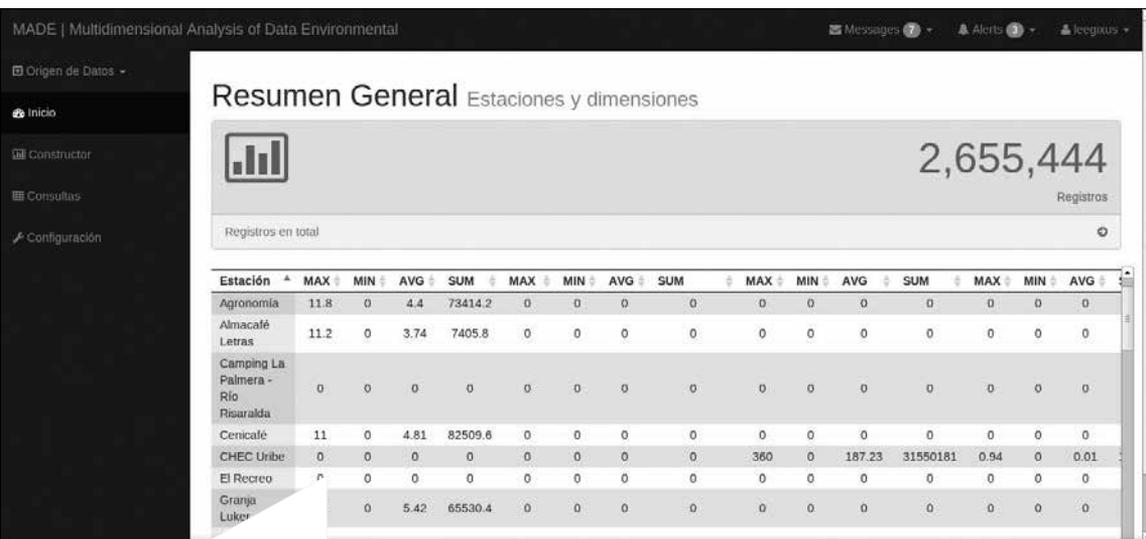


# boletín ambiental

Marzo de 2015

Instituto de Estudios Ambientales IDEA **124**

## Bodega de Datos Ambientales, Procesamiento y Análisis de Información Ambiental



# Bodega de Datos Ambientales, Procesamiento y Análisis de Información Ambiental

EMILCY JULIANA HERNÁNDEZ LEAL,  
estudiante de Maestría en Ingeniería Administrativa

NÉSTOR DARÍO DUQUE MÉNDEZ,  
doctor en Ingeniería

JORGE JULIÁN VELEZ UPEGUI,  
doctor en Planificación y Gestión de Recursos Hidráulicos

Trabajo conjunto Grupo de Trabajo Académico -GTA- Ingeniería Hidráulica y Ambiental,  
Grupo de Ambientes Inteligentes Adaptativos -GAIA- con el apoyo de la Corporación  
Autónoma Regional de Caldas CORPOCALDAS.  
Universidad Nacional de Colombia Sede Manizales

La recolección, almacenamiento, procesamiento y análisis de información de tipo ambiental es un requerimiento que día a día exigen la comunidad y las entidades ambientales locales, con el fin de contribuir con los procesos de planificación ambiental y como apoyo a los procesos de investigación que realizan en la región. Es por este motivo que el monitoreo de condiciones ambientales como las climáticas, meteorológicas e hidrometeorológicas constituyen una fuente de datos que puede aumentar de manera exponencial

La gran cantidad de datos recolectados por las redes de estaciones meteorológicas e hidrometeorológicas de la ciudad de Manizales y del departamento de Caldas ha sido una de las motivaciones que lleva al planteamiento de un trabajo conjunto entre el Grupo de Trabajo Académico -GTA- de Ingeniería Hidráulica y Ambiental y el Grupo de Ambientes Inteligentes

Adaptativos -GAIA-, los dos de la Sede Manizales. Dicho trabajo se concentra en la construcción de un modelo para la administración y análisis de datos ambientales a través de una Bodega de Datos, que facilite la extracción de información relevante para la prevención de desastres naturales y la generación de alertas tempranas.

## ¿Qué es una bodega de datos?

En la actualidad se hace necesario para todas las organizaciones contar con información oportuna, dinámica, centralizada y de fácil acceso; esto con el fin de apoyar los procesos de análisis de datos y la toma de decisiones acertadas en los momen-

tos indicados. Una de las tecnologías que está apoyando el alcance de estos objetivos es la bodega de datos, ya que esta brinda una forma de almacenamiento que permite contar más que con una base de datos, con un sistema de información que

por lo general se implementa siguiendo un modelo multidimensional que facilita los análisis y extracción de datos (Fradel Abril & Castillo Pérez, 2007).

Las bodegas o almacenes de datos están orientados al manejo de grandes volúmenes de datos que provienen de fuentes de diversa índole y que cubren

### ¿Para qué la construcción de una bodega de datos ambientales?

El Instituto de Estudios Ambientales-IDEA-, de la Sede Manizales ha sido pionero en la dirección de proyectos de monitoreo ambiental que buscan contribuir a la generación de alertas tempranas para la prevención y atención de desastres. Dicho monitoreo se hace por medio de las mediciones realizadas por las redes de estaciones meteorológicas e hidrometeorológicas ubicadas en puntos estratégicos de la ciudad de Manizales y del departamento de Caldas.

Este monitoreo es realizado también por otras instituciones públicas y privadas de la región, lo cual hace que la cantidad de datos recolectados haya alcanzado una dimensión importante, puesto que además del creciente número de estaciones,

periodos de tiempo extensos, es decir, datos históricos. El objetivo de una bodega de datos es permitir obtener, a partir de los datos almacenados y de algunas herramientas adicionales, la información deseada para análisis y apoyo a la toma de decisiones. (Bustos Ríos, Duque, & Moreno Laverde, 2011).

se presentan datos desde hace casi 60 años para algunas estaciones.

Este antecedente hace necesario contar con una estrategia formal para el almacenamiento y administración de estos datos climáticos; la bodega de datos (Datawarehouse) ofrece una posibilidad de afrontar gran volumen y lograr tener no solo datos, sino información útil y de fácil acceso para la elaboración de análisis y reportes, así como también para la generación de indicadores que permitan una mejor comprensión del comportamiento climático de las diferentes zonas monitoreadas y su posible influencia en otro tipo de eventos naturales como deslizamientos, flujos torrenciales, inundaciones, sismos, entre otros.

### Ventajas de contar con una bodega de datos ambientales

Entre otras, las ventajas de contar con una bodega para la administración de datos ambientales son:

- Proporcionar información para la toma de decisiones, con el fin de mejorar su calidad, ya que estas pueden ser más oportunas y basadas en eventos históricos.

- Brindar poder de procesamiento, mediante una mayor flexibilidad e integridad de la información.
- Contribuir al proceso de transformación de los datos a la información y de la información al conocimiento.
- Al contar con un modelo dimensional, una bodega de datos puede admitir nuevas dimensiones y hechos, a medida que se requiera el estudio de nuevos fenómenos y eventos naturales medibles.

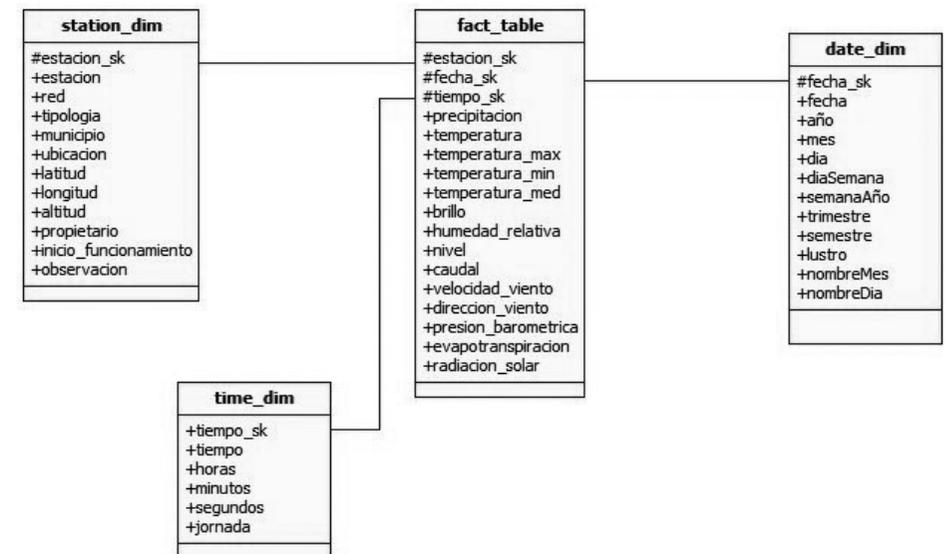
que contiene los datos para el análisis, rodeada de las tablas de dimensiones que ayudan a explicar la tabla de hechos («IBM Knowledge Center», 2013). Para este caso, los atributos de la tabla de hechos son: estación\_sk, fecha\_sk, tiempo\_sk, precipitación, temperatura, temperatura\_max, temperatura\_min, temperatura\_med, brillo, humedad relativa, nivel, caudal, velocidad\_viento, dirección\_viento, presión\_barométrica, evapotranspiración y radiación\_solar.

Las dimensiones básicas son estación, fecha y tiempo. En la figura 1 se puede apreciar el modelo diseñado. Como se aprecia en la figura, la tabla de hechos contiene principalmente los datos de las diferentes variables monitoreadas por las estaciones y en las tablas de dimensiones se da granularidad a aspectos que apoyan y explican la tabla de hechos.

### Avances del proyecto

Para la elaboración de la bodega de datos ambientales se ha empleado un Modelo en Estrella. Este esquema en estrella es un modelo de datos que tiene una tabla de hechos (o fact table)

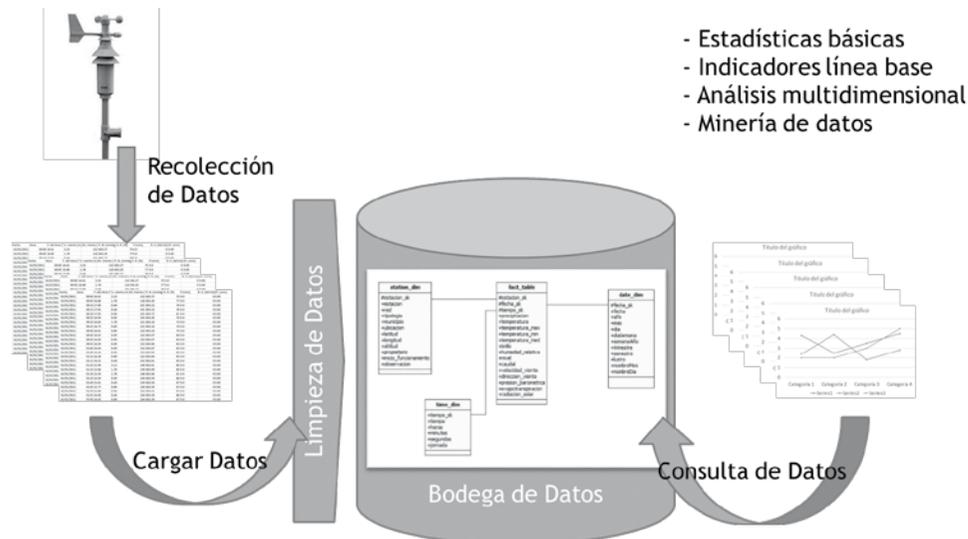
Figura 1. Estructura de la bodega de datos



En la figura 2 se presenta un esquema de la forma en la que se realiza el procesamiento de los datos, que va desde el momento en que los datos medidos por las estaciones de monitoreo ingresan en tiempo real al software central de recolección de datos; luego se hace la carga de datos, pasando primero por un proceso de limpieza, con el fin de garantizar la integridad y validez

de los datos. Posterior a dicha limpieza los datos son almacenados en la bodega de datos, de donde luego se pueden hacer diferentes análisis a través de la generación de estadísticas sencillas, indicadores, cubos OLAP (algunos avances se han alcanzado y se pueden encontrar en Duque-Méndez, Orozco-Alzate, & Vélez, 2014) y minería de datos.

Figura 2. Esquema del procesamiento de datos



### Trabajos futuros

Hasta el momento se ha avanzado en el diseño de la bodega y la carga de datos históricos a la misma. Como trabajo futuro se plantea la construcción de aplicativos que permitan afrontar los procesos de análisis de datos y hacer

el refinamiento del esquema de almacenamiento para la inclusión de nuevas variables y dimensiones. También se está trabajando en las estrategias para el filtrado y carga automática de los datos a la bodega.

### Conclusiones

El trabajo adelantado entre dos grupos de investigación adscritos al IDEA de la Sede Manizales ha dejado ver el potencial que se puede desarrollar al unir las capacidades de cada una de las disciplinas, consiguiendo con esto resultados muy positivos tanto en la solución informática como con la herramienta para el apoyo al monitoreo ambiental.

La bodega de datos ambientales podrá convertirse en un centro de datos muy fuerte, donde no solo se almacenen y administren los datos asociados a los proyectos del IDEA, sino que también se alberguen los datos recolectados por CORPOCALDAS y otras instituciones, de tal forma que se convierta en un proyecto a escala regional.

### Bibliografía

Bustos Ríos, L. S., Duque, N. D., & Moreno Laverde, R. (2011). Modelo de una bodega de datos para el soporte a la investigación bioinformática. *Scientia Et Technica*, 16(49), 145-152.

Duque-Méndez, N. D., Orozco-Alzate, M., & Vélez, J. J. (2014). Hydro-meteorological data analysis using OLAP techniques. *DYNA*, 81(185), 160. doi:10.15446/dyna.v81n185.37700

Fradel Abril, D. O., & Castillo Pérez, J. N. (2007). Estado actual de las tecnologías de bodega de datos y OLAP aplicadas a bases de datos espaciales. *Ingeniería e Investigación*, 27, 58-67.

IBM Knowledge Center. (2013, Enero 1). [CT701]. Recuperado 5 de febrero de 2015, a partir de [http://www-01.ibm.com/support/knowledgecenter/SS9UM9\\_8.1.0/com.ibm.datatools\\_dimensional.ui.doc/topics/c\\_dm\\_star\\_schemas.html?lang=es](http://www-01.ibm.com/support/knowledgecenter/SS9UM9_8.1.0/com.ibm.datatools_dimensional.ui.doc/topics/c_dm_star_schemas.html?lang=es)



**Instituto de Estudios Ambientales - IDEA -**  
*Teléfono: 8879300 Ext. 50190 / Fax 8879383*  
*Cra 27 #64-60 / Manizales - Caldas*  
*<http://idea.manizales.unal.edu.co>*  
*[idea\\_man@unal.edu.co](mailto:idea_man@unal.edu.co)*