

Tipo de artículo: Artículo original  
Temática: Tecnologías de la información y las telecomunicaciones  
Recibido: 04/03/19 | Aceptado: 09/07/19 | Publicado: 26/09/19

## **Sistema para la gestión y análisis de datos de una red de sensores inalámbricos basado en un almacén de datos.**

### *System for the management and data analysis of a wireless sensor network based on a data warehouse*

Alejandro Cantero Díaz<sup>1\*</sup>, María Margarita Goire Castilla<sup>2</sup>, Yasser Quintana Cassulo<sup>3</sup>

<sup>1</sup> Universidad de Oriente, Cuba, [acantero@uo.edu.cu](mailto:acantero@uo.edu.cu)

<sup>2</sup> Universidad de Oriente, Cuba, [mgoire@uo.edu.cu](mailto:mgoire@uo.edu.cu)

<sup>3</sup> Universidad de Oriente, Cuba, [yasser.quintana@uo.edu.cu](mailto:yasser.quintana@uo.edu.cu)

Alejandro Cantero Díaz: [acantero@uo.edu.cu](mailto:acantero@uo.edu.cu); [alejandroc@elecstg.une.cu](mailto:alejandroc@elecstg.une.cu)

---

#### **Resumen**

Usar de forma efectiva toda la información generada por una red de sensores inalámbricos es todo un reto. El gran volumen de datos que generan los sensores presenta gran variedad y su generación se produce con mucha celeridad. Esto hace difícil la interpretación de estos datos por parte de los usuarios, volviéndose poco útiles en el apoyo a la toma de decisiones, si no están procesados adecuadamente. En este trabajo se desarrolló un sistema de gestión y análisis de los datos generados por redes de sensores inalámbricos, basado en un almacén de datos. Este sistema permitió la estandarización de los datos y su almacenamiento en una estructura multidimensional. Posibilitó la supervisión remota a través de la web, del entorno donde se encuentra desplegada la red de sensores inalámbricos. También permitió analizar los datos desde diferentes escenarios históricos, y evaluar su comportamiento y evolución en un ambiente multidimensional, o sea, mediante la combinación de diferentes perspectivas o dimensiones. El análisis de las series temporales construidas con los datos almacenados en el almacén de datos, permitió realizar predicciones, a partir del comportamiento histórico de los datos. Proporcionando elementos que apoyan la toma de decisiones, a partir de lo que puede ocurrir en el futuro. El sistema fue desarrollado usando Hefestos como metodología de desarrollo y herramientas como Pentaho Data Integration para la carga de los datos en el almacén de datos y Pentaho BI Server para el análisis de estos mediante un navegador web.

**Palabras clave:** redes de sensores inalámbricos, gran volumen de datos, almacén de datos, gestión y análisis de datos.

### **Abstract**

*Effectively using all the information generated by a wireless sensor network is a challenge. The large volume of data generated by the sensors presents great variety and its generation occurs very quickly. This makes it difficult for users to interpret these data, making them little useful in supporting decision-making, if they are not properly processed. In this work, a system of management and analysis of the data generated by wireless sensor networks, based on a data warehouse was developed. This system allowed the standardization of the data and its storage in a multidimensional structure. It enabled remote supervision through the web, of the environment where the wireless sensor network it is. It also allowed to analyze the data from different historical scenarios, and to evaluate its behavior and evolution in a multidimensional environment, that is, by combining different perspectives or dimensions. The analysis of the time series constructed with the data stored in the data warehouse allowed predictions to be made, based on the historical behavior of the data. Providing elements that support decision making, based on what may happen in the future. The system was developed using Hephæstus as a development methodology and tools such as Pentaho Data Integration for loading the data in the data warehouse and Pentaho BI Server for the analysis of these through a web browser.*

**Keywords:** *wireless sensor networks, large volume of data, data warehouse, data management and analysis*

---

## **Introducción**

Una de las tecnologías que experimenta un gran desarrollo en los últimos años, cuya aplicación implica la generación de grandes volúmenes de datos, es la denominada red de sensores inalámbricos o WSN (por su acrónimo en inglés Wireless Sensor Networks). Una WSN constituye un tipo de red ad-hoc autónoma, integrada por decenas o cientos de nodos inteligentes (motes); caracterizados por tener poca capacidad de procesamiento y una fuente de alimentación típica basada en baterías. Los nodos sensores que componen la red, son pequeños sistemas informáticos embebidos, encargados de recoger datos del entorno como temperatura y humedad, entre otros, teniendo la posibilidad de comunicarse de forma inalámbrica (Figueredo, 2016).

En las WSNs, los datos recolectados y procesados por los nodos sensores son transmitidos hacia un nodo central conocido como gateway (Yinbiao, 2014), (Iacono, Godoy, 2012). El gateway se encarga de recopilar la información de todos los nodos de la red y a su vez, puede estar conectado a una computadora o a un dispositivo capaz de almacenar gran cantidad de información. Este actúa como elemento de interconexión entre la WSN y una red TCP/IP para monitorizar los datos recolectados por los nodos sensores (Figueredo, 2016), (Yinbiao, 2014).

La reducción en tamaño y costo de los sensores, la diversidad de la gama de tipos de sensores disponibles y la evolución de los medios de comunicación inalámbricos, posibilitan la ampliación del alcance de las WSNs (Farrah, El Manssouri, Ouzzif, 2015), las cuáles se presentan como la solución para aplicaciones como el monitoreo de

fenómenos meteorológicos, las erupciones volcánicas o los terremotos, la vigilancia, la seguridad y la supervisión de entornos industriales. (*Mayhua, Ludeña, 2016*), (*Nellore, Hancke, 2016*), (*Sharma, 2014*), (*Aponte, Gómez, Gómez, Sánchez, Alcina, Teixido, 2018*).

Miles de kilómetros cuadrados pueden ser supervisados a través de los años mediante una WSN, convirtiéndose en una herramienta de mucha utilidad en diversas esferas de la vida del ser humano. Por ejemplo, en la agricultura, son fácilmente desplegables pudiendo cubrir áreas con diferentes superficies. Los datos generados por estas redes suelen ser finalmente usados para supervisar el consumo de agua o las condiciones meteorológicas locales (*Yu, Wu, Han, Zhang, 2013*), (*Quiroga, Jaramillo, Campo, Chanchi, 2016*).

La utilidad que tiene una WSN en sus múltiples aplicaciones, pudiera verse limitada sin el adecuado tratamiento al gran volumen de datos colectados, impidiendo la utilización de los mismos en la toma de decisiones estratégicas y tácticas. La mayoría de los sistemas de administración de bases de datos relacionales, ya no son capaces de administrar y procesar adecuadamente el volumen de datos generado por una WSN (*Yinbiao, 2014*), (*Hernández, Duque, Cadavid, 2017*).

Una de las tecnologías que está concebida para gestión y análisis de grandes volúmenes de datos, es la denominada Almacenes de Datos o DW (del inglés Data Warehouse). Los sistemas basados en DWs son considerados una de las soluciones clásicas para este tipo de problemas, relacionado con la gestión y análisis de grandes volúmenes de datos. Los sistemas basados en DW son repositorios diseñados para facilitar la confección de informes y la realización de análisis de datos con mayor rapidez. Pueden ser completamente separados del sistema de información principal; lo cual significa una ganancia enorme en el rendimiento de los sistemas, cuando se ejecuten las consultas.

Este es un tema actual y que adquiere mucha importancia, fundamentalmente cuando se manejan grandes volúmenes de datos. Los cuales necesitan ser analizados y procesados para cumplir un objetivo y con ello tomar decisiones oportunas.

Un sistema basado en DW posibilita la extracción de datos de sistemas operacionales y fuentes externas, permite la integración y homogeneización de los datos, provee información transformada y estandarizada, para que ayude en el proceso de toma de decisiones estratégicas y tácticas. Convierte los datos operacionales en una herramienta competitiva, debido a que pone a disposición de los usuarios indicados la información pertinente, correcta e integrada, en el momento que se necesita.

Un ejemplo de aplicaciones basadas en DW para WSN se presenta en el artículo titulado “*An Approach to Analyze Large Scale Wireless Sensors Network Data*” de los autores Soufiane Farrah, Hanane El Manssouri y Mohamed Ouzzif, el cual propone el desarrollo de una herramienta para el análisis de datos captados y el comportamiento de los sensores en la WSN (Farrah, El Manssouri, Ouzzif, 2015). Otro ejemplo lo constituye el artículo titulado “*Use of data warehouse to manage data from wireless sensors networks that monitor pollinators*” de los autores Ricardo Augusto Gomes da Costa y Carlos Eduardo Cugnasca, donde se muestra el uso de un sistema basado DW, como una alternativa para administrar datos recolectados por una WSN que supervisa abejas (Gomes, Cugnasca, 2010).

Considerando lo planteado anteriormente, es posible determinar que el gran volumen de datos recopilado por la WSN, posibilita supervisar características del entorno donde sea desplegada, pero estos datos solos o simplemente recopilados, son interpretados por los usuarios con mucha dificultad y no ayudan a tomar una decisión. Para usar esta información de una forma productiva, debe organizarse en un repositorio o base de datos, y tener una interfaz de fácil acceso, a través de la cual el usuario pueda ver información consolidada sobre diferentes comportamientos y ser capaz de tomar decisiones estratégicas (Farrah, El Manssouri, Ouzzif, 2015).

A partir de la situación problemática antes expresada, se plantea el siguiente problema de investigación: ¿Cómo gestionar y analizar grandes volúmenes de datos generados por una WSN? En el contexto investigativo se define como objetivo general desarrollar un sistema para la gestión y análisis de grandes volúmenes de datos generados por una WSN.

## **Materiales y métodos**

### **Almacenes de datos**

El almacenamiento de datos o DWH (por su acrónimo en inglés Data Warehousing) es el proceso que posibilita la extracción de datos de sistemas operacionales y fuentes externas, permite la integración y homogeneización de todos los datos, provee información que ha sido transformada y estandarizada, para que ayude en el proceso de toma de decisiones estratégicas y tácticas. Convierte los datos operacionales en una herramienta competitiva, debido a que pondrá a disposición de los usuarios indicados la información pertinente, correcta e integrada, en el momento que se necesita. Este almacenamiento simple, completo y consistente de datos obtenidos desde una variedad de fuentes, disponibles para el usuario final de forma tal que puedan utilizarlos (Ricardo, 2010).

Para que el DWH pueda cumplir con sus objetivos, es necesario que la información que se extrae, transforma y consolida, sea almacenada de manera centralizada en una base de datos con estructura multidimensional denominada DW (*Ricardo, 2010*). Entre las principales características de un DW se puede destacar que: es orientado al negocio, integrado, variable en el tiempo y no volátil.

- Orientado al Negocio – La información se clasifica en bases a los aspectos que son de interés.
- Integrado – Implica que todos los datos de diversas fuentes, deben ser consolidados en una instancia antes de ser agregados al almacén de datos.
- Variable en el Tiempo – Todos los datos almacenados en un DW son identificados por un instante de tiempo específico (*Gomes, Cugnasca, 2010*).
- No Volátil – La información es útil para el análisis y la toma de decisiones solo cuando es estable. Los datos una vez que entran en el DW no cambian (*Ricardo, 2010*).

Con la descripción de las características del almacén de datos se puede detallar a continuación la arquitectura del DWH.

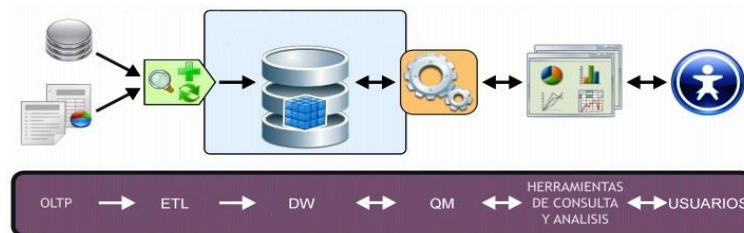


Figura 1. Esquema del proceso de almacenamiento de datos (*Ricardo, 2010*).

Los elementos de este esquema son los siguientes:

- *OLTP (On Line Transaction Processing)*: representa toda aquella información transaccional que se genera diariamente, además, de las fuentes externas con las que puede llegarse a disponer.
- *ETL*: los ETL extraen datos de las diversas fuentes que se requiera, los transforma para resolver posibles problemas de inconsistencias entre los mismos y finalmente, después de haberlos depurado se procede a su carga en el DW. Estos procesos permiten manipular, integrar y transformar los datos, para posteriormente cargar los resultados obtenidos en el DW.
- *DW*: Es una base de datos con estructura multidimensional, constituye una colección de datos orientada al negocio, integrada, no volátil y variante en el tiempo.

- *Query Manager*: Permite realizar la gestión y ejecución de consultas relacionales, tales como join y agregaciones, y de consultas propias del DW para el análisis de los datos.
- Herramientas de consulta y análisis: son sistemas que permiten al usuario realizar la exploración de datos del DW. Básicamente constituyen el nexo entre el depósito de datos y los usuarios.
- El procesamiento analítico en línea OLAP (*On Line Analytic Processing*): es la componente más poderosa de los sistemas basados en DW, siendo el motor de consultas especializado del depósito de datos. Su principal objetivo es el de brindar respuestas rápidas a preguntas complejas, para interpretar la situación existente y tomar decisiones.
- Usuarios: son aquellos que se encargan de tomar decisiones y de planificar las acciones a realizar, es por ello que se hace tanto énfasis en la integración y limpieza de datos, para poder conseguir que la información posea toda la calidad posible.

### **Modelo y cubo multidimensional**

El modelo multidimensional es una técnica de diseño lógico que intenta presentar los datos en un estándar que se rige por una disciplina que aplica el modelo entidad-relación, pero con algunas restricciones importantes. Cada modelo está compuesto por una tabla con múltiples llaves llamada tabla de hechos y un grupo de pequeñas tablas llamadas dimensiones. Cada tabla dimensión tiene una llave primaria que corresponde exactamente con una de las múltiples llaves en la tabla de hechos (*Ricardo, 2010*).

El modelo multidimensional cuenta con varias estructuras para su modelado entre ellas el Esquema Estrella (Star Schema), el Copo de Nieve (Snowflake Schema) y el Esquema Constelación (Constellation Scheme). Esquema estrella: Consta de una tabla de hechos central y de varias tablas de dimensiones relacionadas a esta, a través de sus respectivas claves. Esquema copo de nieve: Representa una extensión del modelo en estrella cuando las tablas de dimensiones se organizan en jerarquías de dimensiones. Esquema constelación: Este modelo está compuesto por una serie de esquemas en estrella.

Para acceder a los datos del almacén, se pueden ejecutar consultas sobre algún cubo multidimensional previamente definido. Dicho cubo debe incluir los siguientes objetos: indicadores, atributos, jerarquías, entre otros, basados en los campos de las tablas de dimensiones y de hechos, que se deseen analizar (*Ricardo, 2010*).

## **Hefesto como metodología de desarrollo**

Hefesto es una metodología cuya propuesta está fundamentada en una amplia investigación, comparación de metodologías existentes y experiencias en procesos de confección de DWs. Esta metodología permite la construcción de un DW de forma sencilla, ordenada e intuitiva (*Ricardo, 2010*).

Los pasos para la aplicación de esta metodología son los siguientes:

1. Análisis de requerimientos
2. Análisis de los OLTP
3. Modelo lógico del Almacén de Datos
4. Procesos ETL

## **Procesamiento analítico en línea: OLAP**

Lo realmente interesante en OLAP, no es la ejecución de simples consultas tradicionales, sino la posibilidad de utilizar operadores tales como *drill-up* (de lo específico a lo general), *drill-down* (de lo general a lo específico), para explorar profundamente la información (*Ricardo, 2010*), (*Gomes, Cugnasca, 2010*).

Mediante la exploración profunda de la información es posible construir series de tiempo, las que permiten caracterizar el comportamiento de un indicador en un periodo de tiempo determinado, utilizando los intervalos de tiempo que defina el usuario, posibilitando la realización de predicciones sobre el comportamiento futuro de este indicador.

## **Series Temporales.**

Una Serie Temporal (ST) es una sucesión de observaciones de una variable en distintos instantes de tiempo. Dichos instantes de tiempo, suelen ser de tipo regular (cada día, cada mes, cada año, entre otros). El análisis de STs posibilita describir el comportamiento de los datos, pudiendo hacer distinción de los diferentes componentes que lo provocan. También es posible predecir valores futuros de la variable a partir de información disponible de la serie, a través de métodos de predicción, como el suavizado exponencial. (*García, 2014*)

## **Suavizado exponencial**

Los métodos de suavizado tienen como propósito obtener valores ponderados de una ST y después extrapolarlos para obtener un pronóstico. Los métodos de suavizado exponencial se caracterizan por ponderar con mayor peso a las

observaciones más recientes y con menor a las más antiguas. Dentro de estos métodos existen varios tipos, conocidos en la literatura como suavizado exponencial de uno, dos y tres parámetros. (Sánchez, Poveda, 2006).

- El suavizado exponencial simple (un parámetro): Es la forma general de los suavizados exponenciales. Es útil sólo para aquellas series que tengan un patrón horizontal, por lo que ignora las tendencias y estacionalidades.
- La doble suavización exponencial de tendencia o de Holt (dos parámetros): Contempla las tendencias en las series, pero no las estacionalidades.
- El suavizado exponencial tendencia y estacionalidad o de Holt – Winters (tres parámetros): Contempla la tendencia y la estacionalidad de las series.

### Arquitectura del sistema de gestión y análisis de datos generados por una red de sensores inalámbricos basado en un Almacén de Datos

La arquitectura para el sistema de gestión y análisis de los datos recolectados por una WSN, está basada en un DW. Esta propuesta se sustenta en la necesidad de modificar, optimizar y generar métodos y modelos de almacenamiento y tratamiento de grandes volúmenes de datos generados por una WSN.

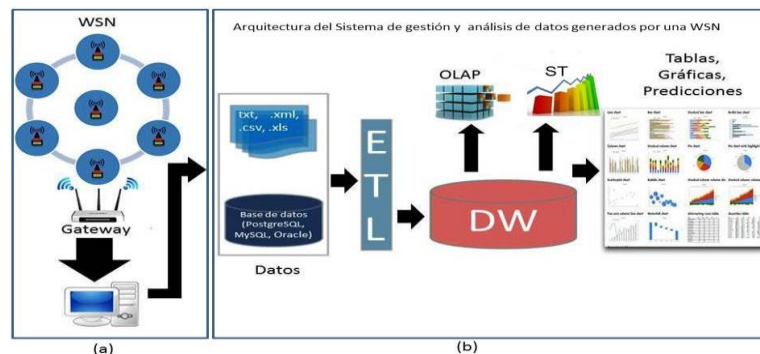


Figura 2. Esquema general de la solución propuesta para la gestión y análisis de los datos generados por WSN basada en un DW.

La figura 2 (b) ilustra la arquitectura, que comienza a partir del almacenamiento de los datos generados por la WSN, en una computadora, mediante la estructura determinada (base de datos o ficheros). Esta información que se genera diariamente por parte de una WSN, se considera la fuente de datos, que es el elemento inicial en esta arquitectura. Esta fuente de datos se denomina en la literatura OLTP.

El siguiente elemento que presenta la arquitectura, es el conjunto de procesos ETL. Lo que constituye la actividad más importante, comenzando con la extracción de datos desde el OLTP. Luego de extraer los datos, se realiza el



proceso de transformación, que consiste en la normalización de los datos, con el objetivo de resolver problemas de inconsistencia en los mismos. Los problemas de este tipo generalmente consisten, en que los datos extraídos están expresados en diferentes unidades de medida. Por ejemplo, la temperatura pudiera estar expresada en ° C y ° F. Por lo tanto, es necesaria una conversión a la unidad apropiada. Generalmente un DW contiene datos de un periodo extenso de tiempo, una granularidad pequeña, como la hora o el minuto, suele ser descartada (*Gomes, Cugnasca, 2010*)

Después de extraer y transformar los datos, se procede a cargar esta información en el DW, facilitando la aplicación sobre estos datos de herramientas y técnicas OLAP (*Ricardo, 2010*). La utilización de estas herramientas posibilita interpretar la situación existente y tomar decisiones. Las herramientas OLAP, aportan al usuario la capacidad de explorar profundamente y de forma remota, los datos generados por una WSN, permitiéndoles conocer el comportamiento de una variable determinada, captada por un sensor determinado, en el transcurso de un año, o de varios años, posibilitando identificar los meses o los días del año, en que esa variable tiene un comportamiento diferente, compararlo con el comportamiento de otra variable e identificar la relación existente entre ambas. Con estas herramientas, el usuario puede realizar análisis cualitativos, mediante la presentación de gráficos de distintos tipos, que ayuden a entender mejor el comportamiento de las variables del entorno donde esté desplegada la WSN. También son capaces, a partir de las consultas al DW, de construir STs de los datos almacenados.

Estas STs pueden ser construidas según los intereses del usuario, permiten caracterizar el comportamiento de un indicador en un periodo de tiempo determinado, posibilitando la realización de predicciones sobre el comportamiento futuro de este indicador, a través de métodos de predicción, como el suavizado exponencial. La aplicación de los métodos de suavizado exponencial a las STs construidas con los datos almacenados en el DW, posibilita pronosticar el comportamiento futuro de estos datos. Proporcionando elementos que apoyan la toma de decisiones, en función de lo que puede ocurrir en el futuro.

La materialización de esta arquitectura es posible mediante la construcción de un DW, en función de la gestión y el análisis de los datos generados por una WSN.

### **Modelo lógico del Almacén de Datos**

El modelo lógico del DW se observa en la figura 3. Este se basa en un modelo de esquema estrella. Se define una tabla de hechos llamada H\_Sensado que contiene todas las mediciones realizadas a través de la WSN, y claves externas de las tablas de dimensiones D\_Fecha y D\_Sensor. La tabla dimensión D\_Fecha contiene datos del instante de tiempo en el que fue realizada la medición, y D\_Sensor presenta los datos del sensor que realiza esta medición.

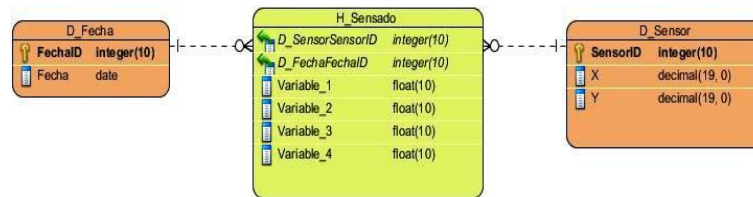


Figura 3. Modelo lógico para el almacenamiento de los datos en el DW.

## Herramientas para el desarrollo del Almacén de Datos

Toda la información generada por una WSN es almacenada en un DW, para ser analizada y procesada de forma remota, a través de la web, apoyando el proceso de tomar decisiones según comparaciones, cálculos estadísticos, gráficos y predicciones. Para lograr este objetivo es necesaria la utilización de la herramienta ETL Pentaho Data Integration 5.3 (PDI), el servidor de bases de datos PostgreSQL 9.1, servidor web Apache Tomcat embebido en la herramienta Pentaho BI Server 4.8, el visor OLAP STPivot y la herramienta Pronóstico del Microsoft Excel 2016.

## Datos para validar la arquitectura propuesta

Para la implementación y validación del sistema de gestión y análisis de los datos colectados por una WSN, se considera factible la utilización los datos ambientales de 10 estaciones meteorológicas ubicadas en el noroeste de México, al servicio del Sistema Meteorológico Nacional de este país. Los datos extraídos de estas estaciones, consisten en valores diarios de precipitación y temperaturas desde el año 1980 hasta el año 2000. Estos están públicos en Internet y son accedidos mediante la url: <http://peac-bc.cicese.mx/datosclim/dcbc.php>.

Cada estación meteorológica proporciona un fichero de texto que contiene las mediciones diarias de temperatura mínima, temperatura máxima, temperatura promedio y precipitaciones. Es importante destacar el alto grado de similitud que presentan estos datos con los emitidos por una WSN real.

## Resultados y discusión

Se obtuvo un sistema que es accesible mediante un navegador web, lo que posibilita la supervisión remota de los datos a través de internet o una intranet, según convenga.

Mediante las funcionalidades que posee *StPivot*, los usuarios tienen la posibilidad de conocer el comportamiento de la temperatura o las precipitaciones, en una región determinada, en el transcurso de un año, o de varios años.

Identificando los meses o los días del año, en que la temperatura o las precipitaciones tienen un comportamiento diferente.

El usuario tiene también la posibilidad de comparar el comportamiento un indicador respecto al otro e identificar la relación existente entre ambos. También pueden realizar análisis cualitativos, mediante la presentación de gráficos de distintos tipos, que ayuden a entender mejor el comportamiento de las variables climáticas, en las regiones donde están desplegados los nodos sensores.

La figura 4 ilustra el comportamiento de las precipitaciones y la temperatura máxima registrada por el sensor 2001 desde el año 1990 hasta el año 2000, pudiendo evidenciarse las diferencias en los valores de precipitaciones entre los años, sin embargo, la temperatura máxima promedio presenta valores similares en estos 10 años. Entre las facilidades proporcionadas por *StPivot* para realizar este tipo de análisis, se encuentra la posibilidad de mostrar gráficos de distintos modelos. De esta forma el usuario puede comprender mejor la información, visualizada desde distintas perspectivas.

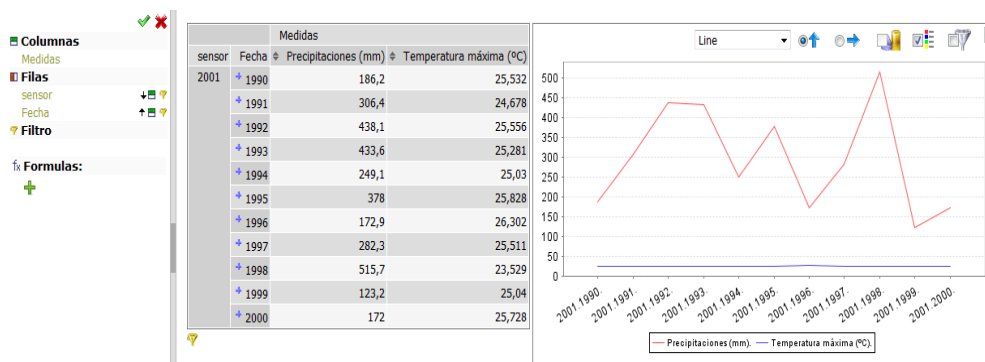


Figura 4. Análisis de los valores de precipitaciones y temperatura máxima registrados por el sensor 2001 desde 1990 hasta 2000.

Otra de las funcionalidades de *StPivot* consiste, en que a partir de las consultas realizadas al DW, permite construir series temporales. Mediante las cuales es posible conocer el comportamiento de un indicador en un periodo de tiempo determinado, utilizando los intervalos de tiempo que defina el usuario y exportar en formato XLS esta información, como se observa en la figura 5. Donde se construye una serie tiempo de 19 años con intervalos mensuales, que describe el comportamiento de la temperatura promedio registrada por el sensor 2003.

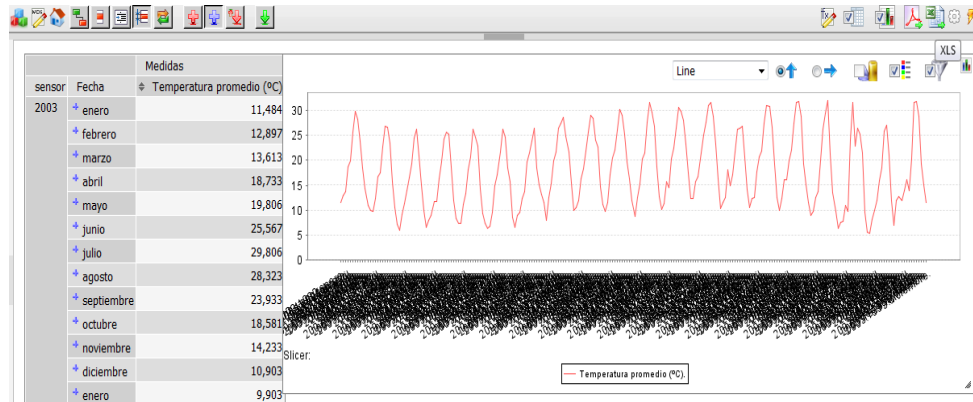


Figura 5. Serie temporal del comportamiento de la temperatura promedio.

Los resultados de la predicción realizada por la herramienta Pronóstico de Microsoft Excel 2016 a partir de la serie temporal construida por *StPivot* y su comparación con los datos reales de temperatura promedio mensual, registrados por el sensor 2003 en el año 2000, pueden observarse en la tabla 1.

Tabla 1. Resultados de la predicción realizada.

Fecha	Temperatura Promedio real (TP) (°C)	Predicción TP (°C)	Límite inferior de la Predicción TP (°C)	Límite superior de la Predicción TP (°C)
ene-00	9	11	7	16
eb-00	10	13	7	19
mar-00	12	15	8	23
abr-00	19	18	9	26
may-00	25	21	12	30
jun-00	30	26	16	35
jul-00	32	29	18	40
ago-00	33	30	18	41
sep-00	30	26	14	38
oct-00	22	19	7	32
nov-00	18	14	1	27
dic-00	13	12	-2	25

El 100% de los valores reales de temperatura promedio mensual, se encuentra entre los valores concebidos el intervalo de confianza establecido en el pronóstico realizado. Se puede observar además la similitud existente entre los valores reales y los valores pronosticados.

La figura 6 contiene una gráfica que ilustra el resultado de la predicción realizada. Destacándose la similitud existente entre la línea azul y la naranja, correspondientes al comportamiento del indicador temperatura promedio y a la predicción realizada para este indicador respectivamente. Las líneas amarilla y gris corresponden a los límites de

confianza de la predicción realizada en este caso, observándose que todos los valores reales de las mediciones de este indicador están comprendidos dentro de este intervalo de confianza.

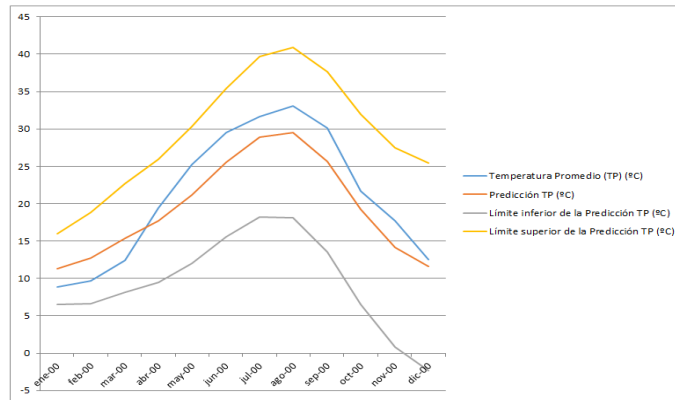


Figura 6. Resultados de la predicción

En la figura anterior se puede comprobar la efectividad de esta herramienta en la predicción de valores futuros de un indicador o variable determinada a partir de datos históricos que reflejen el comportamiento de estos indicadores.

Esto constituye un aporte novedoso en el análisis de los datos generados por una WSN. La posibilidad de obtener una predicción con un grado adecuado de similitud, a lo que realmente puede ocurrir en el futuro, es un elemento de suma importancia en el proceso de toma de decisiones. Permitiendo a los usuarios conocer anticipadamente el posible comportamiento de una variable climática, como es el caso del ejemplo anterior.

## Conclusiones

El gran volumen de datos que genera una WSN, debe ser procesado adecuadamente, para que sea de utilidad a los usuarios en la toma de decisiones. Un sistema basado en un DW es una alternativa adecuada para la gestión y análisis de los datos generados por una WSN. Considerando la posibilidad que tienen estos sistemas de aportar varios elementos que facilitan al usuario la interpretación de los datos y la toma de decisiones.

La metodología Hefestos permite guiar el proceso que garantiza la construcción de un DW para el almacenamiento adecuado de los datos generados por una WSN. Las herramientas de la suite Pentaho y la herramienta Pronóstico del Microsoft Excel 2016 posibilitan la implementación de un sistema basado en un DW para la gestión y análisis de los datos generados por una WSN. Este sistema permite la estandarización de grandes volúmenes de datos y su almacenamiento en una estructura multidimensional. Posibilitando analizar estos datos desde diferentes escenarios

históricos, y evaluar su comportamiento y evolución en un ambiente multidimensional. Es capaz también de realizar predicciones a partir de los datos históricos almacenados en el DW. Proporcionando elementos importantes que ayudan en el proceso de toma de decisiones en base a lo que puede pasar en el futuro.

## Referencias

FIGUEREDO, M. Propuesta de diseño de red de sensores inalámbricos para la programación del riego en la caña de azúcar. Maestría en Telemática. Universidad Central de las Villas. Santa Clara. Cuba, 2016.

YINBIAO, S. Internet of Things: Wireless Sensor Networks. [En línea]. International Electrotechnical Commission. 2014, [Consultado el: 04 de febrero de 2018]. Disponible: <https://www.iec.ch/whitepaper/pdf/iecWP-internetofthings-LR-en.pdf>

IACONO, L.; GODOY, P. Estudio de la Integración entre WSN y redes TCP/IP. [En línea]. Memoria de Trabajos de Difusión Científica y Técnica, núm. 10. Julio 2012, [Consultado el: 05 de febrero de 2018] Disponible: [http://www.um.edu.uy/docs/5\\_estudio\\_de\\_la\\_integracion\\_entre\\_WSN\\_redes%20TCP\\_IP.pdf](http://www.um.edu.uy/docs/5_estudio_de_la_integracion_entre_WSN_redes%20TCP_IP.pdf)

FARRAH, S; EL MANSSOURI, H; OUZZIF, M. An Approach to Analyze Large Scale Wireless Sensors Network Data. [En línea]. IRJCS. Issue 5, Volume 2, May 2015. [Consultado el: 05 de febrero de 2018]. Disponible: <http://www.irjcs.com/volumes/vol2/iss6/02.JNCS10083.pdf>.

MAYHUA, E; LUDEÑA, J. Sistema de riego por goteo automático utilizando una red de sensores inalámbricos. [En Línea]. Arequipa. Volumen 7, 69-92, Año 2016. [Consultado el: 06 de febrero de 2018]. Disponible: <http://ucsp.edu.pe/investigacion/wp-content/uploads/2017/01/4.-Sistema-de-riego-por-goteo-automatico.pdf>.

NELLORE, K; HANCKE, G. A Survey on Urban Traffic Management System Using Wireless Sensor Networks. [En Línea]. sensor, Enero 2016. [Consultado el: 2 de marzo de 2018]. Disponible: <http://www.mdpi.com/1424-8220/16/2/157>.

SHARMA, P. Wireless Sensor Networks for Environmental Monitoring. [En Línea]. IJSRET, Noviembre 2014. [Consultado el: 2 de marzo de 2018]. Disponible: <https://pdfs.semanticscholar.org/f850/1a47830512c7056470de88bcad92a1ae9507.pdf>

APONTE, J; GÓMEZ, J; GÓMEZ, F; SÁNCHEZ, M; ALCINA, J; TEIXIDO, P. An Efficient Wireless Sensor Network for Industrial Monitoring and Control. [En Línea]. sensor, January 2018. [Consultado el: 2 de marzo de 2018]. Disponible: [https://www.researchgate.net/publication/322371843\\_An\\_Efficient\\_Wireless\\_Sensor\\_Network\\_for\\_Industrial\\_Monitoring\\_and\\_Control](https://www.researchgate.net/publication/322371843_An_Efficient_Wireless_Sensor_Network_for_Industrial_Monitoring_and_Control).

YU, X; WU, P; HAN, W; ZHANG, Z. A survey on wireless sensor network infrastructure for agriculture. [En Línea]. Computer Standards & Interfaces, pp. 59-64, enero 2013. [Consultado el: 06 de febrero de 2018]. Disponible: [https://www.researchgate.net/profile/Imed\\_Romdhani/post/I\\_am\\_interested\\_to\\_develop\\_a\\_system\\_using\\_sensors\\_for\\_agricultural\\_field\\_so\\_anyone\\_suggest\\_me\\_what\\_should\\_I\\_do/attachment/59d64346c49f478072eabe66/AS%3A273808659812357%401442292479264/download/WSN\\_Agriculture\\_Survey.pdf](https://www.researchgate.net/profile/Imed_Romdhani/post/I_am_interested_to_develop_a_system_using_sensors_for_agricultural_field_so_anyone_suggest_me_what_should_I_do/attachment/59d64346c49f478072eabe66/AS%3A273808659812357%401442292479264/download/WSN_Agriculture_Survey.pdf).

QUIROGA, E; JARAMILLO, S; CAMPO, W; CHANCHÍ, G. Propuesta de una Arquitectura para Agricultura de Precisión Soportada en IoT. [En Línea]. RISTI, 2016. [Consultado el: 06 de febrero de 2018]. Disponible: DOI: 10.17013/risti.24.39-56.

HERNÁNDEZ, E; DUQUE, N; CADAVID, J. Big Data: una exploración de investigaciones, tecnologías y casos de aplicación. [En Línea]. TecnoLógicas, vol. 20, no. 39, mayo - agosto, 2017. [Consultado el: 8 de marzo de 2018]. Disponible: <http://www.scielo.org.co/pdf/teclo/v20n39/v20n39a02.pdf>

GOMES, R; CUGNASCA, C. Use of data warehouse to manage data from wireless sensors networks that monitor pollinators. [En Línea]. Researchgate, January 2010. [Consultado el: 06 de febrero de 2018]. Disponible: [https://www.researchgate.net/publication/221422898\\_Use\\_of\\_Data\\_Warehouse\\_to\\_Manage\\_Data\\_from\\_Wireless\\_Sensors\\_Networks\\_That\\_Monitor\\_Pollinators](https://www.researchgate.net/publication/221422898_Use_of_Data_Warehouse_to_Manage_Data_from_Wireless_Sensors_Networks_That_Monitor_Pollinators)

RICARDO, B. Data Warehousing: Investigación y Sistematización de Conceptos – Hefesto: Metodología propia para la Construcción de un Data Warehouse. Córdoba, Argentina, Julio 2010. 145 pág.

GARCÍA, C. Modelización de la volatilidad mediante modelos autor regresivos de heterocedasticidad condicional. Maestría Oficial en Estadística Aplicada. Universidad de Granada, Granada, España, 2014.

SÁNCHEZ, J; POVEDA, G. Aplicación de los métodos Mars, Holt-Winters y ARIMA Generalizado en el pronóstico de caudales medios mensuales en ríos de Antioquia. [En Línea]. Meteorología Colombiana, No.10, pp 36-46, marzo 2006. [Consultado el: 10 de marzo de 2018]. Disponible: [http://ciencias.bogota.unal.edu.co/fileadmin/content/geociencias/revista\\_meteorologia\\_colombiana/numero10/10\\_04.pdf](http://ciencias.bogota.unal.edu.co/fileadmin/content/geociencias/revista_meteorologia_colombiana/numero10/10_04.pdf)