

## MODELACIÓN INVERSA APLICADA AL ANÁLISIS DEL INVENTARIO DE EMISIONES DE LA ZONA METROPOLITANA DE GUADALAJARA, MÉXICO

Alberto MENDOZA<sup>1</sup> y Marisa R. GARCÍA<sup>2</sup>

<sup>1</sup> Departamento de Ingeniería Química, Tecnológico de Monterrey, Campus Monterrey, Ave. Eugenio Garza Sada No. 2501, Monterrey Nuevo León, México, C. P. 64849, mendoza.alberto@itesm.mx

<sup>2</sup> Centro de Calidad Ambiental, Tecnológico de Monterrey, Campus Monterrey, Ave. Eugenio Garza Sada No. 2501, Monterrey Nuevo León, México, C.P. 64849.

*(Recibido noviembre 2009, aceptado junio 2011)*

Palabras clave: modelación atmosférica, calidad del aire, ozono, fuentes de emisión

### RESUMEN

Uno de los componentes con mayor grado de incertidumbre en el proceso de aplicar modelos complejos de calidad del aire es el inventario de emisiones. Si dicha información contiene errores, las conclusiones resultantes de la aplicación del modelo se verían empañadas. En este trabajo, se presenta el uso de una técnica de modelación inversa para ajustar el inventario de emisiones de la Zona Metropolitana de Guadalajara, México, el cual se compara contra el inventario de emisiones original. El sistema de modelación utilizado estima, a través de un proceso iterativo, factores de escalamiento a las emisiones de un caso base de tal manera que, al aplicar dichas correcciones, se minimiza el error entre las concentraciones observadas y las simuladas por el modelo directo. Los resultados obtenidos indican que, durante el caso de estudio analizado, las emisiones diarias de CO tendrían que estar sujetas a variaciones globales que van desde reducciones de ~16% a incrementos de ~60%. Las emisiones de NO<sub>x</sub> y SO<sub>2</sub> aparentan estar subestimadas, requiriéndose aumentos del orden de ~100 % a ~150 % sobre el caso base para los NO<sub>x</sub> y de 20% a ~140% del valor base para el caso del SO<sub>2</sub>. Los cambios sugeridos por el modelo inverso aumentan significativamente el desempeño del modelo directo en el sentido de su capacidad de replicar las observaciones, incluyendo las de O<sub>3</sub>, un contaminante secundario. Sin embargo, no todas las discrepancias entre observaciones y valores simulados quedan resueltas al, por ejemplo, haber tenido que limitar el escalamiento de las emisiones hora a hora sin diferenciar entre subdominios, debido al número limitado de estaciones de monitoreo disponibles en la zona, y por el hecho real que el modelo directo no es perfecto.

Key words: atmospheric modeling, air quality, ozone, emission sources

### ABSTRACT

Emission inventories are one of the sources of highest uncertainty in the process of applying complex air quality models. If there are errors in such information, the conclusions that result from the model application can be clouded. This work presents the application of an inverse modeling technique to adjust the emissions inventory of the Guadalajara Metropolitan Area, Mexico, which is then compared against the original

inventory. The modeling system estimates, by means of an iterative process, scaling factors to the emissions of a base case in such a way that after applying such corrections the error between observed and simulated concentrations is minimized. The obtained results indicate that, during the analyzed case study, the daily CO emissions needed to be globally varied from reductions in the order of ~16% to increments of ~60%. NO<sub>x</sub> and SO<sub>2</sub> emissions are both apparently underestimated: NO<sub>x</sub> emissions were increased in the range of 100% to 150% with respect to the base case, while SO<sub>2</sub> emissions increased in the range of 20% to 140% the nominal value. The changes suggested by the inverse model significantly increase the performance of the direct model, in the sense of its ability to replicate the observations, including O<sub>3</sub>, a secondary pollutant. However, not all discrepancies between observations and model-derived concentrations were solved due to the limitation of only scaling hourly domain-wide emissions but not by sub-domains since not enough monitoring stations were available and by the fact that the direct model is not perfect.

## INTRODUCCIÓN

Los modelos tridimensionales de calidad del aire empleados para estudiar la dinámica de contaminantes atmosféricos han demostrado ser una de las herramientas más robustas para estudiar la problemática ambiental de cuencas atmosféricas (Russell y Dennis 2000, Russell 2008). Estos modelos permiten construir campos de concentración que varían espacial y temporalmente; asimismo, permiten determinar explícitamente relaciones emisor-receptor para múltiples contaminantes y fuentes, lo cual permite identificar las estrategias de reducción de emisiones que brindan el mayor beneficio a los receptores. Actualmente, la mayoría de estos modelos operan en modo dinámico, es decir, después de su inicialización la solución obtenida por el modelo es función exclusiva de la integración de la ecuación de conservación de especies químicas respecto al tiempo, sujeta a condiciones meteorológicas, tasas de emisión y condiciones frontera predefinidas. Típicamente, los modelos de calidad del aire se aplican, primero, a uno o varios episodios históricos para los cuales los niveles de contaminantes están suficientemente bien caracterizados y los datos para alimentar el modelo están disponibles, con el fin de evaluar el desempeño del mismo. Esto se logra comparando las observaciones puntuales con las concentraciones arrojadas por el modelo y calculando una serie de estadísticos que cuantifican el desempeño. Una vez que se decide que el modelo describe apreciablemente bien el caso o los casos base, se emplea para análisis subsecuentes. Por ejemplo, en experimentos numéricos para evaluar estrategias de control (ver Cohan *et al.* 2007, Bergin *et al.* 2008), análisis de sensibilidad (Yang *et al.* 1997, Hakami *et al.* 2003, Napelenok *et al.* 2006) y análisis de procesos (Jang *et al.* 1995, O'Neill *et al.* 2005, Kimura *et al.* 2008).

Sin embargo, las aplicaciones típicas de modelos de calidad del aire tienden a no coincidir, en cierto grado, con las mediciones reportadas para los contaminantes atmosféricos (Russell y Dennis 2000, Hogrefe *et al.* 2008). Esta desviación puede tener su origen en errores en los datos suministrados al modelo (p. ej., campos meteorológicos o emisiones) o en los parámetros del modelo mismo (p. ej., las constantes cinéticas del mecanismo fotoquímico). Asimismo, la resolución temporal y espacial del modelo también juega un papel en determinar la precisión de los resultados. Los datos arrojados por el modelo son valores promedio sobre un volumen (representado por una celda computacional dentro de la malla tridimensional empleada para resolver numéricamente las ecuaciones de transporte y transformación química), mientras que los valores observados son puntuales (aunque pueden llegar a tener una cierta representatividad espacial). Así, no se esperaría que los valores simulados coincidieran completamente con los valores observados. Sin embargo, cuando las diferencias son significativas, es indicativo de problemas en la aplicación del modelo que tienen que resolverse, ya que de otra manera las conclusiones que se pudieran derivar del ejercicio de modelación pudieran ser incorrectas.

Uno de los procedimientos formales para resolver las discrepancias entre valores observados y valores simulados es la modelación inversa. La modelación inversa permite directa o indirectamente incorporar información observada (monitoreada) a un modelo determinístico al modificar sus entradas o sus parámetros. Así, este esquema intenta obtener una mejor descripción de los procesos que el modelo simula al añadir información del estado real del sistema. Entre los esquemas numéricos que se han empleado para tal fin en modelos de calidad del aire se encuentran los

filtros Kalman (Chang *et al.* 1997, Napelenok *et al.* 2008), la asimilación variacional (Elbern y Schmidt 1999, Elbern *et al.* 2007), las ecuaciones adjuntas (Davydova-Belitskaya *et al.* 2001, Hakami *et al.* 2005) y la inversión matricial directa (Brown 1993, Mendoza-Domínguez y Russell 2000).

En este trabajo se presenta la aplicación de una técnica de modelación inversa basada en inversión matricial directa descrita por Mendoza-Domínguez y Russell (2000) para investigar la simulación realizada de un caso base sobre la Zona Metropolitana de Guadalajara (ZMG), para la cual se empleó un modelo de calidad del aire con un inventario de emisiones con cierto grado de incertidumbre que, a su vez, pudo influir en discrepancias observadas entre valores observados y simulados. Así, el objetivo general de este trabajo es explorar el uso de la técnica de modelación inversa mencionada en conjunto con el modelo de calidad del aire, con el fin de obtener un inventario de emisiones ajustado, de modo que el desempeño del modelo se vea beneficiado. Esto daría información sobre posibles áreas de mejora en el inventario de emisiones que tendrían que ser analizadas detalladamente en caso de que las modificaciones sean de escenarios plausibles o bien indiquen otras posibles fuentes de incertidumbre en el proceso de modelación.

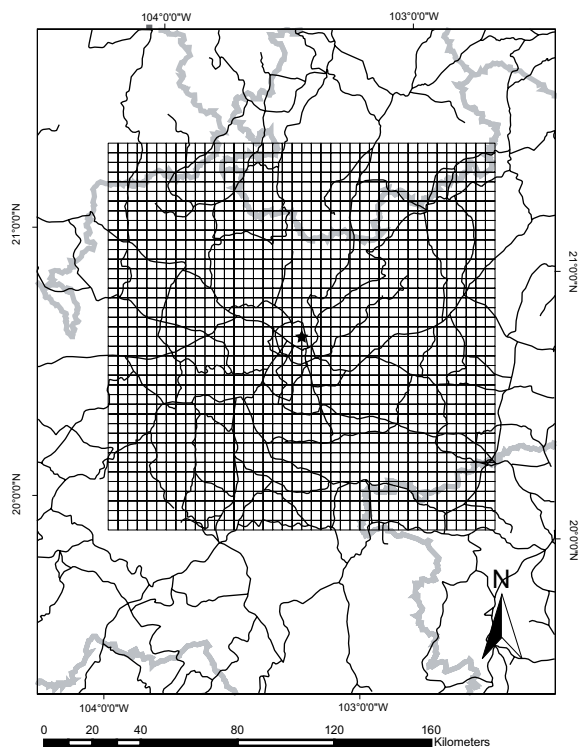
## CASO DE ESTUDIO

El área de estudio, la ZMG, ubicada en el centro del estado de Jalisco, es una de las regiones con los niveles más altos de contaminantes atmosféricos del país. Según datos de 2005, y de acuerdo a un indicador ponderado de la frecuencia de días en los que se excede el límite normado de calidad del aire y de la magnitud de las concentraciones alcanzadas, la ZMG fue la ciudad más contaminada del país respecto a sus niveles de NO<sub>2</sub> (precursor de O<sub>3</sub>), y la segunda respecto a O<sub>3</sub> y CO (Zuk *et al.* 2007). La principal fuente de contaminantes atmosféricos en la ZMG es el sector transporte, el cual contribuye con cerca del 75 % del total de emisiones (Gobierno del estado de Jalisco *et al.* 1997). El parque vehicular está constituido principalmente por vehículos ligeros y camionetas *pick-up* a gasolina (~91 %); los vehículos a diesel representan poco más del 5 % del parque vehicular. Los sectores industrial y de servicios representan el 5 % del total de emisiones, siendo su principal contribución las emisiones de SO<sub>2</sub> (~68 %) e hidrocarburos (40 %) (Gobierno del Estado de Jalisco *et al.* 1997). Los giros industriales con mayor contribución al inventario de emisiones son los de productos de consumo

alimenticio (25 %), minerales no metálicos (20 %) y productos de impresión (12 %). El 20 % restante del total de emisiones proviene de fuentes naturales. Los episodios de contaminación atmosférica en la ZMG se ven exacerbados por las condiciones topográficas y meteorológicas de la región (Gobierno del Estado de Jalisco *et al.* 1997). Las zonas montañosas que rodean la ciudad (Sierra Madre Occidental y Eje Neovolcánico) constituyen una barrera física natural para la circulación del viento. Además, los sistemas anticiclónicos que se forman tanto en el Golfo de México como en el Océano Pacífico ocasionan gran estabilidad atmosférica sobre la región que inhibe el mezclado vertical del aire. El viento dominante en la ZMG proviene del oeste con un 15.5 % de la frecuencia total, siguiéndole el viento proveniente del este con un 7.5 %. Los periodos de calma o con vientos menores a 1 m/s alcanzan una frecuencia mayor al 44 %. Finalmente, debido a su latitud (20 °N), la ZMG recibe una cantidad apreciable de radiación solar, lo cual hace que su atmósfera sea altamente fotorreactiva.

Como caso de estudio se empleó el caso base reportado por Mendoza y García (2009) para una modelación de la dinámica de contaminantes fotoquímicos en la ZMG. Este caso es en particular apropiado al tener de origen un inventario de emisiones que, dado su proceso de conformación, se sospecha pudiera tener un grado de incertidumbre tal que los resultados no pudieran estar representando correctamente algunos procesos atmosféricos de la zona. En dicho estudio se aplicó el modelo de calidad del aire CIT (California/Carnegie Institute of Technology), versión 3.0 (McRae *et al.* 1982a, Russell *et al.* 1988, Harley *et al.* 1993) a un episodio ocurrido del 16 al 18 de mayo de 2001. El dominio de modelación consiste en una malla computacional de 40 x 40 celdas en la horizontal, cada una de 4 x 4 km (**Fig. 1**) y seis niveles en la vertical para una altura total de 3100 m. Cabe notar que el modelo se comienza a ejecutar con datos del 14 de mayo, y así se aprovechan los dos primeros días para estabilizar la respuesta del modelo y minimizar el impacto de las condiciones iniciales en los campos de concentración derivados para los días clave del episodio. Los detalles de la generación de los campos meteorológicos y la especificación de las condiciones iniciales y de frontera no se reportan al estar descritos por Mendoza y García (2009).

El inventario de emisiones empleado en el caso de estudio reportado por Mendoza y García (2009) toma como base el inventario oficial de 1995 para la ZMG. El inventario tuvo que ser extendido para cubrir los municipios adicionales que se incluyen en el dominio



**Fig. 1.** Definición del dominio de modelación alrededor de la ZMG (★). Se presenta un recuadro interior que representa los límites del dominio de modelación y, como referencia, los límites estatales y principales vías de comunicación

de modelación y no son propiamente parte de la ZMG. Asimismo, tuvo que ser escalado del año base (1995) al año de modelación (2001). Además, el inventario tuvo que ser segregado espacialmente con base en la densidad estimada en la ubicación de las fuentes de emisión y se aplicaron perfiles temporales de emisión para obtener la variación de las emisiones hora a hora. Finalmente, las emisiones de los compuestos orgánicos volátiles tuvieron que ser sometidos a un proceso de “especiación” química para poder describir el comportamiento de especies individuales o grupos de pseudoespecies a partir de los datos agregados del inventario original. Todo este proceso conlleva a tener un inventario que fue construido en base a múltiples suposiciones, lo cual lo hace sujeto a tener incertidumbres. Aunado a lo anterior, la posibilidad de que el inventario del año base (1995) pudiera tener fallas de origen, también afectaría los resultados de la simulación.

## EL SISTEMA DE MODELACIÓN

El sistema de modelación empleado consiste en un modelo directo (el modelo de calidad del aire) que

determina campos de concentración para el dominio de modelación seleccionado, acoplado a un módulo de modelación inversa. El módulo de modelación inversa estima factores de escalamiento a valores de datos de entrada al modelo o parámetros del mismo que lograrían, una vez aplicados, acercar los valores de concentración de contaminantes atmosféricos simulados por el modelo directo a los valores observados.

*Modelo directo.* En este estudio se empleó el mismo modelo de calidad del aire que la aplicación original de Mendoza y García (2009): el CIT v. 3.0, con la diferencia de que se activó el módulo de análisis directo de sensibilidad con la que cuenta el modelo. El CIT es un modelo euleriano tridimensional de segunda generación, capaz de simular la dinámica de contaminantes gaseosos emitidos en una determinada cuenca atmosférica. El modelo cuenta con módulos que procesan el transporte, reacción química y remoción por depositación en seco de contaminantes atmosféricos, sujeto a condiciones meteorológicas, topográficas y de uso de suelo locales. Los detalles de la formulación matemática del CIT se encuentran reportados extensamente en la literatura, por lo que no se repiten aquí (McRae *et al.* 1982a, Russell *et al.* 1988, Harley *et al.* 1993). Basta comentar que el modelo resuelve la ecuación de conservación de especies químicas sobre un volumen predefinido, dividiendo dicho volumen en un arreglo de celdas tanto horizontales como verticales, las cuales forman una malla en tres dimensiones. El transporte horizontal se resuelve en este estudio empleando un esquema de advección parabólico (Srivastava *et al.* 2000), el cual tiene la propiedad de conservar la masa y proporciona una solución monotónica (Odman 1998). Mendoza y García (2009) compararon el uso de este esquema de transporte contra un esquema implícito basado en un método de elementos finitos de cuarto orden (McRae *et al.* 1982b) demostrando una superioridad marginal del primero para la aplicación a la ZMG.

El tratamiento de las reacciones fotoquímicas está basado en el mecanismo cinético SAPRC90 (Carter 1990), el cual contempla 89 especies químicas y 207 reacciones. No se incluye en el modelo un módulo para el tratamiento de la dinámica de aerosoles atmosféricos, únicamente se ajustan los valores de las concentraciones de  $\text{HNO}_3$  y  $\text{NH}_3$  en fase gaseosa después del paso de cinética química resolviendo el equilibrio químico entre estas dos especies y el nitrato de amonio en fase aerosol (Russell *et al.* 1988). La pérdida de material a la superficie de cualquier especie química está sujeta exclusivamente a la depositación en seco, la cual se trata empleando un método basado en resistencias superficiales (Wesely 1989).



Como se mencionó con anterioridad, el CIT también cuenta con un módulo de análisis directo de sensibilidad que permite calcular de manera muy eficiente coeficientes locales de sensibilidad de primer orden (Yang *et al.* 1997). Los coeficientes de sensibilidad ( $s_{ij}$ ) de las concentraciones arrojadas por el modelo ( $c_i$ ) a cambios en las entradas o parámetros del mismo ( $p_j$ ), es decir  $s_{ij} = \partial c_i / \partial p_j$ , se calculan al obtener las derivadas de las ecuaciones que gobiernan la evolución de la concentración de contaminantes en el aire (es decir, la ecuación atmosférica de continuidad para cada especie química), y resolviendo numéricamente dichas ecuaciones de sensibilidad. Cabe mencionar que los coeficientes obtenidos no sólo incluyen el efecto debido al transporte de especies, sino también a los efectos de reacción química en las especies químicamente reactivas. Asimismo, los coeficientes de sensibilidad brindan información del cambio de concentración en una celda del dominio debido a cambios en entradas o parámetros en todo el dominio o subdominio espacial que se esté analizando. Debido a que las magnitudes de las entradas al CIT, así como sus parámetros de entrada difieren en rangos de magnitud muy variados, los coeficientes de sensibilidad a diferentes entradas o parámetros también difieren en varios órdenes de magnitud, por lo que el módulo numérico del CIT en realidad calcula un coeficiente seminormalizado de sensibilidad. Este coeficiente seminormalizado se define como  $\partial c_i / \partial \varepsilon_j$ , donde  $\varepsilon_j = p_j / p_j^0$  es una variable de escalamiento con un valor nominal de uno;  $p_j$  es el valor escalado del parámetro o la entrada al modelo; y  $p_j^0$  es el valor nominal del parámetro o la entrada al modelo. Los coeficientes de sensibilidad tienen unidades de ppmv por cada por ciento de incremento en el parámetro  $p_j$ .

**Modelo inverso.** Una vez que el CIT resuelve la ecuación atmosférica de continuidad para especies químicas y, al mismo tiempo, calcula los coeficientes de sensibilidad de cada especie química a diferentes entradas predefinidas del modelo, como pudieran ser las fuentes de emisión (p. ej., cambios en la concentración de  $O_3$  al modificar el nivel de emisiones de  $NO_x$ ), se procede a emplear el módulo de modelación inversa desarrollado por Mendoza-Domínguez y Russell (2000). Dicho módulo permite calcular factores de ajuste, en este caso a las tasas de emisión, los cuales, una vez aplicados al inventario de emisiones, permitirían obtener un mejor desempeño del modelo directo. Esto se hace minimizando la diferencia entre los valores observados y los datos arrojados por el modelo, utilizando información adicional de los monitoreos y, de conocerse, las propiedades de

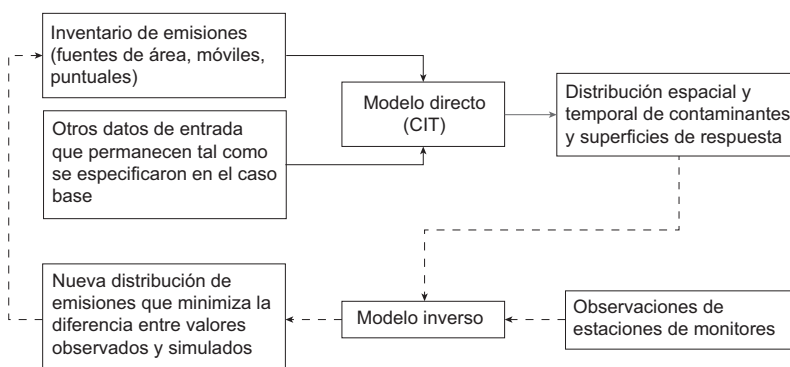
las emisiones. Matemáticamente, el modelo inverso tiene la forma:

$$m = (G^T W_e G + W_m)^{-1} G^T W_e d \quad (1)$$

donde:  $d$  es un vector con la diferencia entre los valores de concentración simulados y observados para cada especie química de interés;  $m$  es un vector con los escalamientos sugeridos a las tasas de emisión de las diferentes fuentes analizadas;  $G$  es una matriz que contiene los coeficientes de sensibilidad de cada especie química a todos los parámetros o entradas de interés (fuentes de emisión en este caso);  $W_e$  es una matriz diagonal con pesos que incorporan información adicional conocida de los datos monitoreados (por ejemplo, incertidumbre de cada valor monitoreado), y  $W_m$  es una matriz diagonal que incorpora información conocida sobre las propiedades de las emisiones.  $W_m$  trabaja como una función de penalización a los cambios sugeridos a las emisiones, de tal manera que restringe los valores que pueda tomar  $m$  y, a su vez, estabiliza la solución del modelo inverso, el cual intrínsecamente proviene de un problema numérico mal condicionado. Una vez resuelto este sistema lineal, el factor de escalamiento obtenido para cada fuente de emisión se aplica a las emisiones a la atmósfera alimentadas al CIT y se vuelve a ejecutar el modelo directo. El proceso completo se repite hasta obtener una respuesta estable del sistema (**Fig. 2**). Este proceso iterativo es necesario ya que los fenómenos atmosféricos, incluyendo las reacciones químicas atmosféricas, son no lineales en su respuesta.

El esquema de modelación inversa planteado puede realizarse de tal manera que se obtengan factores de escalamiento a las emisiones de una especie química al nivel de todo el dominio o por subdominios. Así mismo, se puede obtener un solo factor de escalamiento para las 24 horas que conforman un día en particular, o bien se pueden obtener factores de forma horaria. Existe una tercera opción, la cual es por episodio, lo cual arrojaría un solo factor de escalamiento para todos los días analizados de una simulación dada. Para el caso del presente trabajo, se obtuvieron factores de escalamiento horarios. La resolución espacial de las observaciones limitó el análisis por subdominios, por lo que se optó por calcular factores de escalamiento aplicables a todo el dominio. Con esto se hace la suposición de que la distribución espacial de emisiones es aproximadamente correcta.

Cabe hacer notar que en este proceso de modelación inversa se hace la suposición de que las



**Fig. 2.** Esquema del proceso de obtención de factores de escalamiento del inventario de emisiones del caso base, de acuerdo al acoplamiento de un modelo directo y un módulo de modelación inversa

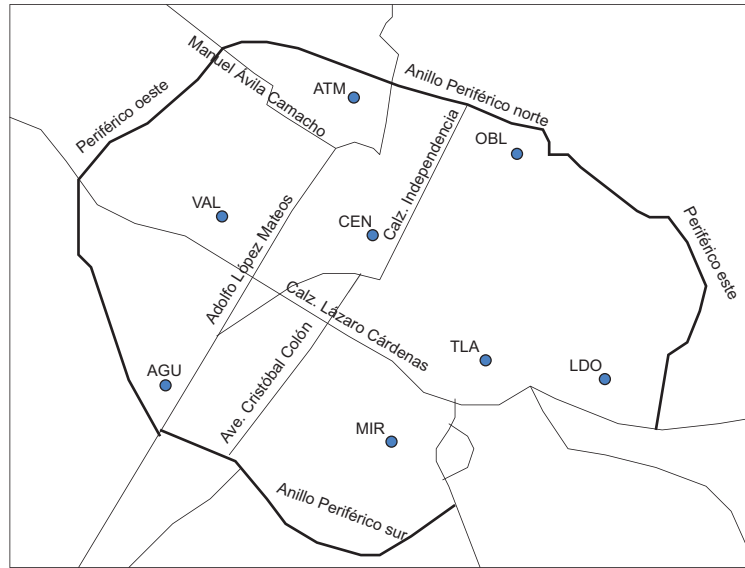
emisiones son la única fuente de incertidumbre en las entradas del CIT y que el modelo es “perfecto”. Mendoza-Domínguez y Russell (2001) aplicaron el mismo esquema de modelación a un dominio de dimensiones similares al empleado aquí, pero para otra cuenca atmosférica. En dicha aplicación, concluyeron que la probabilidad de que incertidumbres en condiciones iniciales y de frontera, o en campos de cobertura de nubes, pudieran estar contribuyendo tanto como las incertidumbres en las emisiones, era baja. Otros estudios han establecido, de manera similar, que, en aplicaciones de modelos de calidad del aire, los inventarios de emisiones son la mayor fuente de incertidumbre (Placet *et al.* 2000, Sawyer *et al.* 2000). Aún así, es claro que el modelo en realidad no es perfecto y que todas las entradas y parámetros involucran cierto grado de incertidumbre. El modelo inverso ajusta los parámetros a la mejor solución posible sin necesariamente considerar la plausibilidad física de la solución. Esto se tiene que considerar en la interpretación de los resultados obtenidos.

### APLICACIÓN DEL SISTEMA DE MODELACIÓN A LA ZMG

El caso base para la aplicación aquí presentada se tomó del trabajo desarrollado por Mendoza y García (2009). Se empleó la misma configuración espacial del modelo. Asimismo, se utilizaron los mismos campos meteorológicos, información de la distribución del uso de suelo, inventario de emisiones y condiciones iniciales y de borde. Para el caso particular de la construcción de los campos meteorológicos (viento, temperatura y humedad), se empleó un modelo de diagnóstico (Goodin *et al.* 1979, 1980) que se alimentó con información registrada por las estaciones meteorológicas ubicadas dentro del dominio de modela-

ción, incluyendo la derivada por las estaciones de la Red Automática de Monitoreo Atmosférico (RAMA) de la ZMG. La altura de la capa de mezcla se obtuvo a partir de radiosondeos realizados dos veces al día por el Observatorio Colomos de la CNA. La altura de la capa de mezcla se consideró espacialmente homogénea en todo el dominio, pero variable hora a hora. La distribución horaria se estimó con base en el perfil horario promedio de julio de 1995, tomando como referencia los datos observados. Al igual que en la aplicación anterior, los dos primeros días de ejecución del CIT se tomaron como un período de estabilización del modelo y tiempo necesarios para minimizar la influencia de las condiciones iniciales en el resto de la simulación.

Los datos de concentración de contaminantes atmosféricos, requeridos para llevar a cabo el proceso de modelación inversa, se obtuvieron de la RAMA de la ZMG (Fig. 3). Particularmente, se usaron datos reportados de los niveles de CO, NO, NO<sub>2</sub>, O<sub>3</sub> y SO<sub>2</sub>. En la base de datos original de observaciones, también se tenía disponible información sobre niveles observados de hidrocarburos no metánicos totales. Sin embargo, al no tener una forma directa de relacionar dichas observaciones con el subconjunto de compuestos orgánicos volátiles (COV) empleados por el modelo en su mecanismo fotoquímico y que están presentes en el inventario de emisiones, se optó por no usar dicha información. Es decir, no se intentó encontrar factores de escalamiento a las emisiones de COV contenidos en el inventario alimentado al modelo. Así, la matriz *G* se construyó con los coeficientes de sensibilidad calculados para cambios en las concentraciones de CO, NO, NO<sub>2</sub>, O<sub>3</sub> y SO<sub>2</sub> a cambios en las emisiones de CO, SO<sub>2</sub> y NO<sub>x</sub>, para todo el dominio de modelación, de tal manera que *m* es un vector que reporta los factores de escalamiento sugeridos para las mismas fuentes de emisiones: CO, SO<sub>2</sub> y NO<sub>x</sub>.



**Fig. 3.** Localización de las estaciones de monitoreo (•) que conforman la Red Automática de Monitoreo Atmosférico de la ZMG. Zona Norte: ATM (Atemajac), OBL (Oblatos); Zona Poniente: AGU (Águilas), VAL (Vallarta); Zona Centro: CEN (estación Centro); Zona Sur: MIR (Miravalle); Zona Oriente: LDO (Loma Dorada), TLA (Tlaquepaque)

La matriz  $W_e$  se calculó de la siguiente manera (Mendoza-Domínguez y Russell 2000):

$$\omega_{ik} = \frac{1}{N_i} \cdot \frac{1}{\sigma_{ik}^2} \cdot \omega_{ik}^* \quad (2)$$

En esta notación,  $\omega_{ik}$  representa los elementos diagonales de  $W_e$  siendo los primeros  $k$  elementos las  $k$  mediciones (promedio de una hora) que corresponden a la especie  $i = 1$ ; los siguientes  $k$  elementos son las mediciones de la segunda especie, y así sucesivamente.  $N_i$  es el número total de mediciones válidas para la especie  $i$ ;  $\sigma_{ik}$  es la desviación estándar de la  $k$ -ésima observación de la especie  $i$ ; y  $\omega_{ik}^*$  son pesos adicionales que se usan durante las primeras iteraciones para acelerar convergencia y sus definiciones se pueden encontrar reportadas en la literatura (Mendoza-Domínguez y Russell 2000). Debido a que no se tiene un valor de  $\sigma_{ik}$  para cada observación, se usaron valores promedio de coeficientes de variación para estimar las incertidumbres correspondientes:

$$\sigma_{ik} = CV_i \times O_{ik} \quad (3)$$

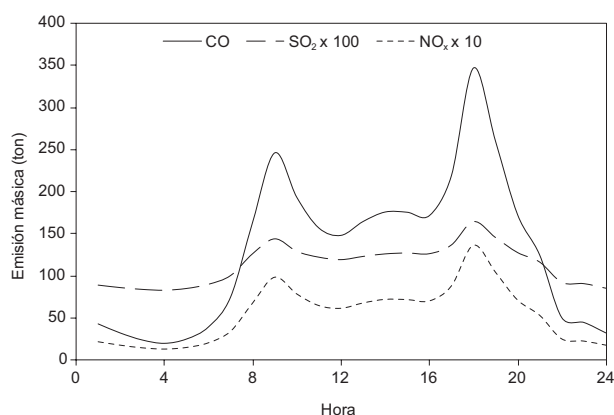
donde  $CV_i$  es el coeficiente de variación de la especie  $i$ , y  $O_i$  es la  $k$ -ésima observación de la especie  $i$ . Los valores de  $CV_i$  empleados fueron:  $O_3$  5 %,  $NO$  10 %,  $NO_2$  20 %,  $SO_2$  15 % y  $CO$  15 % (NRC 1991).

Finalmente, la matriz  $W_m$  se estimó siguiendo el

proceso iterativo sugerido por Mendoza-Domínguez y Russell (2000). En primera instancia, se establece que  $W_m = \lambda I$ , donde  $\lambda$  es una constante positiva estimada por argumentos estocásticos (Hoerl y Kennard 1976, Aldrin 1997) e  $I$  es la matriz identidad. En caso de que ciertos valores resultantes del vector  $m$  ( $m_k$ ) excedan un cierto límite fijado *a priori*, valores específicos de la diagonal de  $W_m$  se recalculan iterativamente para “penalizar” el cambio sugerido a los valores  $m_k$  (positivo o negativo). En el caso de la aplicación presentada aquí, se dio una alta flexibilidad en los cambios que pudieran estimarse para los tres diferentes conjuntos de fuentes de emisión.

## RESULTADOS

Las emisiones reportadas en el inventario original de la ZMG (1995) son: 2460 ton/día de  $CO$ , 102 ton/día de  $NO_x$  22 ton/día de  $SO_2$ , y 394 ton/día de hidrocarburos (HC) (Gobierno del Estado de Jalisco *et al.* 1997). Con los ajustes indicados para escalar el inventario al año de simulación (2001) y extenderlo a cubrir todo el dominio de modelación, el inventario del caso base resulta en 3111 ton/día de  $CO$ , 130 ton/día de  $NO_x$ , 28 ton/día de  $SO_x$  ( $SO_2 + SO_3$ ) y 454 ton/día de HC. El perfil de emisiones másicas de  $CO$ ,  $SO_2$  y  $NO_x$  para el último día del caso base se presenta en la **Fig. 4**, en donde se reporta la suma de emisiones para todo el dominio para cada hora.



**Fig. 4.** Perfil de emisión másica para el dominio de modelación (18 de mayo de 2001)

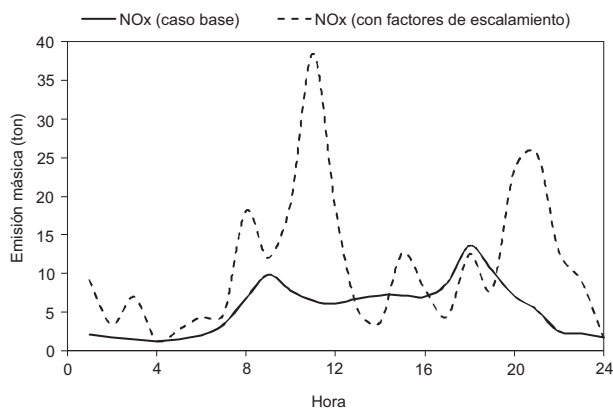
Cabe resaltar que, dada la forma en que se estimó el inventario de emisiones usado en este trabajo, el porcentaje de contribución de las diversas fuentes al inventario global se mantuvo igual respecto al inventario de 1995 (esta es una posible fuente de error y pudiera ser investigada en un futuro a través de pruebas de sensibilidad). De manera particular, el 99.7 % de las emisiones de CO, el 91.0 % de las emisiones de NO<sub>x</sub> y el 30.4 % de las emisiones de SO<sub>x</sub> se estima provienen de las fuentes móviles. Así, los factores de ajuste obtenidos para las emisiones de CO y NO<sub>x</sub> tras la aplicación del modelo inverso sugerirían cambios predominantemente para las fuentes móviles (considerando que en efecto la contribución que se establece en el inventario original sea la correcta). Esto es relevante ya que otros estudios han encontrado que la variabilidad diaria de las emisiones de las fuentes móviles en centros urbanos puede llegar a ser importante e incluso es de esperarse que la proporción emitida de algunos contaminantes (p. ej., CO vs. NO<sub>x</sub>) no sea necesariamente constante (Niemeier 2003, Harley *et al.* 2005, Parrish 2006). Esta variabilidad en las emisiones de las fuentes móviles puede tener su origen en diversos factores. Por ejemplo, Hao *et al.* (2000) indican que las emisiones de CO son más sensibles a variaciones en las velocidades de tránsito y los modos de manejo que las emisiones de NO<sub>x</sub>. Zhang *et al.* (1995) mencionan que la presencia de vehículos altamente contaminantes (*high emitters*) puede contribuir a cerca del 50 % del total de emisiones de CO, por lo que su presencia o ausencia en la red vial pudiera provocar variaciones significativas en las emisiones totales. Por otro lado, congestionamientos locales pueden incrementar de un 50 a un 100 % las emisiones de CO, mientras que las emisiones de NO<sub>x</sub> son altamente influenciadas por la presencia de vehículos pesados a diesel (Corsmeier

*et al.* 2005). En resumen, variaciones en los flujos espaciales y temporales de vehículos, composición de la flota vehicular, así como en los modos de operación y manejo, que ocurren cotidianamente en una metrópoli traen consigo cambios diarios en las emisiones totales. Evidentemente es difícil modelar dichas variaciones y es precisamente en estos casos donde las técnicas de modelación inversa pueden ser útiles para identificar posibles ajustes a los inventarios de emisiones obtenidos a través de los modelos de emisiones.

Una vez que se aplicó el proceso de modelación inversa y se sumaron los cambios horarios al inventario de emisiones base, se estimó que las emisiones de CO deberían reducirse de manera global en un 16 % el primer día (mayo 16), aumentarse en un 60 % el segundo, y aumentar en un 64% el tercer día. Para el caso de los NO<sub>x</sub>, todos los días se reporta un aumento global: 152 % el primer día, 146 % el segundo y 104 % el tercero. Finalmente, para el caso del SO<sub>2</sub>, el aumento sugerido es de 72 % el primer día, 20 % el segundo y 137 % el tercero. Considerando que el mayor emisor de CO y NO<sub>x</sub> son las fuentes móviles, los resultados del segundo y tercer día indicarían cierto nivel de congruencia, es decir, un aumento en las emisiones de CO va acompañado de un aumento en las emisiones de NO<sub>x</sub>. Sin embargo, es de esperarse que los incrementos no necesariamente debieran ser iguales tal como se mencionó arriba. Además, hay otras fuentes de emisión presentes en la zona que también contribuyen y no necesariamente fueron modeladas correctamente en el inventario base. Sin embargo, para el primer día de modelación, los resultados se alejan de lo encontrado en los dos últimos días. Esto pudiera ser debido a un efecto residual de las condiciones iniciales que aún pudieran estar afectando los niveles de los contaminantes simulados, o bien a una limitación del modelo inverso. Esto se discute más adelante.

A manera de ejemplo de los resultados obtenidos tras la aplicación del modelo inverso, la **Fig. 5** ilustra el perfil de emisión base de NO<sub>x</sub> para el tercer día de simulación y el perfil obtenido una vez aplicados los factores de escalamiento al inventario sugeridos por la modelación inversa. Para este caso, se puede apreciar que el modelo inverso sugiere un perfil con mayores variaciones temporales que el original, pero aun así, sigue una tendencia acorde a lo que se esperaría fueran las horas de mayor y menor emisión (p. ej., horas pico de circulación vehicular). De manera general, se obtuvieron resultados similares para el resto de los días y especies emitidas. Así, la modela-





**Fig. 5.** Perfil de emisión másica de NOx para las simulaciones con el inventario base y el inventario corregido (18 de mayo de 2001)

ción inversa pudiera estar capturando variaciones que no es posible representar con los modelos y técnicas actuales de estimación de emisiones y globalmente estaría indicando que el inventario de emisiones base está subestimado, con excepción de las emisiones de CO del primer día (mayo 16). No obstante, una limitación de esta técnica de modelación es que no puede identificar el origen de las variaciones observadas. Por ejemplo, en la **Fig. 5** se aprecia una mayor cantidad de emisiones alrededor de las 21 h que durante las 8 h. Esto lo único que llega a sugerir es que la mezcla de modos de operación, composición de la flota, nivel de actividad de las fuentes de emisión, etc., es diferente en ambas horas. Una posible respuesta pudiera ser que hay más vehículos pesados diesel circulando a las 21 h, cuya influencia *per capita*

en las emisiones de NOx es mayor a los vehículos a gasolina, pero eso (y otras posibles explicaciones) es meramente una hipótesis a este nivel de análisis. Identificar dichas razones está fuera del alcance de este trabajo, sin embargo los resultados obtenidos pueden dar pie a explorar dichas opciones.

El modelo inverso trabaja bajo un esquema de minimización de la diferencia entre los valores observados y simulados (en términos de una norma  $L_2$ ), quedando un error residual que el modelo no puede explicar. Además, por la misma forma de la ecuación 1, el estimador de  $m$  es un estimador sesgado. Así, no es de esperarse que la concordancia entre los valores observados y simulados sea total tras la aplicación del modelo inverso, pero sí es de esperarse un mejor desempeño del modelo de calidad del aire después de aplicar los factores de escalamiento sugeridos por el modelo inverso. Los **cuadros I, II y III** muestran los estadísticos de desempeño del CIT calculados para cada día para cada una de las especies químicas cuyas emisiones fueron ajustadas directamente por el proceso de modelación inversa: CO, NOx y SO<sub>2</sub>. Los estadísticos presentados son los recomendados para evaluaciones de desempeño de modelos de calidad del aire respecto a O<sub>3</sub> (Doll *et al.* 1991); no existen guías acordadas que indiquen un desempeño aceptable o no del modelo para las concentraciones de CO, NOx y SO<sub>2</sub>. No obstante, los valores presentados en los **cuadros I y II** indican que, de manera general, el CIT está describiendo mejor la dinámica de CO y NOx, mientras que para el SO<sub>2</sub> los resultados son mixtos. Para CO y NOx, el sesgo normalizado dis-

**CUADRO I.** EVALUACIÓN ESTADÍSTICA DEL DESEMPEÑO DEL CIT PARA CO (CB: CASO BASE, MI: MODELACIÓN INVERSA)<sup>a</sup>

	16 de mayo		17 de mayo		18 de mayo	
	CB	MI	CB	MI	CB	MI
Sesgo normalizado, %	30.2	-15.9	-34.1	-17.3	-26.8	-15.6
Error normalizado, %	65.6	35.8	58.6	48.5	60.6	46.2
ECMN	0.89	0.37	0.88	0.41	0.87	0.44
RCECM (ppmv)	2.02	1.08	1.39	1.08	1.33	1.05
RCECM <sub>s</sub> (ppmv)	0.64	0.62	1.10	0.66	1.01	0.62
RCECM <sub>n</sub> (ppmv)	1.91	0.88	0.86	0.86	0.86	0.84
IDA	0.48	0.75	0.55	0.72	0.60	0.76

<sup>a</sup> Los estadísticos se calculan tomando como base el residual  $r_i = P_i - O_i$ , donde  $O_i$  y  $P_i$  son la  $i$ -ésima concentración observada y simulada, respectivamente. El sesgo normalizado es  $1/N \sum (r_i/O_i)$ , donde  $N$  es el número de pares válidos que dan origen a un valor de  $r_i$  y la sumatoria corre de  $i=1$  a  $N$ . De manera similar, el error normalizado es  $1/N \sum [r_i/O_i]$ , el error cuadrado medio normalizado (ECMN) es  $1/N \sum (r_i/O_i)^2$  y la raíz cuadrada del error cuadrado medio (RCECM) es  $[1/N \sum (r_i)^2]^{1/2}$ . La contribución sistemática al RCECM (esto es, RCECM<sub>s</sub>) se obtiene a partir de  $[1/N \sum (\bar{r}_i)^2]^{1/2}$ , donde  $\bar{r}_i = P_i - O_i$ , y  $P_i = a + bO_i$  ( $a$  y  $b$  son coeficientes de una regresión lineal). Para el RCECM se cumple que  $RCECM^2 = RCECM_s^2 + RCECM_n^2$ , donde RCECM<sub>n</sub> es la contribución no sistemática. Finalmente, el índice diario de ajuste (IDA) es  $1 - [N RCECM^2 / \sum (P_i - M_o + |O_i - M_o|)^2]$ , donde  $M_o$  es el valor observado promedio dado por  $1/N \sum O_i$ . Note que un sesgo positivo indica que los valores simulados son mayores a los observados

**CUADRO II.** EVALUACIÓN ESTADÍSTICA DEL DESEMPEÑO DEL CIT PARA NO<sub>x</sub> (CB: CASO BASE, MI: MODELACIÓN INVERSA)<sup>a</sup>

	16 de mayo		17 de mayo		18 de mayo	
	CB	MI	CB	MI	CB	MI
Sesgo normalizado, %	-28.1	68.0	-61.2	8.4	-51.1	-13.2
Error normalizado, %	77.8	121.7	70.4	78.6	67.0	61.9
ECMN	2.05	1.20	2.17	0.71	2.05	0.83
<i>RCECM</i> (ppbv)	54.6	60.1	42.0	38.7	43.1	36.7
<i>RCECM<sub>s</sub></i> (ppbv)	40.9	33.8	39.4	18.5	39.7	26.1
<i>RCECM<sub>n</sub></i> (ppbv)	36.2	49.7	14.6	34.0	16.9	25.8
IDA	0.39	0.33	0.47	0.56	0.47	0.59

<sup>a</sup>Ver el **cuadro I** para definiciones de los estadísticos

**CUADRO III.** EVALUACIÓN ESTADÍSTICA DEL DESEMPEÑO DEL CIT PARA SO<sub>2</sub> (CB: CASO BASE, MI: MODELACIÓN INVERSA)<sup>a</sup>

	16 de mayo		17 de mayo		18 de mayo	
	CB	MI	CB	MI	CB	MI
Sesgo normalizado, %	50.7	55.6	-10.2	-12.8	-13.1	10.1
Error normalizado, %	105.4	115.2	72.6	69.5	66.9	79.9
ECMN	1.24	1.20	1.22	1.26	1.99	1.43
<i>RCECM</i> (ppbv)	8.1	7.9	5.7	5.6	8.1	7.7
<i>RCECM<sub>s</sub></i> (ppbv)	5.8	5.6	4.8	4.9	7.3	6.4
<i>RCECM<sub>n</sub></i> (ppbv)	5.6	5.6	3.0	2.8	3.5	4.2
IDA	0.38	0.42	0.44	0.45	0.42	0.48

<sup>a</sup> Ver el cuadro I para definiciones de los estadísticos

minuyó en todos los casos, con excepción del 16 de mayo en donde aumentó para los NO<sub>x</sub>. Así mismo, el ECMN disminuyó para ambos contaminantes en los tres días del episodio; el ECMN tiende a ser un mejor indicador que el error normalizado como una métrica del buen desempeño espacial y temporal del modelo (Hanna 1988). En lo general, la *RCECM* se redujo y su componente no sistemático de la *RCECM* pasó a tener mayor peso en su cálculo. Esto último da una indicación de que el error residual está conformado de manera importante por variaciones (ruido) que están fuera del alcance de la resolución del modelo directo. Finalmente, el IDA, que se puede tomar como un indicador integral de desempeño, mejoró en casi todos los casos; la excepción nuevamente fueron los NO<sub>x</sub> durante el 16 de mayo. La evaluación de desempeño para los NO<sub>x</sub>, para el primer día, muestra que el modelo está teniendo problemas en reconciliar los valores observados y los simulados: se tiene un sesgo normalizado de +68 % (contra el original de -28 %) y un IDA que baja; sin embargo, se tiene una mejora en el ECMN y en el componente sistemático de la *RCECM*. El impacto de esto es que pudiera existir una aparente discrepancia entre los ajustes sugeridos para las emisiones de NO<sub>x</sub> entre el primer día y los

últimos dos días de simulación, que a su vez tiene que ver con lo que se observa para el CO durante el primer día (incremento en emisiones sugeridas para los NO<sub>x</sub> y decremento en las emisiones sugeridas para el CO –para el primer día– siendo que en principio el mayor contribuyentes a estas emisiones son las fuentes móviles). Por un lado, pudiera tenerse una influencia de las condiciones iniciales que estaría causando esta aparente discrepancia en los resultados, mientras que por otro lado los resultados mismos pudieran sugerir una limitación del modelo inverso, que al no incorporar explícitamente otras restricciones físicas, está realizando un ajuste “a ciegas”. Todo ello es relevante para poder atribuirle el justo valor a los resultados obtenidos, particularmente para el primer día de análisis. Para el caso del SO<sub>2</sub>, los resultados presentados en el **cuadro III** muestran una mejora en el ECMN, IDA y la *RCECMs* durante el primer y tercer día; pero solamente una mejora marginal en el IDA para el segundo día.

Si bien se observa una mejoría en el desempeño del modelo para describir las concentraciones de CO, NO<sub>x</sub> y, marginalmente, SO<sub>2</sub>, es de interés analizar el efecto de los cambios en emisiones sobre un contaminante secundario como lo es el O<sub>3</sub>. Si bien en el

**CUADRO IV.** EVALUACIÓN ESTADÍSTICA DEL DESEMPEÑO DEL CIT PARA O<sub>3</sub> (CB: CASO BASE, MI: MODELACIÓN INVERSA)<sup>a,b</sup>

	16 de mayo		17 de mayo		18 de mayo	
	CB	MI	CB	MI	CB	MI
EEPCD, %	28.7	5.9	-7.1	-2.5	16.4	-10.2
Sesgo normalizado, %	55.1	7.2	10.2	4.3	20.3	0.1
Error normalizado, %	66.1	32.1	19.3	20.1	36.5	15.6
ECMN	0.34	0.13	0.07	0.05	0.19	0.04
RCECM (ppbv)	59.0	34.6	34.8	27.2	31.1	21.9
RCECM <sub>s</sub> (ppbv)	47.7	22.1	30.7	18.5	20.7	15.1
RCECM <sub>n</sub> (ppbv)	34.8	26.6	16.4	19.9	23.2	15.8
IDA	0.61	0.88	0.86	0.93	0.89	0.95

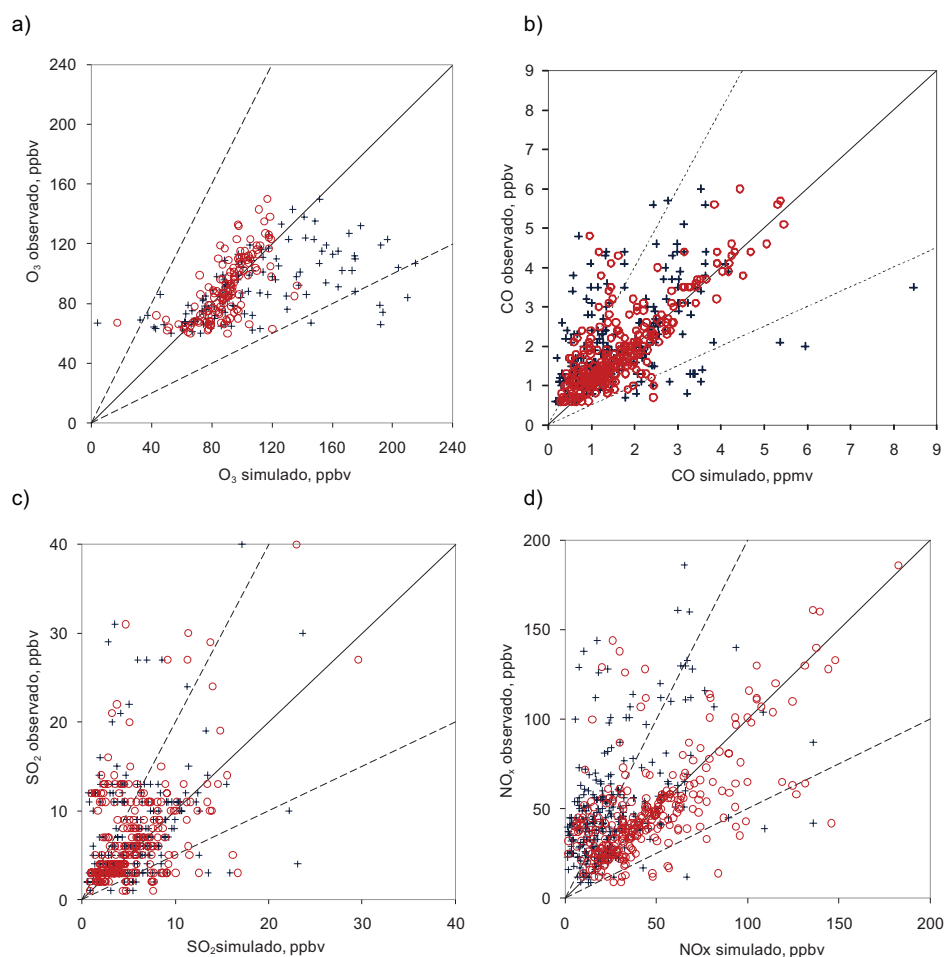
<sup>a</sup> Ver el **cuadro I** para definiciones de los estadísticos; además, EEPCD es el error en la estimación del pico de concentración diario. <sup>b</sup> Valor de corte para el ozono: 60 ppbv (Doll *et al.* 1991, Harley *et al.* 1993)

proceso de modelación inversa se usa información de cómo cambia la concentración de O<sub>3</sub> con respecto a cambios en cada una de las fuentes de emisión usadas, típicamente los factores de escalamiento son gobernados por las especies primarias: el coeficiente de sensibilidad del O<sub>3</sub> con respecto a un cambio en las emisiones de CO tiene menos preponderancia que el coeficiente de sensibilidad del CO con respecto a un cambio en las emisiones de CO. En el **cuadro IV** se reportan los estadísticos de desempeño para el O<sub>3</sub>. En este caso, los valores guía son un sesgo normalizado inferior a  $\pm 15\%$ , un error normalizado menor al 35 % y un error en la predicción del pico máximo de ozono inferior a  $\pm 20\%$  (Doll *et al.* 1991). Los resultados obtenidos indican que el desempeño del CIT mejoró sustancialmente en todas las métricas estadísticas calculadas (solamente durante el 17 de mayo se tuvo un ligero incremento en el error normalizado), con lo que se pudiera establecer que, con el inventario de emisiones ajustado, se podría estar mejorando la descripción de la dinámica de las especies primarias, lo cual a su vez también mejora, como sería de esperarse, la descripción de la dinámica de las especies secundarias.

La **Fig. 6** ilustra la mejoría observada en el desempeño del CIT respecto a las concentraciones simuladas de CO, NO<sub>x</sub> y O<sub>3</sub> tras la aplicación del inventario de emisiones ajustado: las parejas de observación/predicción tienden a compactarse sobre el área acotada por las líneas 1:2 y 2:1, y alineadas sobre la línea 1:1. Sin embargo, la dispersión en el comportamiento de las parejas observación/predicción para el SO<sub>2</sub> es notoria. Cabe mencionar que las gráficas para CO, NO<sub>x</sub> y SO<sub>2</sub> fueron construidas con la mitad de las parejas de datos disponibles con el fin de no sobrecargar las imágenes; los datos que se tomaron fueron para las horas pares. De particular interés es

el hecho de que el O<sub>3</sub> pasó de estar sobreestimado en el caso base a estar ligeramente subestimado en el caso con el inventario corregido. Es conocido que los niveles de ozono están gobernados por la interacción entre los NO<sub>x</sub> y los HC (en presencia de radiación solar). Dado que el valor de las emisiones de HC no fue modificado aquí, se estaría teniendo la respuesta del O<sub>3</sub> exclusivamente a las modificaciones en los NO<sub>x</sub>. El error remanente en el O<sub>3</sub> pudiera ser función de errores remanentes en las emisiones de HC. Aun así, la respuesta obtenida es de importancia ya que estaría indicando que la atmósfera en la ZMG tiene una condición de inhibición por NO<sub>x</sub>: al aumentar los NO<sub>x</sub>, disminuye el O<sub>3</sub>, y a la inversa, si disminuyen los NO<sub>x</sub>, aumenta el O<sub>3</sub>. Esta información es relevante para el establecimiento de estrategias de control de emisiones que pudieran ser efectivas en la zona.

Los factores de escalamiento al inventario de emisiones se estimaron bajo la suposición de que la distribución espacial de las mismas era aproximadamente correcta. Así, las emisiones de todo el dominio se escalaban hora a hora por el valor sugerido por la modelación inversa. Esta decisión se tomó con base en el número limitado de estaciones de monitoreo disponibles en la región de modelación. Una forma de establecer qué tan válida es la anterior premisa, es estimar el desempeño del CIT por zonas. El **cuadro V** presenta los coeficientes de correlación lineal ( $r^2$ ) entre valores observados y simulados para las cinco zonas en que fue dividida la presencia de estaciones de monitoreo. Tanto para el CO, NO<sub>x</sub> y O<sub>3</sub>, el desempeño mejoró en todas las zonas respecto al caso base. Para el CO, el peor desempeño se obtuvo para la zona Norte, mientras que para el NO<sub>x</sub> y el O<sub>3</sub> el peor desempeño se obtuvo para las zonas Norte y Oriente. Cabe notar que el desempeño por zonas entre NO<sub>x</sub> y O<sub>3</sub> correlaciona bastante bien. Respecto



**Fig. 6.** Gráficas de dispersión para parejas de valores de concentración observados y simulados: a) O<sub>3</sub>, b) CO, c) SO<sub>2</sub>, y d) NO<sub>x</sub>. (+) Simulación con el inventario base; (o) simulación con emisiones corregidas. Las líneas guía representan las proporciones 2:1, 1:1 y 1:2.

**CUADRO V.** COEFICIENTES DE CORRELACIÓN ( $r^2$ ) ENTRE VALORES OBSERVADOS Y SIMULADOS POR ZONA PARA EL CASO BASE (CB) Y LA MODELACIÓN INVERSA (MI).<sup>a</sup>

Zona	O <sub>3</sub>		CO		NO <sub>x</sub>		SO <sub>2</sub>	
	CB	MI	CB	MI	CB	MI	CB	MI
Norte	0.45	0.68	0.03	0.23	0.02	0.09	0.11	0.06
Poniente	0.73	0.82	0.41	0.57	0.32	0.44	0.29	0.26
Centro	0.50	0.89	0.21	0.76	0.31	0.67	-	-
Sur	0.70	0.86	0.26	0.95	0.38	0.33	0.09	0.07
Oriente	0.25	0.70	0.50	0.89	0.12	0.29	0.22	0.56

<sup>a</sup> Ver Fig. 3 para definición de las estaciones del sistema de monitoreo ambiental que pertenecen a cada zona. En todos los casos se emplean todas las parejas de datos (observados vs. simulados) válidos.

al caso del SO<sub>2</sub>, el desempeño del modelo empeoró tras la aplicación de las correcciones al inventario de emisiones en todas las zonas excepto la Oriente, en donde la mejoría fue notable; en el resto de las estaciones, el desempeño es muy pobre. Un análisis más profundo de los datos de monitoreo indica que

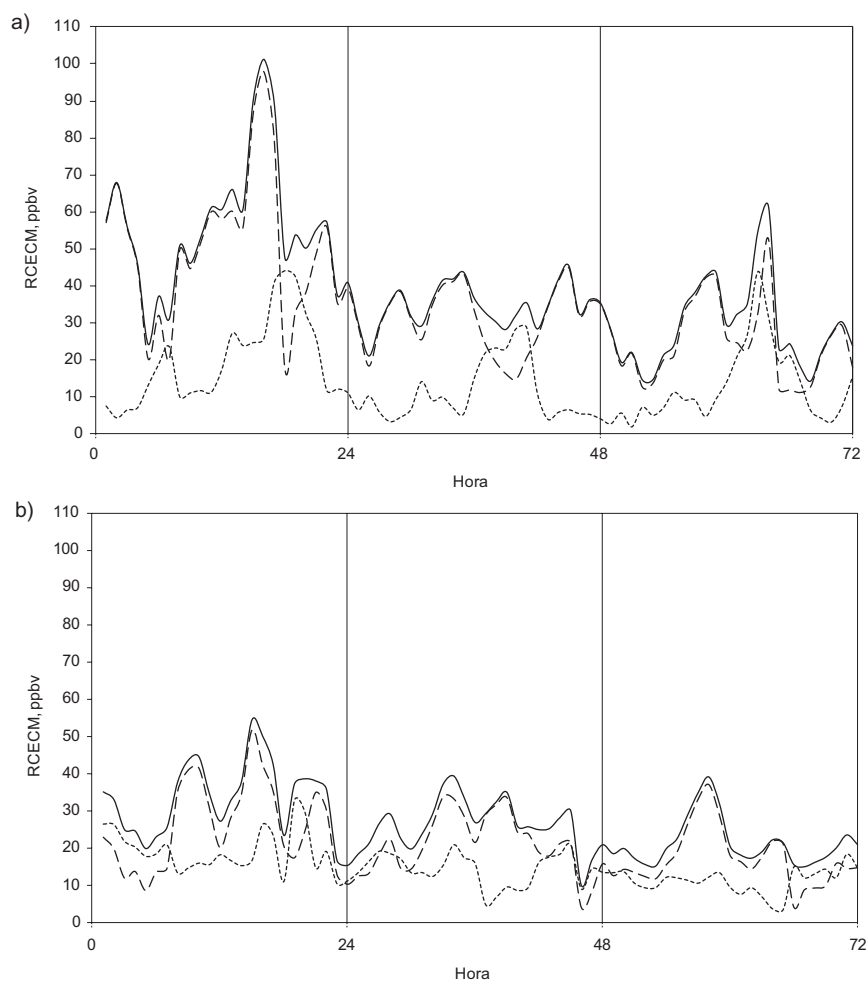
las máximas concentraciones de SO<sub>2</sub> se reportaron en la zona Oriente y estarían gobernando el proceso de asimilación: el modelo inverso estaría tratando de conciliar las diferencias presentadas en las estaciones de la zona Oriente en detrimento del resto de las estaciones. Con el fin de indagar las razones por



las cuales el desempeño respecto al  $\text{SO}_2$  es tan dispar respecto a  $\text{CO}$ ,  $\text{NO}_x$  y  $\text{O}_3$  se llevó a cabo una prueba de la homogeneidad en los campos de concentración observados. A tal efecto, se llevó a cabo un proceso de estimación de concentraciones en los puntos de observación de las estaciones de monitoreo a través de un proceso de retención de datos (McNair *et al.* 1996). Sistemáticamente, se obtuvieron concentraciones estimadas en cada una de las ocho estaciones de monitoreo empleando una técnica de interpolación espacial de pesos calculados con base en el cuadrado del inverso de la distancia entre estaciones (Goodin *et al.* 1979) en donde los datos de siete estaciones eran empleados para estimar la concentración de la octava. Posteriormente, los valores estimados se compararon con los datos reales observados y se estimaron las  $r^2$  correspondientes. Los resultados fueron como sigue:  $\text{CO}$  0.32,  $\text{NO}$  0.42,  $\text{SO}_2$  0.02,  $\text{NO}_2$  0.05 y  $\text{O}_3$  0.60.

Estos resultados dan una indicación de que el campo de concentración de  $\text{O}_3$  es el que tiende a ser más homogéneo y, entre contaminantes primarios, el  $\text{CO}$  tendería a ser más homogéneo que el  $\text{SO}_2$ . Así, podría esperarse que variaciones en el posicionamiento espacial de las emisiones de  $\text{CO}$  tuvieran un menor impacto en el error de las concentraciones simuladas relativo al posicionamiento de las emisiones de  $\text{SO}_2$ . Por ende, habría que detallar mejor la correcta ubicación espacial de las emisiones de este último contaminante.

Finalmente, se exploraron las series de tiempo obtenidas para el estadístico  $RCECM$  con la finalidad de evaluar de manera global si el modelo está representando correctamente la dinámica asociada al  $\text{O}_3$  respecto al tiempo. La **Fig. 7** muestra la evolución del  $RCECM$  y sus dos componentes: error sistemático ( $RCECM_s$ ) y no sistemático ( $RCECM_n$ ). Se puede



**Fig. 7.** Series de tiempo para la  $RCECM$ , y sus componentes sistemático y no sistemático: a) simulación con el inventario base, b) simulación con emisiones corregidas. Línea sólida:  $RCECM$  total; línea discontinua:  $RCECM_s$ ; línea punteada:  $RCECM_n$

notar que el *RCECM* disminuye significativamente a todas horas para la modelación con el inventario ajustado respecto al caso base. Para el caso base, los mayores valores del error global tienden a ocurrir durante las horas de actividad fotoquímica (horario diurno), sobre todo durante el primer día de modelación reportado. Por otro lado, en el caso de la simulación con el inventario ajustado, el error global se acota apreciablemente durante el primer día de simulación y en los subsecuentes días el *RCECM* oscila entre 20 y 30 ppbv. La permanencia del componente sistemático (*RCECM<sub>s</sub>*) indica que todavía es posible realizar mejoras a las entradas del modelo con el fin de obtener mejores desempeños.

Como se comentó anteriormente, una suposición importante en este estudio es la afirmación de que se tiene un modelo perfecto y que la única fuente de error en la modelación tiene como origen el inventario de emisiones. Otras fuentes de error evidentemente existen por lo que los resultados se deben de tomar como una parametrización del inventario de emisiones al caso de estudio, y debe de haber reservas sobre indicar que el inventario obtenido es el correcto. Por ejemplo, Grell *et al.* (2004) demostraron que el uso de modelos de calidad del aire “fuera de línea” como el usado aquí (es decir, aquellos que resuelven la ecuación de continuidad para especies químicas desacoplada de la solución de los campos meteorológicos) tienden a tener problemas respecto a su contraparte (modelos “en línea”) al momento de resolver el transporte vertical de contaminantes, sobre todo bajo condiciones con alta convectividad vertical. En el caso presentado el dominio estuvo influenciado por cielos despejados y vientos débiles (el promedio de la intensidad máxima entre 6 y 18 hrs para los tres días del episodio fue de 2.1 m/s), con lo cual se reduce el efecto convectivo vertical. Sin embargo, hay cierto grado de incertidumbre residual no cuantificada que se debe de tener presente. De igual manera, Niemeir (2003) indica que la distribución espacial de las emisiones tiende a ser dinámica; en este estudio se consideró que la distribución seleccionada era la adecuada. Y así como los dos ejemplos anteriores, existen otros fenómenos que por su complejidad típicamente quedan fuera del alcance del sistema de modelación. Entonces, al aplicar el modelo inverso se estaría generando un inventario parametrizado para compensar esos posibles errores adicionales por una deficiente descripción de los fenómenos meteorológicos o la complejidad de la variación del inventario de emisiones, entre otros. Asimismo, se trabaja bajo la suposición que las mediciones

obtenidas de los sistemas de monitoreo de calidad del aire son correctos. Finalmente, los resultados obtenidos en nuestro estudio son válidos para el episodio estudiado, y pudieran interpretarse como una guía para episodios con condiciones similares. Habría que realizar aplicaciones adicionales para estudiar las variaciones que se pudieran tener durante diferentes estaciones del año, ya que es reconocido que hay una variación inherente en las emisiones durante dichas estaciones.

## CONCLUSIONES

Los experimentos numéricos desarrollados en este trabajo indican que, mediante un proceso formal de modelación inversa, es posible encontrar ajustes al inventario de emisiones del caso base que permiten tener un mejor desempeño global del modelo directo. De particular interés fue el hecho de que un ajuste en el inventario de las especies primarias resultó en un mejor desempeño del O<sub>3</sub> (especie secundaria), lo cual indicaría una mejor descripción general de las especies contaminantes en el dominio de modelación. Considerando que el inventario original estima que las fuentes móviles son la principal fuente de emisiones de CO y NO<sub>x</sub>, el inventario de emisiones ajustado para CO y NO<sub>x</sub> para los últimos dos días de simulación aparentaría tener mayor plausibilidad que el obtenido para el primer día, ya que incrementos en las emisiones de una de las especies va acompañado de un incremento en las emisiones de la otra (pese a que las magnitudes no son idénticas). Para el primer día de simulación, los estadísticos de desempeño de NO<sub>x</sub> son pobres, lo cual estaría indicando que el modelo inverso está teniendo problemas en reconciliar la información de valores simulados y observados, lo cual arrojaría una limitación conocida del modelo inverso: al no incorporar explícitamente restricciones físicas, estaría realizando un ajuste del inventario “a ciegas”. No obstante, los resultados obtenidos dan indicación de que, químicamente, la atmósfera de la ZMG tiene una condición de estar inhibida por NO<sub>x</sub>: reducciones en las emisiones de este contaminante causarían un incremento en los niveles de O<sub>3</sub>. Finalmente, y a pesar de la mejora en el desempeño del modelo directo al aplicar los cambios sugeridos por el modelo inverso en las emisiones, un análisis espacial de los resultados indicó que las emisiones de las zonas Norte y Oriente requieren revisión adicional. Asimismo, las emisiones de SO<sub>2</sub> no están siendo descritas adecuadamente, lo cual resulta en desempeños pobres para esta especie.

## AGRADECIMIENTOS

Se agradece el apoyo recibido por el Tecnológico de Monterrey, a través de la Cátedra de Investigación en Ingeniería de la Contaminación Atmosférica (número de apoyo CAT-186), para la realización de este trabajo.

## REFERENCIAS

- Aldrin M. (1997). Length modified ridge regression. *Comput. Stat. Data An.* 28, 377-398.
- Bergin M.S., Russell A.G., Odman M.T., Cohan D.S. y Chameides W.L. (2008). Single-source impact analysis using 3D air quality models. *J. Air Waste Manage.* 58, 1351-1359.
- Brown M. (1993). Deduction of emissions of source gases using an objective inversion algorithm and a chemical transport model. *J. Geophys. Res.* 98, 12639-12660.
- Carter W.P.L. (1990). A detailed mechanism for the gas-phase atmospheric reactions of organic compounds. *Atmos. Environ.* 24, 481-518.
- Chang M.E., Hartley D.E., Cardelino C., Haas-Laursen D. y Chang W.-L. (1997). On using inverse methods for resolving emissions with large spatial inhomogeneities. *J. Geophys. Res.* 102, 16023-16036.
- Cohan D.S., Boylan J.W., Marmur A. y Maudood K. (2007). An integrated framework for multipollutant air quality management and its application in Georgia. *Environ. Manage.* 40, 545-554.
- Corsmeier U., Imhof D., Kohler M., Kühlwein J., Kurtenbach R., Petrea M., Rosenbohm E., Vogel B. y Vogt U. (2005). Comparison of measured and model-calculated real-world traffic emissions. *Atmos. Environ.* 39, 5760-5775.
- Davydova-Belitskaya V., Skiba, Y.N., Martínez A. y Bulgakov S. (2001). Modelación matemática de los niveles de contaminación en la ciudad de Guadalajara, Jalisco, México. Parte II. Modelo numérico de transporte de contaminantes y su adjunto. *Rev. Int. Contam. Ambie.* 17, 97-107.
- Doll D.C., Scheffe R.D., Meyer E.L. y Chu S.-H. (1991). EPA-450/4-91-013. Guideline for regulatory application of the Urban Airshed Model. Office of Air Quality, Planning and Standards, United States Environmental Protection Agency, Research Triangle Park, NC. 89 pp.
- Elbern H. y Schmidt H. (1999). A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modeling. *J. Geophys. Res.* 104, 18583-18598.
- Elbern H., Strunk A., Schmidt H. y Talagrand O. (2007). Emission rate and chemical state estimation by dimensional variational inversion. *Atmos. Chem. Phys.* 7, 3749-3769.
- Grell G. A., Knoche R., Peckman S.E. y McKeen S. A. (2004). Online versus offline air quality modeling on cloud-resolving scales. *Geophys. Res. Lett.* 31, L16177, doi:10.1029/2004GL020175.
- Gobierno del Estado de Jalisco, Secretaría de Medio Ambiente, Recursos Naturales y Pesca, y Secretaría de Salud (1997). *Programa para el Mejoramiento de la Calidad del Aire en la Zona Metropolitana de Guadalajara 1997-2001*. Instituto Nacional de Ecología, México, D.F., 240 pp.
- Goodin W.R., McRae G.J. y Seinfeld J.H. (1979). A comparison of interpolation methods for sparse data: Application to wind and concentration fields. *J. Appl. Meteorol.* 18, 761-771.
- Goodin W.R., McRae G.J. y Seinfeld J.H. (1980). An objective analysis technique for constructing three-dimensional urban-scale wind fields. *J. Appl. Meteorol.* 19, 98-108.
- Hakami A., Odman M.T. y Russell A.G. (2003). High-order, direct sensitivity analysis of multidimensional air quality models. *Environ. Sci. Technol.* 37, 2442-2452.
- Hakami A., Henze D. K., Seinfeld J. H., Chai T., Tang Y., Carmichael G. R. y Sandu A. (2005). Adjoint inverse modeling of black carbon during the Asian Pacific Regional Aerosol Characterization Experiment. *J. Geophys. Res.* 110, D14301, doi: 10.1029/2004JD005671.
- Hanna S.R. (1988). Air quality model evaluation and uncertainty. *JAPCA.* 38, 460-412.
- Hao J., He D., Wu Y., Fu L. y He K. (2000). A study of the emission and concentration distribution of vehicular pollutants in the urban area of Beijing. *Atmos. Environ.* 34, 453-465.
- Harley R., Russell A.G., McRae G.J., Cass G.R. y Seinfeld J.H. (1993). Photochemical modeling of the Southern California Air Quality Study. *Environ. Sci. Technol.* 27, 378-388.
- Harley R.A., Marr L.C., Lehner J.K. y Giddings S.H. (2005). Changes in motor vehicle emissions on diurnal to decadal time scales and effects on atmospheric composition. *Environ. Sci. Technol.* 39, 5356-5362.
- Hoerl A.E. y Kennard R.W. (1976). Ridge regression iterative estimation of the biased parameter. *Commun. Stat. A-Theor.* A5, 77-88.
- Hogrefe C., Civerolo K. L., Hao W., Ku, J.-Y., Zalewsky E.E. y Sistla G. (2008). Rethinking the assessment of photochemical modeling systems in air quality planning applications. *J. Air Waste Manage.* 58, 1086-1099.
- Jang J.C.C., Jeffries H.E. y Tonnesen S. (1995). Sensitivity of ozone to model grid resolution – II: Detailed

- process analysis for ozone chemistry. *Atmos. Environ.* 29, 3101-3114.
- Kimura Y., McDonald-Buller E., Vizuete W. y Allen D.T. (2008). Application of a Lagrangian process analysis tool to characterize ozone formation in Southeast Texas. *Atmos. Environ.* 42, 5743-5759.
- McNair L.A., Harley R.A. y Russell A.G. (1996). Spatial inhomogeneity in pollutant concentrations, and their implications for air quality model evaluation. *Atmos. Environ.* 30, 4291-4301.
- McRae G.J., Goodin W. y Seinfeld J.H. (1982a). Development of a second-generation mathematical model for urban air pollution – I. Model formulation. *Atmos. Environ.* 16, 679-696.
- McRae G.J., Goodin W.R. y Seinfeld J.H. (1982b). Numerical solution of the atmospheric diffusion equation for chemically reacting flows. *J. Comp. Phys.* 45, 1-42.
- Mendoza-Domínguez A. y Russell A.G. (2000). Iterative inverse modeling and direct sensitivity analysis of a photochemical air quality model. *Environ. Sci. Technol.* 34, 4974-4981.
- Mendoza-Domínguez A. y Russell A.G. (2001). Estimation of emission adjustments from the application of four-dimensional data assimilation to photochemical air quality modeling. *Atmos. Environ.* 35, 2879-2894.
- Mendoza A. y García M.R. (2009). Aplicación de un modelo de calidad del aire de segunda generación a la Zona Metropolitana de Guadalajara, México. *Rev. Int. Contam. Ambie.* 25, 73-85.
- Napelenok S.L., Cohan D.S., Yongtao H. y Russell A.G. (2006). Decoupled direct 3D sensitivity analysis for particulate matter (DDM-3D/PM). *Atmos. Environ.* 40, 6112-6121.
- Napelenok S.L., Pinder R.W., Gilliland A.B. y Martin R.V. (2008). A method for evaluating spatially-resolved NO<sub>x</sub> emissions using Kalman filter inversion, direct sensitivities, and space-based NO<sub>2</sub> observations. *Atmos. Chem. Phys.* 8, 5603-5614.
- Niemeier D.A. (2003). The impact of incorporating spatial, temporal variability on running stabilized mobile emissions inventories. *Atmos. Environ.* 37, Supplement No. 2, S27-S37.
- NRC (1991). *Rethinking the Ozone Problem in Urban and Regional Air Pollution*. National Academy Press. Washington, EUA. 500 pp.
- Odman M.T. (1998). EPA/600/R-97/142. Research on Numerical Transport Algorithms for Air Quality Simulation Models. Office of Research and Development, United States Environmental Protection Agency. Washington, D.C. 62 pp.
- O'Neill S.M. y Lamb B.K. (2005). Intercomparison of the Community Multiscale Air Quality Model and CAL-GRID Using Process Analysis. *Environ. Sci. Technol.* 39, 5742-5753.
- Parrish D.D. (2006). Critical evaluation of US on-road vehicle emission inventories. *Atmos. Environ.* 40, 2288-2300.
- Placet M., Mann C.O., Gilbert R.O. y Niefer M.J. (2000). Emissions of ozone precursors from stationary sources: a critical review. *Atmos. Environ.* 34, 2183-2204.
- Russell A.G., McCue K.F. y Cass G.R. (1988). Mathematical modeling of the formation and transport of nitrogen-containing pollutants I: Model evaluation. *Environ. Sci. Technol.* 22, 263-271.
- Russell A. y Dennis R. (2000). NARSTO critical review of photochemical models and modeling. *Atmos. Environ.* 34, 2283-2324.
- Russell A.G. (2008). EPA Supersites program-related emissions-based particulate matter modeling: initial applications and advances. *J. Air Waste Manage.* 58, 289-302.
- Sawyer R.F., Harley R.A., Cadle S.H., Norbeck J.M., Slott R. y Bravo H.A. (2000). Mobile sources critical review: 1998 NARSTO assessment. *Atmos. Environ.* 34, 2161-2181.
- Srivastava R.K., McRae, D.S. y Odman, M.T. (2000). An adaptive grid algorithm for air quality modeling. *J. Comp. Phys.* 165, 437-472.
- Wesely M.L. (1989). Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models. *Atmos. Environ.* 23, 1293-1304.
- Yang Y.-J., Wilkinson J.G. y Russell A.G. (1997). Fast, direct sensitivity analysis of multidimensional photochemical models. *Environ. Sci. Technol.* 31, 2859-2868.
- Zhang Y., Stedman H.D., Bishop G.A., Guenther P.L. y Beaton S.P. (1995). Worldwide on-road vehicle exhaust emissions study by remote sensing. *Environ. Sci. Technol.* 29, 2286-2294.
- Zuk M., Rojas Bracho L. y Tzintzun Cervantes M.G. (2007). *Tercer almanaque de datos y tendencias de la calidad del aire en nueve ciudades mexicanas*. Instituto Nacional de Ecología, México, D.F.; 116 p.