_1+1}=j | \mathbf{S}_{t_1}=i) = P(\mathbf{S}_{t_2+1}=j | \mathbf{S}_{t_2}=i) \quad (6)$$

$$t_1, t_2 \in [2, T] \wedge t_1 \neq t_2$$

Suposición de la independencia de las salidas: la salida actual (\mathbf{O}_i) es estadísticamente independiente de las salidas previas ($\mathbf{O}_{i-1}, \mathbf{O}_{i-2}, \dots, \mathbf{O}_1$), formalmente:

$$P(\mathbf{O} | \mathbf{S}, \boldsymbol{\lambda}) = \prod_{t=1}^T P(\mathbf{O}_t | \mathbf{S}_t, \boldsymbol{\lambda}) \quad (7)$$

Finalmente, un modelo HMM dadas las tres suposiciones anteriores permite resolver tres tipos de problemas:

- Proceso de aprendizaje/ entrenamiento: El modelo HMM es ajustado en función de las muestras O. Es decir, dada la secuencia O, se determina el mejor modelo con parámetros $\boldsymbol{\lambda}$ que se ajuste a la secuencia.
- Proceso de evaluación: Dada la secuencia de observaciones O, la secuencia de estados S y los parámetros $\boldsymbol{\lambda}$, se determina la probabilidad de que dicho modelo haya generado la secuencia O dada la secuencia S, es decir, resolver la ecuación (7).
- Proceso de decodificación: Se determina la secuencia de estados S “más probable” que generó la secuencia de observación O. Es decir, encontrar la secuencia S que maximiza el



número de estados correctos correspondientes a las observaciones. Este problema es resuelto a través del algoritmo de Viterbi [5]

Sánchez en su trabajo [6] muestra como se puede usar un modelo HMM para encontrar patrones temporales a través de resolver los problemas descritos en el siguiente orden: 1. aprendizaje, 2. evaluación y 3. decodificación. En su trabajo se usa un 75% del conjunto de datos para el entrenamiento, un 25% para la evaluación, y finalmente, se realiza la decodificación total del conjunto de datos para el agrupamiento de muestras. Para este artículo, se ha procedido de la misma manera, de forma que, en el proceso de aprendizaje, las observaciones de entrada son las muestras de demanda media-horaria y se entrenan los modelos usando el algoritmo de maximización de la expectativa (EM) para encontrar los parámetros λ [5, 7]. El número de muestras del conjunto de datos es de 1460 correspondiente a 4 años (2014-2017) de demanda de la EEQ. En el proceso de evaluación, un total de 400 modelos son entrenados, de los cuales se selecciona el modelo cuyo $\log(P(O | S, \lambda))$ es el mayor [5]. Para evaluar la calidad de los perfiles de demanda, se utiliza una box plots [8] con el objetivo de observar la varianza de dichos agrupamientos. La Fig. 2 muestra un ejemplo de perfil de demanda identificado con el ID=5 donde 68 días han tenido un comportamiento similar. Se observa que la desviación estándar no excede los 20 MW, para este perfil de demanda. Un modelo poco entrenado, dará desviaciones superiores y por ende su agrupamiento no será adecuado.

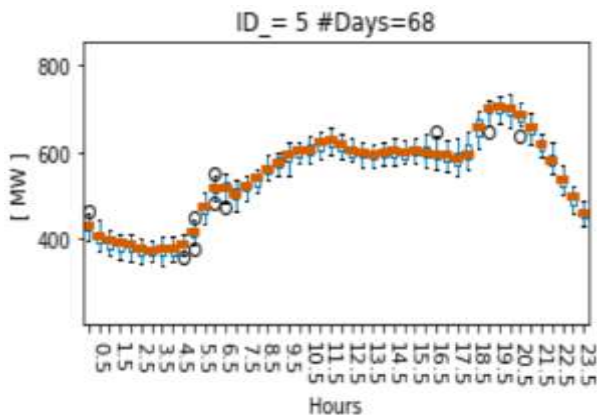


Figura 2: Perfil de demanda cuyo ID=5, donde 68 días tienen un comportamiento similar

La manera de agrupar los días, se realiza a través del proceso de decodificación. Del modelo mejor entrenado se obtiene un total de $K=57$ agrupamientos. Es decir existe una secuencia S de longitud T que es la permutación de estos K estados, cada uno correspondiente a cada observación, creando de esta manera un mapeo entre observaciones y los estados ocultos que “generaron” estas muestras. La Fig. 3 muestra un ejemplo hipotético donde se observa cómo se asocian el agrupamiento ID=2, en los cuales tres días son similares.

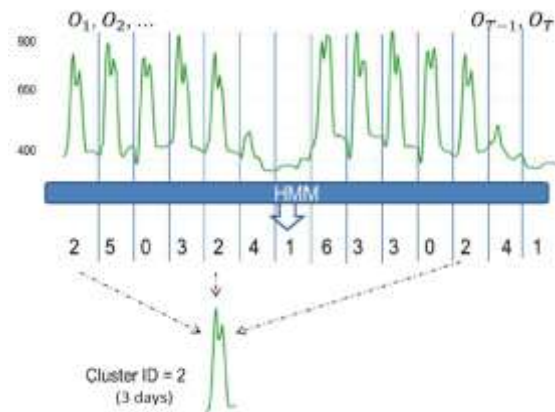


Figura 3: Ejemplo hipotético. La secuencia observada O ha sido generada por la secuencia de estados ocultos $S=[2, 5, 0, 3, 2, \dots, 4, 1]$. El agrupamiento ID=2 tiene 3 días similares [6].

2.2.2 Descubrimiento de las familias de perfiles de demanda

Sánchez [6] en su trabajo propone el uso de un *Hierarchical Agglomerative Clustering* (HAC), con el objetivo de realizar una comparación entre agrupamientos. En base de lo propuesto, para determinar si la representación del HAC es adecuada, se procedió a evaluar el agrupamiento usando el *Cophenetic correlation coefficient* [9]. El resultado final permite la construcción de un dendograma, donde se aprecian las familias de perfiles de demanda. La Fig. 4 muestra un extracto del dendograma total. En el eje X se encuentra los $K=57$ agrupamientos, mientras que en el eje Y la distancia de disimilitud euclidiana entre agrupamientos. Así por ejemplo, el perfil 39 es muy parecido al perfil 40, y ambos son muy diferentes del perfil 38, que pertenece a otra familia. El racimo en color magenta es un ejemplo de familia de perfil. Esta familia es muy categórica ya que está conformado en un 99.5% de perfiles de demanda correspondientes al día lunes, mientras que el racimo en color cian está conformada por días jueves de máxima demanda. Adicionalmente el perfil 55 contiene días atípicos correspondiente a días feriados. El umbral que permite la distinción de familias es un valor empírico, que es determinado por el analista. Para el ejemplo, un valor de 250, permitió la creación de 6 familias.

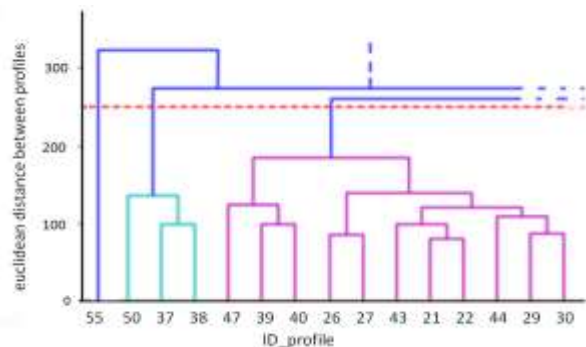


Figura 4: Dendograma. Dos familias de perfiles de demanda (cian y magenta) con distancia euclidiana menor a 250.

2.2.3 Pronóstico de la demanda en tiempo real

Dadas las familias de demanda y la variable en tiempo real, se puede llegar a determinar el perfil de demanda esperado. Para ello se aplica la fórmula (2) y posteriormente se procede al cálculo del área de demanda esperada conforme a (3.4). Esta área sirve como base, para identificar si la demanda del día actual, sigue algún perfil de demanda ya conocido, o de ser el caso, la demanda se está comportando de manera atípica. La Fig. 5 muestra el pronóstico de la demanda en tiempo real. La línea en color rojo es la demanda en tiempo real muestreada cada 15 minutos. La sombra en color verde es el área de demanda esperada. Es decir, se espera que el comportamiento de la demanda real esté dentro de los límites del área. Adicionalmente, en la parte superior del gráfico se ha añadido la desviación estándar del perfil de demanda esperada. Así por ejemplo, se observa que la demanda a las 18:00 podría tener una variación de 30 MW aprox. alrededor de la media. Con este último artefacto visual, el operador tiene un criterio para determinar las horas de máxima variación de la demanda. Finalmente con el objetivo de hacer más robusto al algoritmo, el perfil de demanda esperado es la unión de los tres perfiles más cercanos a la demanda real, con el criterio de máximos y mínimos. Así cada perfil aporta a la creación de valles e inflexiones.

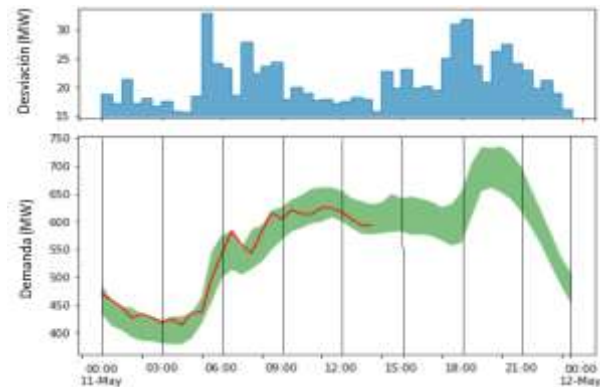


Figura 5: Pronóstico de la demanda en tiempo real.

3. DESARROLLO

3.1. Software utilizado

La exploración de datos ha sido desarrollada sobre Jupyter Python con las librerías asociadas a HMM (hmm_learn¹). Además para el entrenamiento de los modelos, se ha desarrollado una aplicación en Python que sea capaz de realizar el entrenamiento en paralelo, explotando de esta manera la tecnología disponible.

3.2. Arquitectura del software

Al inicio de esta arquitectura se encuentra una base de datos ya validada la cual permite realizar el proceso de aprendizaje/entrenamiento al modelo.

Posteriormente, con la información de tiempo real, obtenida de PI-Server, se evaluó al modelo obtenido en el proceso de entrenamiento. Finalmente los resultados son mostrados en una aplicación en el computador como se ilustra a continuación:

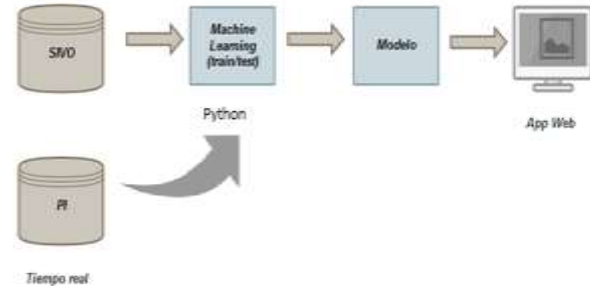


Figura 6: Arquitectura del software

RESULTADOS Y EVALUACIÓN

3.3. Perfiles de demanda

El descubrimiento de perfiles efectuado desde 01-01-2014 hasta 31-12-2017 proporciona un total de 57 perfiles de demanda. La Fig. 2 muestra un ejemplo de ellos. Se obtuvieron 40 perfiles que son categóricos y 17 que son vacilantes. En los perfiles categóricos más del 90% de sus muestras pertenecen a un día en particular. Así por ejemplo, el perfil ID=39 agrupa estrictamente 43 días lunes de los años 2014 a 2017. Un ejemplo de perfil vacilante es el perfil ID=28 que contiene 19 días martes, 23 días miércoles y 9 días jueves, lo que demuestra que los días martes, miércoles y jueves tienen un comportamiento similar que era conocimiento empírico que ya tenía el operador. Adicionalmente, a través del HAC se puede detectar aquellos perfiles que tienen un comportamiento singular, por ejemplo, el perfil ID=54 (véase en Fig. 2 tiene una distancia euclidiana grande referente a otros perfiles. Este perfil agrupa 9 días feriados, entre ellos Navidad, fin de año y 2 de noviembre.

3.4. Familias de demanda

El umbral de distancia igual a 250 permitió encontrar 6 familias, de las cuales 3 son categóricas y 3 son vacilantes. Por ejemplo, como se muestra en la Fig. 7, la familia 4 es vacilante ya que contiene 32% de perfiles típicos del día martes, 35% del día miércoles, 28% del día jueves y 5% para otros días. De la misma manera se verificó que una familia puede contener los perfiles correspondientes a días feriados y otra familia contener los perfiles asociados a días de alta demanda.

¹ <https://hmmlearn.readthedocs.io/en/latest/>

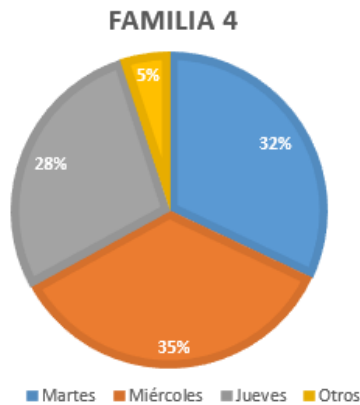


Figura 7: Distribución de la familia 4 de perfiles de demanda

3.5. Pronóstico de demanda en tiempo real

En una primera evaluación del pronóstico en tiempo real, se observó que las bandas del área de demanda esperada, no superan los 35 MW alrededor de la media, esto implica una variación del 9% de la demanda total de la EEQ (761 MW). En un tiempo de evaluación de 30 días se observó que la herramienta acierta el 86% de los casos y el valor de demanda en tiempo real se encuentra dentro de la banda de demanda esperada.

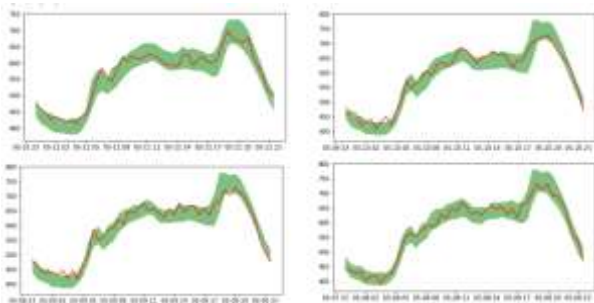


Figura 8: Evaluación de la metodología - ejemplo de días en cumplimiento

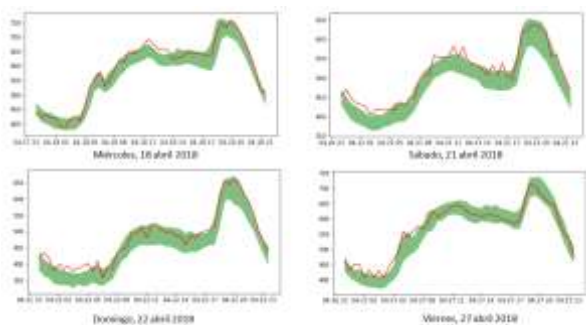


Figura 9: Evaluación de la metodología - ejemplo de días en incumplimiento

Esta evaluación sigue el criterio de control estadístico de calidad, donde se considera como día incumplido, aquel día donde se registre 2 horas (4 valores consecutivos) de desvío, fuera del área de predicción, criterio que se utiliza en el centro de control

del CENACE para realizar un redespacho por desvío de demanda. Se observó que para el mes de evaluación (abril 2018), el 86% (26 días) fueron acertados por la metodología, como se ilustra en la Fig. 8. El 14% restante presenta un incumplimiento como se muestra en la Fig. 9.

4. CONCLUSIONES Y RECOMENDACIONES

La evaluación del pronóstico en tiempo real de la demanda, conforme a los resultados obtenidos en un tiempo de evaluación de 30 días es aceptable. En una futura investigación se deberá evaluar con mayor rigurosidad la herramienta. Como plan piloto se ha comenzado con la demanda de la EEQ, sin embargo dado los resultados obtenidos, se desea ampliar su uso a otros puntos de entrega en el Sistema Interconectado.

El descubrimiento de perfiles de demanda está permitiendo la ganancia de conocimiento. Si bien es cierto, anteriormente se conocía de manera empírica que la demanda de los días lunes tenía un comportamiento característico, gracias a esta investigación se confirma dicha afirmación. Además se pudo verificar que existen perfiles de demanda que son interanuales, y algunos son muy característicos en ciertos meses, lo cual puede estar ligado a temas de clima, o periodos altos de producción.

De los resultados obtenidos, se observa que la metodología tiene aún posibilidades de mejora. Por ejemplo se puede determinar una banda autoajustable cambiando las fórmulas (3,4) donde el valor de α oscile entre 1 a 3 desviaciones estándar. Con esto se consigue ampliar el área de demanda esperada de manera que se disminuya incumplimientos.

REFERENCIAS BIBLIOGRÁFICAS

- [1] D. K. D. Haussler and M. G.R.F. H. Eeckman, “A generalized hidden markov model for the recognition of human genes in dna”, in Proc. Int. Conf. on Intelligent Systems for Molecular Biology, St. Louis, 1996, pp. 134–142.
- [2] D. K. Seymore, A. McCallum, and R. Rosenfeld, “Learning hidden markov model structure for information extraction”, in AAAI-99 workshop on machine learning for information extraction, 1999, pp. 37–42.
- [3] J. A. Bilmes et al., “A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models”, International Computer Science Institute, vol. 4, no. 510, p. 126, 1998.
- [4] Z. Ghahramani, “An introduction to hidden markov models and bayesian networks”, International journal of pattern recognition and artificial

- intelligence, vol. 15, no. 01, pp. 9–42, 2001.
- [5] M. Stamp, “A revealing introduction to hidden markov models”, Department of Computer Science San Jose State University, 2004.
- [6] R. G. Sánchez, “A proposed method for unsupervised anomaly detection for a multivariate building dataset”, M.S. thesis, Dep. of informatics, Fribourg University., Fribourg, Suiza, 2017.
- [7] G. Pfundstein, “Hidden markov models with generalised emission distribution for the analysis of high-dimensional, non-euclidean data”, PhD dissertation, Institut für Statistik, Alemania, 2011.
- [8] D. F. Williamson, R. A. Parker, and J. S. Kendrick, “The box plot: A simple visual method to interpret data”, *Annals of internal medicine*, vol. 110, no. 11, pp. 916–921, 1989.
- [9] S. Saraçlı, N. Dogan, and İsmet, “Comparison of hierarchical cluster analysis methods by cophenetic correlation”, *Journal of Inequalities and Applications*, vol. 2013, no. 1, p. 203, 2013.



Roberto Sánchez Albán.- Nació en Quito en 1986. Recibió su título de Ingeniero Electrónico de la Universidad de las Fuerzas Armadas (ESPE) en 2010; de Master en Computer Science de la Universidad de Fribourg - Suiza en 2017. Se

ejerció como ingeniero SCADA en el área de tiempo real en el Operador Nacional de Electricidad CENACE entre 2009 a 2014. Sus campos de investigación están relacionados con la aplicación de Machine Learning, Artificial Intelligence y Data visualization en la industria, actualmente aplicado al sector eléctrico.



Patricio Barrera Navas.- Nació en Ambato en 1990. Recibió su título de Ingeniero Electrónico de la Escuela Politécnica Nacional en 2016. Actualmente trabaja en el Operador Nacional de Electricidad, CENACE. Sus áreas de interés radican en el análisis post operativo de sistemas de potencia.