

Gestión de Conocimiento para apoyar la Toma de Decisiones a la Contaminación en la ciudad de Santiago de Chile

Santiago Zapata Cáceres
Departamento de Informática y Computación
Facultad de Ingeniería
Universidad Tecnológica Metropolitana
Santiago, Chile
szapata@utem.cl

Luis Escobar Ramirez
Departamento de Informática y Computación
Facultad de Ingeniería
Universidad Tecnológica Metropolitana
Santiago, Chile
laescoba@utem.cl

Resumen - La contaminación del aire es uno de los mayores problemas que afecta a los países en casi cualquier parte del mundo. El incremento en las cantidades de gases y partículas potencialmente dañinas para la salud y el medio ambiente se ha constatado a escala mundial, y cada día es mas obvio que la respuesta a estos problemas se concentra en la gestión de conocimiento para apoyar la toma de decisiones.

En Chile, por ley, se establece la obligación de desarrollar planes de descontaminación. Durante el período otoño-invierno la población de la ciudad de Santiago, se ve afectada por un aumento repentino en los niveles de contaminación del aire. Esta situación es conocida como episodio crítico de contaminación atmosférica. Estos episodios se originan a partir de la convergencia de una serie de factores meteorológicos que impiden la buena ventilación de la cuenca de Santiago y/o debido a un incremento en las emisiones de contaminantes previos al evento.

El criterio para decretar un episodio crítico de contaminación esta referido al índice ICAP, que es generado por los datos de la Red de Monitoreo de gases y partículas que son ingresados a un Modelo de Pronóstico.

Esta investigación busca encontrar conocimiento inteligentes para la predicción y control de problemas medioambientales en la ciudad de Santiago de Chile.

Palabras Clave: Gestión de Conocimiento, Sistema de Soporte de Decisiones (DSS), Extracción de Conocimiento (KDD), Contaminación.

1. INTRODUCCIÓN

1.1. El Problema

La contaminación del aire en la ciudad de Santiago se origina principalmente durante el período otoño-invierno, en estos meses la población de la Región Metropolitana se ve afectada por un aumento repentino en los niveles de contaminación del aire.

Estas situaciones, conocidas genéricamente como episodios críticos de contaminación atmosférica, se producen cuando se registran elevados niveles de concentración de ciertos agentes contaminantes durante períodos de corta duración. Estos episodios se originan a partir de la convergencia de una serie de factores meteorológicos que impiden la buena ventilación de la cuenca de Santiago y también debido a un incremento en las emisiones de los agentes contaminantes.

El Índice de Calidad del Aire por Material Particulado, ICAP [1], [2], [3] es el indicador que a través de los datos emanados del modelo de pronóstico Cassmassi sirve como antecedente para que la autoridad de gobierno pueda determinar si la ciudad de Santiago esta en presencia de un episodio crítico de contaminación. El modelo Cassmassi pronostica el valor máximo de concentración promedio de 24 horas de material particulado respirable PM10, para el período de 00 a 24 horas del día siguiente.

Para este pronóstico de episodios un equipo de meteorólogos expertos prepara diariamente el pronóstico de condiciones meteorológicas asociadas a episodios de contaminación atmosférica para la Región Metropolitana, en base a información meteorológica actualizada, nacional e internacional y a los datos de calidad de aire de la red de monitoreo automático de calidad del aire y meteorología Red MACAM [2], [45].

- Es importante destacar que debido a que el modelo pronostica la calidad del aire del día siguiente, la declaración de un episodio por parte de la autoridad no implica que el aire haya empeorado, sino que podría llegar a empeorar. Es decir, los episodios se decretan en forma preventiva para evitar alcanzar los índices pronosticados, y así proteger la salud de la población.

Esta forma de abordar el problema y tomar decisiones presenta algunas características indeseables, tales como;

- El modelo de Cassmassi considera 8 estaciones de monitoreo en forma aislada y en muchas ocasiones se ha observado que durante la ocurrencia de episodios críticos, las estaciones interactúan entre sí.
- La permanencia del punto de vista reduccionista (en desmedro del enfoque sistémico) ha inducido la ocurrencia de innumerables errores de predicción.
- En un estudio realizado por la CONAMA [3], [48], [49] se manifestó que este modelo es inefectivo para predecir niveles críticos de contaminación. Para un período comprendido entre el 20 de marzo de 2003 al 10 de agosto de 2003 sólo ha acertado 8 veces en la predicción de situaciones críticas (alerta y preemergencia), con 30 falsas alarmas y 12 episodios subestimados.
- Debido al cambio climático global, cada año es más difícil realizar pronósticos confiables, pero este modelo no se ha adaptado para estas nuevas situaciones (es decir, mantiene sus parámetros estables, a pesar de que las condiciones en estudio cambian).
- Otro elemento que se destaca es que este modelo de Cassmassi no se ha actualizado desde que se diseñó, y ya han pasado más de seis años desde su creación, también se ignora que los técnicos del Cenma, organismo de la Universidad de Chile encargado de realizar los pronósticos para la toma de decisiones, están pidiendo desde hace tiempo su renovación.

Con respecto a la norma primaria de calidad del aire para material particulado grueso (PM10) se dice que el modelo actual de predicción sólo centra la atención en los valores medios y no en las variaciones hora por hora, lo que significa que los máximos de contaminación, dentro de cada día, quedan ocultos en el promedio de 24 horas.

1.2. Manejo del problema

En la actualidad y a nivel mundial, el estudio y control de problemas de contaminación medioambiental, se aborda con el apoyo de los llamados DSS (Decision Support Systems, Sistemas de Soporte de Decisiones en castellano) que son una clase específica de sistemas de información que apoyan los procesos de toma de decisiones en las organizaciones. Estos DSS son sistemas interactivos que ayudan a la toma de decisiones facilitando el manejo de los datos, documentos, conocimiento y/o modelos que se usan para resolver problemas dentro de las organizaciones.

Una de los ámbitos en que los DSS son de gran utilidad es en el apoyo a la toma de decisiones para análisis y control de problemas medioambientales, en particular el apoyo que puede brindar la Minería de Datos (Data Mining DM) al permitir extraer patrones, modelos, relaciones, tendencias, etc., que finalmente permiten encontrar “reglas” o “patrones” (“conocimiento”) a partir de los datos y comunicarlos al usuario a través de los DSS.

1.3. Objetivos de la Investigación

El objetivo general de la investigación es proponer una metodología basada en Gestión de Conocimiento como una alternativa al método actualmente usado para medir y tomar decisiones con respecto a las emergencias ambientales en la ciudad de Santiago.

2. GESTION ACTUAL DE LA CONTAMINACIÓN EN LA CIUDAD DE SANTIAGO

2.1. Introducción

Uno de los grandes problemas que posee la Región Metropolitana, especialmente la comuna de Santiago de Chile es la concentración en el ambiente de partículas contaminantes que dañan la salud de las personas que viven y transitan por esta ciudad. Por esto es importante contar con un buen sistema de apoyo a las decisiones que permita analizar los datos para obtener resultados que ayuden a tomar mejores decisiones de prevención para la ciudadanía.

Una entidad responsable que se encarga de velar por el cuidado del medio ambiente es el SESMA [44], quien a través de una red de medición llamada Red MACAM [45] van registrando diariamente los niveles de contaminación en el medio ambiente.

2.2. Red MACAM

La red oficial de Monitoreo Automático de Calidad del Aire y Meteorología de la ciudad de Santiago de Chile (Red MACAM) fue puesta en marcha en 1988 con 5 estaciones de monitoreo, mayoritariamente ubicadas en el sector céntrico de Santiago. En el invierno de 1997, dicha red fue renovada y ampliada a 8 estaciones de monitoreo automático, en lo que hoy se denomina como Red MACAM-2, dependiente del Servicio de Salud Metropolitano del Ambiente (SESMA). El sitio de Internet que permite obtener los datos es el siguiente:

<http://www.minsal.cl/sesma/pvcasesma/default.htm>

2.3. Normas Chilenas de Calidad del Aire

2.3.1. Normas

La Resolución N°369, del Ministerio de Salud del 12 de abril de 1988, publicada en el diario oficial el 26 de Abril de 1988, establece los índices de calidad del aire para determinar el nivel de contaminación atmosférica de la Región Metropolitana.

2.3.2 Índices ICA E ICAP

Tanto el ICA como el ICAP, dan origen a la siguiente calificación de la situación de acuerdo al valor obtenido:

INDICE	CALIFICACION
0 – 100	BUENO
101 – 200	REGULAR
201 – 300	MALO
301 – 400	CRITICO
401 – 500	PELIGROSO

2.4. Modelo Predictivo de Joseph Cassmassi

2.4.1. Introducción

Este modelo es usado para pronosticar los episodios de Alerta, Preemergencia o Emergencia, cuyas medidas que toma la autoridad han sido descritas en la introducción de la presente investigación, la información entregada por el modelo implica adelantarse al aumento excesivo de los Índices de Calidad del Aire considerando las variables meteorológicas y los niveles de contaminación por Material Particulado en suspensión.

El Modelo de Predicción usa ocho estaciones de monitoreo distribuidas geográficamente en la ciudad de Santiago de Chile y para su estimación considera: el comportamiento histórico de la contaminación atmosférica y su asociación con parámetros meteorológicos, niveles de contaminación atmosférica y su variación diaria. Todos estos factores son ingresados a un Polinomio matemático que fue desarrollado por el físico americano Joseph Cassmassi en el año 1999, basado en técnicas estadísticas clásicas de regresión lineal múltiple [47].

2.4.2. Modelo de pronóstico para episodios

El Plan de Prevención y Descontaminación Atmosférica para la Región Metropolitana (PPDA) estableció la necesidad de crear un modelo de pronóstico de episodios de contaminación, debido a la importancia de anticipar las medidas de control y de resguardo a la población frente a la ocurrencia de situaciones críticas.

Por su parte, el Decreto Supremo N° 59 de 1998 del Ministerio Secretaría General de la Presidencia, que establece la norma de calidad primaria para material particulado respirable (MP10), incluyó los requerimientos que debe satisfacer una metodología de pronóstico de calidad de aire.

De acuerdo con dichos requerimientos, mediante resolución N°12.612 de 1998 del Servicio de Salud Metropolitano del Ambiente de la Región Metropolitana (SESMA), se oficializó la primera aplicación de un modelo de pronóstico de calidad de aire en la ciudad de Santiago, en uso a partir de julio de 1998.

Como una forma de optimizar esta herramienta, en 1999 CONAMA encargó un estudio para mejorar la metodología de pronóstico de calidad de aire en la Región Metropolitana. El estudio dio como resultado un nuevo modelo de pronóstico, denominado modelo Cassmassi.

2.4.3. Variables utilizadas

La metodología de pronóstico de concentraciones de PM10 está basada en algoritmos de cálculo desarrollados mediante aplicación de técnicas estadísticas de regresión múltiple, enfocadas a encontrar relaciones entre posibles variables predictivas y la variable a predecir. Los predictores incluyen variables meteorológicas observadas, índices de condiciones meteorológicas observadas y pronosticadas, concentraciones de contaminantes observadas, índices de variaciones esperadas de emisiones y otros.

2.4.4. Potencial meteorológico de contaminación atmosférica (Algoritmos usados)

La aplicación operacional de esta metodología considera dos algoritmos de predicción para cada estación monitorea:

Un primer algoritmo incluye el "Índice de Potencial Meteorológico de Contaminación Atmosférica" (PMCA) pronosticado para el día siguiente. La figura 2.1 muestra que el PMCA tiene una buena correspondencia con los promedios de 24 horas de las concentraciones observadas de PM10.

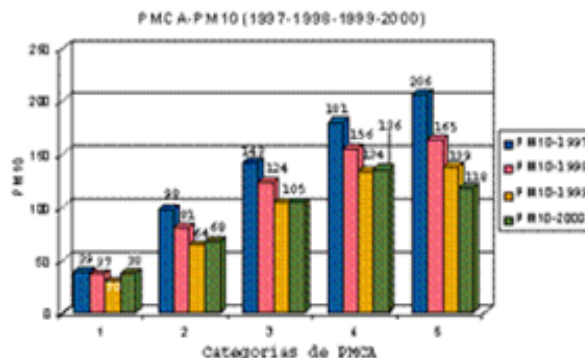


Figura 2.1.- Relación PMCA y contaminación por PM10.

El segundo algoritmo está basado solamente en observaciones (del mismo día y del día anterior). De esta forma, si el primer algoritmo no se puede aplicar por insuficiencia de la información, se utiliza el segundo algoritmo.

2.4.5. Validación del modelo

En 1999/2000, se realizó una evaluación del modelo Cassmassi por parte del Departamento de Geofísica de la Universidad de Chile. Para ello, se utilizó la información del período que se extiende desde el 1 de abril al 17 de septiembre de 1999. De esta manera, se garantizó la independencia de los datos usados para validar el modelo de los usados en su construcción. La validación muestra que el modelo Cassmassi cumple con los requisitos del Decreto Supremo. N° 59 de 1998.

La descripción detallada del modelo Cassmassi se encuentra en el estudio "Improvement of the Forecast of Air Quality and of the Knowledge of the Local Meteorological Conditions in the Metropolitan Region" (Cassmassi 1999), [48], [49].

2.5. Fallas del Modelo Predictivo de Cassmassi

El modelo predictivo es ineficiente para predecir niveles críticos de contaminación, por ejemplo, en el año 2003 solo acertó 8 veces la predicción de situaciones críticas, con 30 falsas alarmas y 12 episodios subestimados. Según estudios (Revista Que Pasa, 16 de Mayo del 2003), [50] indica que una de las causas principales de las fallas es que no se ha actualizado en 5 años, y que además de usar el modelo predictivo de Cassmassi oficialmente, se apoya en otros cinco modelos no oficiales, para estimar el grado de certeza del primero. Por ejemplo, un día cualquiera el pronóstico indicaba que había un 80% de probabilidades de alerta ambiental, pero con una incerteza del 35% de que se transformara en preemergencia ambiental.

Este modelo considera 8 estaciones de monitoreo en forma aislada y en muchas ocasiones se ha observado que durante la ocurrencia de episodios críticos, las estaciones interactúan entre sí. El adoptar el enfoque reduccionista en lugar del enfoque sistémico ha generado la ocurrencia de innumerables errores.

2.6. Decisiones a partir de los resultados del modelo

Diferentes son los efectos nocivos que puede traer a la salud de las personas la contaminación del aire que actualmente enfrenta la ciudad de Santiago. Sin embargo, éstos se tornan especialmente peligrosos en el caso de exposiciones a niveles muy elevados durante cortos períodos de duración, y que es lo que ocurre cuando estamos frente a un episodio crítico. Es entonces fundamental poder anticiparse a las situaciones de mayor riesgo para detener las negativas consecuencias que acarrea una emergencia ambiental.

Los resultados del modelo pronóstico son utilizados como herramienta de información preventiva por parte de las autoridades regionales, de esta manera, si el valor arrojado por el modelo pronóstico supera en alguna de las ocho estaciones:

El **nivel ICAP 200** según este modelo debiera decretarse **Alerta Ambiental**
 El **nivel ICAP 300** según este modelo debiera decretarse **Preemergencia Ambiental**
 El **nivel ICAP 500** según este modelo debiera decretarse **Emergencia Ambiental**

2.7. Costo de los errores del Modelo Predictivo

Según estudios económicos, los costos asociados a los errores del modelo predictivo superaron los \$ 2.000 millones de pesos en el año 2001 unos 2,7 millones de Euros, sobre los \$ 5.000 millones de pesos en el año 2002 unos 6,9 millones de Euros, y se estimaba que para el año 2003 esta cifra podría duplicarse alcanzando valores cercanos a los 12 millones de Dólares, [50], [52].

Como medidas correctivas, estos estudios recomendaban revisar el análisis de efectividad de las medidas adoptadas durante los episodios críticos; evaluar los costos asociados a la paralización de fuentes industriales y en tercer lugar revisar y mejorar la calidad predictiva de los modelos usados.

3. PROPUESTA DE TRABAJO

3.1. Introducción

De acuerdo a lo mostrado anteriormente para la gestión de los problemas ocasionados por la contaminación ambiental en la ciudad de Santiago, concluimos que al problema puede aportarse una solución basada en las Tecnologías de Información, como se detalla a continuación.

El **tratamiento actual** del problema presentado se puede esquematizar como lo muestra la figura 3.1, en la cual se puede destacar que el modelo de Cassmassi (utiliza herramientas estadísticas clásicas), no se ha actualizado desde su concepción inicial hace aproximadamente seis años, no justifica claramente los episodios críticos y basa su predicción en un promedio de comportamiento de las variables involucradas del día anterior, no es sistémico dado que no relaciona el efecto sinérgico de los datos generados por las diferentes estaciones de registro de datos que están distribuidas geográficamente en la ciudad de Santiago.

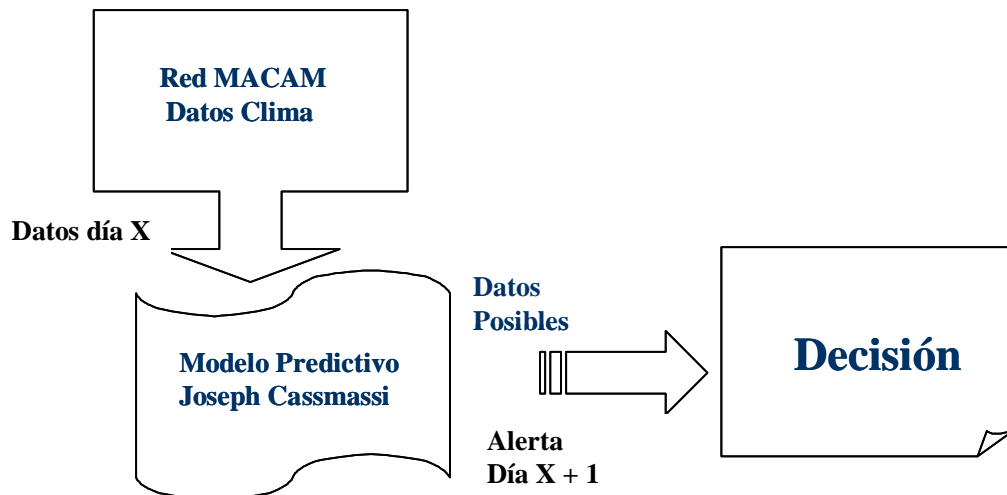


Fig. 3.1. Tratamiento actual de gestión de decisiones ambientales

La **propuesta alternativa**, propone disponer los datos básicos entregados por las estaciones de monitoreo de datos de la Red MACAM y la información meteorológica de la predicción del clima entregada por los servicios meteorológicos de la ciudad, en un Módulo llamado Almacén de Datos Histórico (Datawarehouse), para estructurar estos datos en el almacén de datos puede utilizarse la fecha de creación como una dimensión principal en la futura base de datos multidimensional, las ventajas de estructurar las dimensiones de esta manera permitirá: discriminar las alertas ambientales según la época del año (estacionalidad), evidenciando que habrá niveles críticos probablemente en épocas del año frías (invierno) potenciadas además por el fenómeno climático de la inversión térmica a que se ve afectada la ciudad de Santiago en épocas invernales y frías, además, podrá presentarse información del comportamiento de los datos por días, semanas, meses, etc.

El **módulo KDD** permite incorporar elementos que ayuden en forma inteligente al decisor, o sea el sistema debe ser capaz de realizar procesos de extracción de conocimiento, por ejemplo minería de datos, a la información histórica almacenada en el Datawarehouse y en un proceso de tipo Batch efectúe por ejemplo análisis de comportamiento de las variables involucradas en los últimos periodos, por ejemplo, los últimos dos años generando conocimiento a partir de los procesos de minería de datos realizada sobre los datos básicos históricos contenidos en el almacén de datos. Los resultados obtenidos por el tratamiento de los datos con las herramientas de Data Warehousing y de Minería de Datos serán presentados al usuario para apoyar su toma de decisiones a través de un Módulo DSS Sistema de Soporte de Decisiones, mostrado figura 3.2.

Este **módulo DSS** facilita al usuario decisor aplicar medidas con respecto a los episodios críticos de contaminación (alerta, preemergencia, emergencia), con interfases usuarias que permitan a través de consultas al almacén de datos ver que pasa si (What is/ What if) si se toman decisiones, además de proporcionaren forma automática ayudas o sugerencias que asistan la tomador de decisiones. La predicción de la calidad del aire es un tema apropiado para aplicar extracción de conocimiento y por ello usar minería de datos. Al tener almacenado información acerca de la calidad del aire que contienen una gran cantidad de datos, estos pueden ser analizados para encontrar entre ellos relaciones, regularidades y patrones de comportamiento.

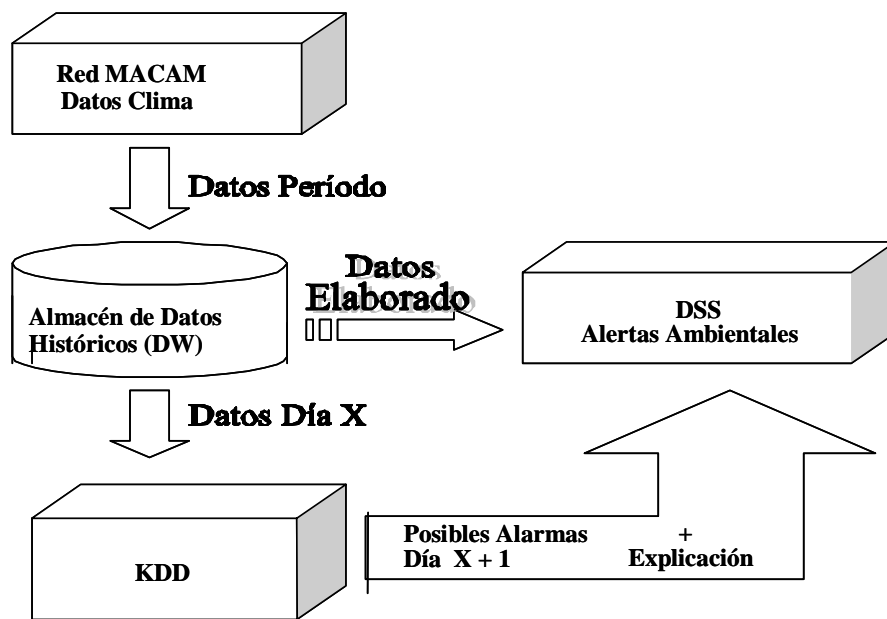


Fig. 3.2. Propuesta alternativa para apoyar la gestión ambiental

El DataWarehouse debe ser diseñado para responder a consultas (no previstas) relativas a encontrar conocimiento, analizado desde distintos puntos de vista.

Por ejemplo:

- Número de preemergencias decretadas en invierno.
- Alertas decretadas correctamente.
- Comuna con más preemergencias durante el año, agrupadas por estación del año

3.2. Diseño del DataWarehouse SESMA

Como primer paso, para cada actividad de la organización a analizar se debe diseñar un esquema estrella o copo de nieve, (Starchemaa y Snowflake). En este caso la actividad a analizar es la **SITUACIÓN**, entonces esta será nuestra “**tabla de hechos**”, la cual contiene la información importante del problema analizado.

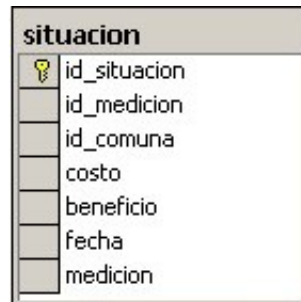


Fig. 3.3. Tabla de Hechos del Modelo Estrella

Esta tabla contiene las columnas *id_situación*, *id_medición* y *id_comuna*, además contiene las columnas *costo* que nos indica el costo asociado a una determinada situación, la columna *beneficio* que nos indica el beneficio asociado a una determinada situación más las columnas *fecha* y *medición* que nos indica la medición efectuada en un día determinado (esta medición fue almacenada anteriormente en la base de datos original OLTP de la Red MACAM)

Cada fila de la tabla de hechos contiene los datos observados (costo, beneficio, medición) de la actividad y de las referencias a las dimensiones que los caracterizan (claves externas: *id_medicion*, *id_comuna*, *fecha*). Generalmente una tabla de hechos representa una relación muchos es a muchos con sus respectivas dimensiones.

El segundo paso es definir nuestras “**tablas dimensiones**” y sus propiedades, son las dimensiones relevantes para nuestras consultas.

Las dimensiones que definimos para nuestro DataWarehouse son: “**medición**”, “**comuna**” y una dimensión de tiempo llamada “**fecha**”.

Con esto y luego de desnormalizar para evitar la concatenación de tablas al momento de realizar las consultas, el esquema estrella queda como se muestra en la figura 3.4.

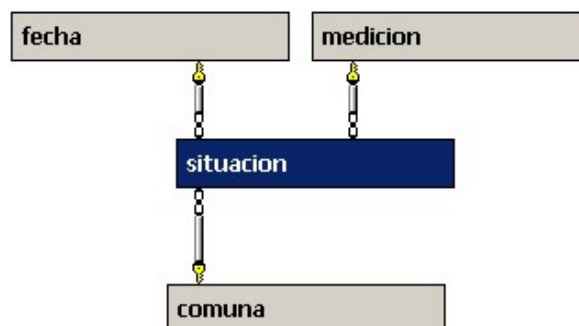


Fig. 3.4. Modelo Estrella DataWarehouse SESMA

3.3. Consultas al Datawarehouse

Estamos en condiciones de hacer consultas, por ejemplo, mostrar fecha en que la “situación” sea “preemergencia” en la “comuna” de “Pudahuel” en el “periodo” de “invierno”, esto se haría de la siguiente forma:

```
select medicion.nombre, fecha.fecha from situacion, medicion, comuna, fecha
where situacion.id_medicion=medicion.id_medicion and
medicion.nombre='preemergencia' and
situacion.id_comuna=comuna.id_comuna and
comuna.nombre='pudahuel' and
fecha.id_fecha=situacion.fecha and
fecha.estacion='invierno'
```

	nombre	fecha
1	preemergencia	2002-07-13 00:00:00
2	preemergencia	2002-07-14 00:00:00

Fig. 3.5. Resultado al consultar fecha y situación en periodo de invierno en la comuna de Pudahuel

En el caso anterior, se introdujo la dimensión de tiempo fecha en el análisis e incluimos un nuevo nivel de agregación sobre el tipo de fecha, esto es conocido como **drill-down**

Otra consulta que nos podría ser de interés para generar reportes que me permitan tomar decisiones sería: número de preemergencias decretadas en invierno, en código esta consulta se expresaría de la siguiente forma:

```
select COUNT (*)
From situacion, medicion, comuna, fecha
where situacion.id_medicion=medicion.id_medicion
and situacion.id_comuna=comuna.id_comuna
and fecha.id_fecha=situacion.fecha
and medicion.nombre='preemergencia'
and fecha.estacion='invierno'
```

	(Sin nombre de columna)
1	3

Fig. 3.6. Resultado al consultar número de preemergencias decretadas en invierno.

En relación a la consulta anterior podríamos mostrar las comunas y la fecha en que se decretó preemergencia durante el invierno. En código esta consulta se expresaría de la siguiente forma:

```
select fecha.fecha, comuna.nombre
From situacion, medicion, comuna, fecha
where situacion.id_medicion=medicion.id_medicion
and situacion.id_comuna=comuna.id_comuna
and fecha.id_fecha=situacion.fecha
and medicion.nombre='preemergencia'
and fecha.estacion='invierno'
```

	fecha	nombre
1	2002-07-13 00:00:00	PUDAHUEL
2	2002-07-14 00:00:00	PUDAHUEL
3	2002-07-13 00:00:00	CERRO NAVIA

Fig. 3.7. Resultado por pantalla consulta fecha y comuna con preemergencia decretada en invierno.

ID	FECHA	MEDICION	COMUNA	SITUACION	COSTO	BENEFICIO
1992	16/06/2002	329	PUDAHUEL	preemergencia	6317	6270
1995	19/06/2002	302	PUDAHUEL	preemergencia	4284	370
2019	13/07/2002	321	PUDAHUEL	preemergencia	3534	4040
2020	14/07/2002	305	PUDAHUEL	preemergencia	2697	550

Fig. 3.8. Vista de una consulta típica

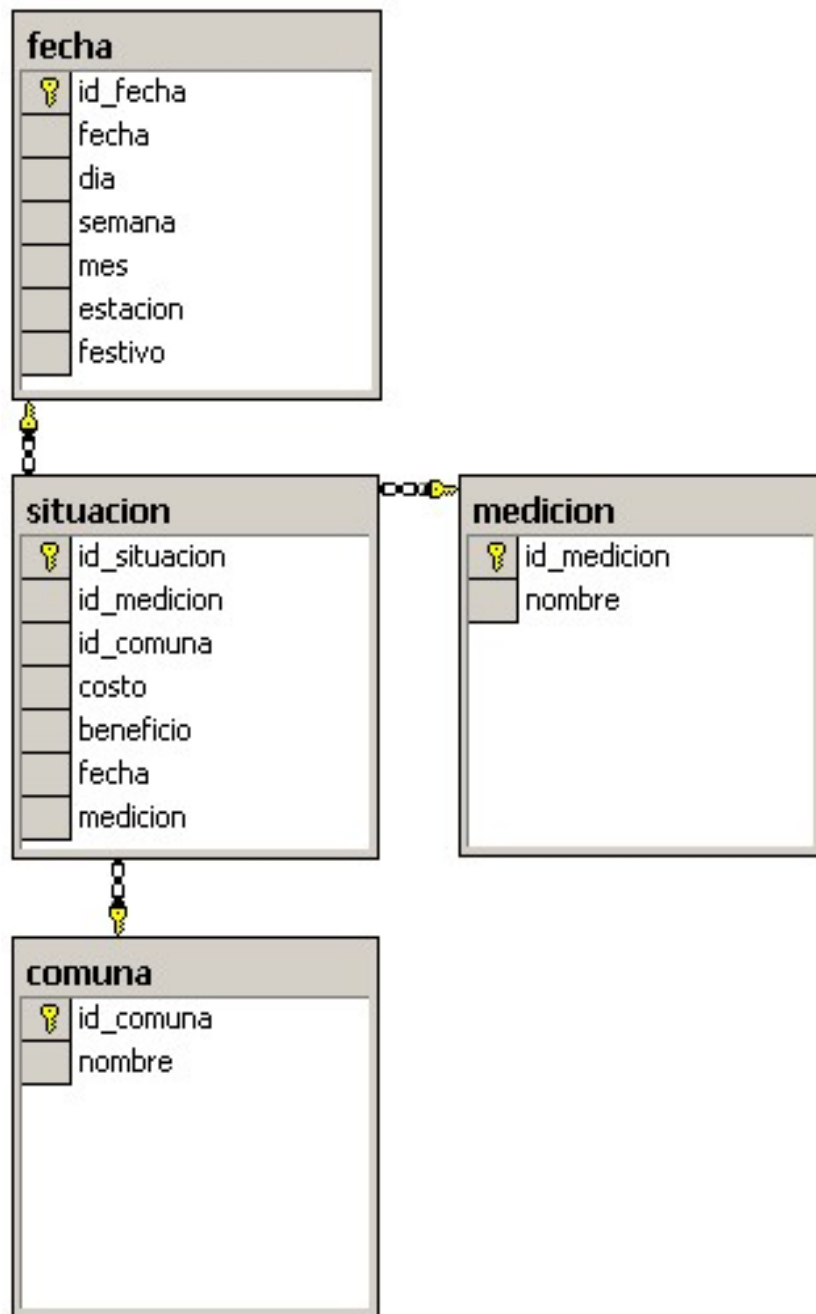


Fig. 3.9. Esquema Estrella de DataWarehouse Sesma

3.5. Requerimientos del Sistema

El motor de base de datos utilizado es SQL Server 2000, sobre un sistema operativo Windows XP Home Edition, con un procesador Intel Pentium IV de 2000 GHz con 256 Mb de RAM, el software utilizado para diseñar la interfaz gráfica para las consultas SQL es Microsoft Visual Basic 6.0, [56], [57].

3.4. Conclusiones

En términos generales, la contaminación del aire puede definirse como cualquier condición atmosférica en que ciertas sustancias alcanzan concentraciones lo suficientemente elevadas sobre su nivel ambiental normal, lo que puede producir un efecto negativo en las personas, en los animales y en la vegetación. En las grandes ciudades, la mayor parte de la contaminación es generada por emisiones del parque automotor y de las fuentes industriales. En la ciudad de Santiago esta situación se agrava debido a su ubicación geográfica y las condiciones meteorológicas que la caracterizan

El modelo predictivo de Cassmassi que se usa actualmente en la ciudad de Santiago pronostica el valor promedio de concentración máxima de material particulado respirable (PM10) para las 24 horas del día siguiente. Lo anterior lo realiza a partir de la información entregada por las estaciones de monitoreo de la Red MACAM que maneja el SESMA, además de la condición meteorológica esperada a escala local. El índice ICAP sirve como antecedente para que el modelo de Cassmassi determine la probable presencia de un episodio crítico de contaminación, debido a lo cual la autoridad correspondiente toma decisiones en forma preventiva.

Se puede concluir que el modelo predictivo de Cassmassi apoya a las autoridades ambientales a tomar decisiones con una precisión cercana al 65%, que en múltiples ocasiones las predicciones no han correspondido a las reales condiciones ambientales que se han presentado al día siguiente y se han tomado decisiones extremas cuando no ha correspondido y no se han tomado decisiones extremas cuando si la situación ambiental lo amerita, todo lo anterior genera altos costos en la salud de las personas y también costos económicos por la paralización de fuentes industriales cuando no ha sido necesario. Tal como hemos indicado en el desarrollo de la presente investigación, el modelo predictivo de Cassmassi se basa en un polinomio matemático que incorpora técnicas estadísticas clásicas, tales como el análisis de regresión lineal múltiple. Se ha criticado fuertemente al modelo por no estar concebido en forma sistémica, es decir no considera la sinergia que produce el efecto de interacción en cada una de las estaciones de las variables involucradas, (material particulado (PM10), ozono, dióxido de nitrógeno, monóxido de carbono, etc.

Se han realizado intentos por encontrar soluciones alternativas, en efecto, se han desarrollado en el mundo modelos de predicción que incorporan en su concepción, técnicas modernas de Redes Neuronales, de Lógica Difusa, de cambios dinámicos generados por los cambios climáticos globales de la tierra, parámetros de flujo de vehículos y ciclos de conducción en la ciudad, etc.

Por ejemplo un grupo de investigadores de la Universidad de Santiago de Chile, en el Departamento de Física, crearon un modelo basado en redes neuronales, que ha sido probado en paralelo con el modelo predictivo de Cassmassi, obteniendo resultados superiores al modelo de Cassmassi. Trabajos realizados por estos investigadores han sido presentados a congresos nacionales e internacionales [55], pero hasta la fecha la autoridad de gobierno no los ha considerado como una alternativa viable para implementarla en la ciudad de Santiago.

Dentro de la revisión realizada al uso de las nuevas Tecnologías de Información como herramientas de apoyo a la gestión de problemas ambientales, se ha encontrado que existen esfuerzos en el mundo por desarrollar herramientas que combinen técnicas de minería de datos y sistemas de soporte de decisiones, pero para aplicaciones específicas a lugares geográficamente definidos.

Las consultas sobre un Data Warehouse son mucho mas sencillas que en una base de datos relacional dado que mucha información se filtra antes de poblar la “tabla de hechos” y sus respectivas “dimensiones” tomando en cuenta además la gran ayuda que nos proporciona el proceso de desnormalizar, con lo cual se obtiene un código más simple y que entrega datos que constituyen información relevante para la toma de decisiones.

El diseño adoptado para el Data Warehouse permite generar conocimiento importante para la Toma de Decisiones acerca la calidad del aire en la ciudad de Santiago de Chile, decisiones basadas en reportes, gráficos, etc. Parte importante de la investigación, además del diseño del Data Warehouse, esta en diseño de un sistema poderoso para gestionar los datos y que pueda entregar información útil y confiable para la toma de decisiones. Al tener un almacenamiento multidimensional se facilitará tratamientos a los datos, y técnicas como la Minería de Datos permitirá extraer conocimiento no explorado a partir de los datos básicos contenidos en el almacén de datos históricos, intentando además que el propio sistema propuesto pueda justificar la predicción ante el decisor (explicación).

Esta característica proporciona una consistencia mayor al sistema propuesto con respecto a lo entregado actualmente bajo el enfoque de usar el modelo de Cassmassi, ya que el decisor tiene antecedentes en cuanto al porque se ha producido una alerta y además podrá discriminar si la alerta se ha producido por un alza en los limites permitidos en las variables que se relacionan con los gases o en las variables que se relacionan con el material particulado o puede explicar que la conjunción de algunas variables esta gatillando un evento de alerta ambiental. .

Por ejemplo, al usar minería de datos puede que sea útil construir árboles de decisión, ya que la predicción que se desea realizar se hará en base a categorías (bueno, regular, malo, critico). Otro aporte de la propuesta es que incorpora el enfoque sistémico para entregar una predicción, usando toda la sinergia que se puede lograr al tratar los datos de la ciudad como un todo en forma de sistema. En resumen, la propuesta que presentamos en esta investigación es más robusta que el sistema actualmente en uso por las autoridades en lo referente al tratamiento de las emergencias ambientales usando el modelo de Cassmassi y la información entregada por la Red MACAM y los organismos en cargados de las predicciones climáticas.

El trabajo futuro es promisorio, pues encontrar soluciones alternativas al problema de la contaminación en la ciudad de Santiago de Chile se ha transformado en una cruzada de carácter humanitario, ya que los problemas a la salud de la población y a la economía del país y del mundo, están alcanzando niveles que sobrepasan lo aceptable por la comunidad.

4. BIBLIOGRAFIA

- [1] www.sesma.cl/sitio/pag/aire/Indexjs3aire005.asp
- [2] www.sesma.cl/sitio/pag/aire/Indexjs3airesist_monit.asp
- [3] <http://www.conama.cl/portal/1255/channel.html>
- [4] R. L. Ackoff “Science and the Systems Age: Beyond I.E., O.R. and M.S.”. Operations Research, May-June 1973.
- [5] J. G. Saxe (1816 – 1887)
- [6] W. Churchman “The Systems Approach” Dell Pub. (1968)
- [7] <http://w3.mor.itesm.mx/~albreyes/sistemas/>
- [8] Power, D. J. A Brief History of Decision Support Systems. DSSResources.COM, World Wide Web, <http://DSSResources.COM/history/dsshistory.html>, version 2.6, November 12, 2002 .
- [9] Ferguson and Jones. Management Decision System: Computer-Based Support for Decision Making, 1969.

- [10] Gorry, G.A. y Scott-Morton, M.S. (1971): "A Framework for Management Information Systems", Sloan Management Review, vol 13, n° 1, fall 1971., Scott Morton, M.S. (Eds.), The Corporation of the 1990s, Oxford University Press, Oxford, 1991.
- [11] Sprague, R., Carlson E. Building Effective Decision Support Systems, Englewood Cliffs, NJ, Prentice Hall, May 1982
- [12] J. P. Shim, Merrill Warkentin, James F. Courtney, Daniel J. Power, Ramesh Sharda, Christer Carlsson Past, Present, and Future of Decision Support Technology, Decision Support System 32 (2002) 111-126, www.elsevier.com/locate/dsw
- [13] Martha Patricia Bobadilla, El impacto en los negocios del DSS, 07 /05 / 2001
<http://www.claveempresarial.com/soluciones/notas/nota010507b.shtml>
- [14] <http://www.bitam.com.mx/AcercaDeBI.htm>
- [15] Diana Cristina Romero Sánchez, La Importancia De Los GDSS En El Trabajo Colaborativo, dromero@campus.zac.itesm.mx <http://www.netmedia.info/netmedia/articulos.php> , Netmedia S.A.
- [16] Los PDAs, la última frontera para los MSS
http://www.pdaexpertos.com/Articulos/Experiencias_de_Usuarios/7.shtml , Everardo Alpuche M., Ingeniero en Sistemas Computacionales, estudiante de Maestría en Administración de Tecnologías de Información, Universidad Virtual del ITESM, México.
- [17] Turban, Efraim & Arosen, Jack E. (1998). Decision Support Systems and Intelligent Systems. EEUU: Prentice-Hall. Quinta Edición.
- [18] Hernández A., Federico (1997). "Revista Soluciones Avanzadas", Sistema Gerencial Administrativo para el Soporte de Decisiones, Infolatina, 14 de Enero de 2001. [19] Porter, Michael E., Competitive Strategy: Techniques for Analyzing Industries and Competitors, Free Press, New York, 1980.
- [20] Power, D. J. Decision Support Systems Web Tour. World Wide Web, <http://dssresources.com/>, version 4.2, June 15, 2002.
- [21] Scott Morton, M.S. (Eds.), The Corporation of the 1990s, Oxford University Press, Oxford, 1991.
- [22] BARRY DE VILLE, Data Mining en SQL Server 2000,
http://www.w2000mag.com/sqlmag/atrasados/04_mar01/articulos/portada_1.htm
- [23] Tuya Javier, Adenso Díaz; "Los Decisión Support Systems: Arquitectura y Aplicaciones Empresariales", ETS Ingenieros Industriales e Informaticos, Universidad de Oviedo, <http://www.di.uniovi.es/~tuya/pub/ati-98-dss-resumen.html>
- [24] Sprague, R. H., "A Framework For the Development of Decision Support Systems". Management Information Systems Quarterly, No. 4. pp 1-26, 1980. Sprague, R.H., y H.J. Watson, Decision Support for Management, Prentice Hall, New Jersey, 1996.
- [25] Marín Ruiz Nicolás, Introducción al Data Warehousing, Decsai, Universidad de Granada, Granada, España.
- [26] W.J. Frawley, G. Piatetsky-Shapiro y C.J Matheus. "Knowledge Discovery in Database: an Overview". Knowledge Discovery in Database. AAAI-MIT Pres, Menlo Park, California 1991, páginas 1-27.
- [27] U.M. Fayyad "Data Mining to Knowledge Discovery: making sence out of data". En IEEE Expert Vol. 5, octubre 1996, páginas 20-25.
- [28] Artículo de Internet con el título "Data Mining: torturando los datos hasta que confiesen", por Luis Carlos Molina Felix, coordinador del programa de data Mining (UOC), 2001
- [29] IBSNAT. 1989. Decision Support System for Agrotechnology Transfer (DSSAT) Version 2.1, Dept. of Agronomy and Soil Science, College of Tropical Agriculture and Human Resources. University of Hawaii, Honolulu.
www.mic.hawaii.edu/dev_tech/software/dssat.html
- [30] Copyright 1995-2002 by: ESS Environmental Software and Services GmbH AUSTRIA
<http://www.ess.co.at/AIRWARE/>
- [31] <http://www.ucar.edu/communications/newsreleases/2003/mdss.html><http://www.amazings.com/ciencia/noticias/270103a.html>
- [32] <http://www.baronams.com/EDSS/>
- [33] http://strategis.ic.gc.ca/Ces_Web/providers_info_.cfm?CES_ESTBLMT_NO=2641&target=English
- [34] <http://www.riks.nl/PROJECTS/KRIM> <http://www.krim.uni-bremen.de/englisch/indexenglisch.html>
- [35] <http://www.aaas.org/international/lac/plata/baethgeninfo.shtml>
<http://www.inia.org.uy/disciplinas/agroclima/index.html>
- [36] Biss, A. "Dynasty Triage Advisor Enables Medical Decision-Support", 2002, at URL DSSResources.COM. www.dynasty.com.

- <http://dssresources.com/cases/dynasty.html>
- [37] <http://www.mega-supply.com/websites/megasupply/ap1-2.htm> <http://www.ap1soft.com/>
- [38] <http://www.iiasa.ac.at/Research/WAT/docs/desert.html>.
- [39] http://www.netmeye.net/enespañol/experiencias/Soporte_de_decisiones.htm
- <http://www.inta.gov.ar/bariloche/ssd/rn>
- [40] <http://www.water.ncsu.edu/watershedss/index3.html>
- [41] <http://www.riks.nl/projects/Xplorah>.
- http://www.riks.nl/RiksGeo/projects/xplorah/XplorahBrochurePrint_es.pdf
- [42] <http://www.geogra.uah.es/Proyectos/sedis.htm>
- [43] <http://www2.ing.puc.cl/gescopp/investigacion.html>
- [44] SESMA (Servicio de Salud Metropolitano del Ambiente), www.sesma.cl
- [45] RED MACAM (Red de Monitoreo Automático de Calidad del Aire y Meteorología), www.conama.cl/rm/568/article-1114.html
- [46] Universidad de Santiago de Chile, "Normativa Ambiental en Aire. (Fuentes Fijas)", <http://lauca.usach.cl/ima/sesma-1.htm>
- [47] <http://www.cleanairnet.org/lac/1471/article-40847.html>.
- [48] <http://www.conama.cl/rm/568/article-1183.html>.
- [49] <http://www.conama.cl/rm/568/article-2581.html>.
- [50] <http://www.quepasa.cl/revista/2003/05/16/t-16.05.QP.NAC.CONTAMINACION.html>
- [51] <http://www.geofisica.cl/papers/pherman.htm>.
- [52] <http://www.sustentable.cl/portada/Descontaminacion/2453.asp>
- [53] <http://www.sesma.cl/sitio/pag/aire/Indexjs3aire005.asp>
- [54] WEKA (Waikato Environment for Knowledge Analysis) <http://www.cs.waikato.ac.nz/~ml/>
- [55] <http://www.geofisica.cl/English/pics3/FUM6.htm>
<http://www.geofisica.cl/Meteorol.htm>
- [56] <http://www.microsoft.com/spain/sql/>
- [57] <http://www.microsoft.com/latam/sql/>