

ANEXO 1

BASE CONCEPTUAL DEL ANÁLISIS DE INCERTIDUMBRE

COPRESIDENTES, EDITORES Y EXPERTOS

Copresidentes de la Reunión de expertos sobre metodologías intersectoriales de la estimación de la incertidumbre y la calidad de los inventarios

Taka Hiraishi (Japón) y Buruhani Nyenzi (Tanzanía)

EDITORES REVISORES

Richard Odingo (Kenya)

Grupo de Expertos: Base conceptual del análisis de incertidumbre

COPRESIDENTES

Ian Galbally (Australia) y Newton Paciornik (Brasil)

AUTORES DE LOS DOCUMENTOS DE ANTECEDENTES

Ian Galbally (Australia), Newton Paciornik (Brasil) y Milos Tichy (República Checa)

AUTORES COLABORADORES

Wiley Barbour (Estados Unidos), Leandro Buendia (IPCC-NGGIP/TSU), Franck Jacobs (Antigua), Naoki Matsuo (Japón), Richard Odingo (Kenya), Daniela Romano (Italia) y Milos Tichy (República Checa)

Índice

ANEXO 1 - BASE CONCEPTUAL DEL ANÁLISIS DE INCERTIDUMBRE

A1.1	INTRODUCCIÓN	A1.5
A1.2	CONCEPTOS ESTADÍSTICOS	A1.5
A1.2.1	Expresión de la incertidumbre.....	A1.5
A1.2.2	Muestra individual, valor medio e intervalo de confianza	A1.6
A1.2.3	Elección de la medida apropiada de la incertidumbre.....	A1.7
A1.2.4	Funciones de probabilidad	A1.8
A1.2.5	Orientación sobre las buenas prácticas para seleccionar una función de densidad de probabilidad.....	A1.9
A1.2.6	Caracterización de las funciones de densidad de probabilidad para los análisis de incertidumbre.....	A1.9
A1.3	FUENTES DE INCERTIDUMBRE EN LOS INVENTARIOS	A1.10
A1.4	EVALUACIÓN, REGISTRO Y PROPAGACIÓN DE INCERTIDUMBRES EN LOS INVENTARIOS.....	A1.11
A1.4.1	Determinación y registro de incertidumbres en los datos de entrada.....	A1.11
A1.4.2	Muestreo representativo, algoritmos y covarianza.....	A1.13
A1.4.3	Propagación de incertidumbres	A1.16
A1.4.4	Propagación de incertidumbres en todo el inventario	A1.18
A1.4.5	Covarianza y autocorrelación.....	A1.19
A1.4.6	Compilación sistemática de incertidumbres en componentes de inventarios.....	A1.19
A1.5	APLICACIONES.....	A1.20
A1.5.1	Importancia de las diferencias de un año a otro y las tendencias en los inventarios.....	A1.20
A1.5.2	Combinación ("empalme") de métodos.....	A1.21
A1.5.3	Análisis de sensibilidad y establecimiento de prioridades de investigación para inventarios nacionales	A1.22
A1.6	ASPECTOS QUE REQUIEREN INVESTIGACIÓN.....	A1.23
	REFERENCIAS	A1.24

Figura

Figura A1.1	Diagrama de flujo y árbol de decisiones para acciones relativas a la representatividad de los datos	A1.14
-------------	---	-------

Cuadro

Cuadro A1.1	Estimaciones nacionales de la masa de desechos enviados a vertederos en 1990.....	A1.7
-------------	---	------

ANEXO 1 - BASE CONCEPTUAL DEL ANÁLISIS DE INCERTIDUMBRE

A1.1 INTRODUCCIÓN

Para estimar las incertidumbres de los inventarios, es necesario adoptar un enfoque estructurado para la elaboración de una metodología. Se requiere lo siguiente:

- un método para determinar las incertidumbres de cada uno de los términos empleados en el inventario;
- un método para totalizar las incertidumbres de cada uno de los términos en el inventario;
- un método para determinar la importancia de las diferencias de año a año y de las tendencias a largo plazo de los inventarios, tomando en cuenta la información sobre las incertidumbres;
- la comprensión de las posibles utilidades de esa información, entre ellas la identificación de áreas que requieren más investigación y observaciones, y la cuantificación de la importancia de los cambios que ocurren año a año y a largo plazo en los inventarios nacionales de gases de efecto invernadero;
- la conciencia de que pueden existir otras incertidumbres, tales como las que surgen de definiciones inexactas que no se pueden tratar por medios estadísticos.

En el presente anexo se abordan los fundamentos de los conceptos que se emplean en otras partes de este informe para analizar las incertidumbres de los inventarios de gases de efecto invernadero. Por último, se examinan algunas cuestiones relativas a las incertidumbres de los inventarios que requieren mayor investigación.

A1.2 CONCEPTOS ESTADÍSTICOS

Ciertos conceptos y términos estadísticos básicos son fundamentales para comprender la incertidumbre en los inventarios de gases de efecto invernadero. Esos términos tienen significados en el lenguaje común, significados particulares en la literatura estadística y, en algunos casos, significados específicos por lo que respecta a la incertidumbre de los inventarios. El lector encontrará las definiciones en el Glosario del Anexo 3, el documento del Órgano Subsidiario de Asesoramiento Científico y Tecnológico de la CMCC (SBSTA-UNFCCC; 1999) y la *Guía para la expresión de la incertidumbre en las mediciones* de la Organización Internacional de Normalización (ISO, 1993).

El procedimiento de estimación de incertidumbres en los inventarios de gases de efecto invernadero se basa en ciertas características de la variable de interés (cantidad de entrada) estimada a partir de su correspondiente conjunto de datos. La información ideal comprende:

- la media aritmética (la media) del conjunto de datos;
- la desviación estándar del conjunto de datos (la raíz cuadrada de la varianza);
- la desviación estándar de la media (el error estándar de la media);
- la distribución de probabilidad de los datos;
- las covarianzas de la cantidad de entrada con otras cantidades de entrada empleadas en los cálculos del inventario.

A1.2.1 Expresión de la incertidumbre

Un aspecto importante del análisis de incertidumbre tiene que ver con los modos de expresar las incertidumbres relacionadas con las estimaciones individuales o con todo el inventario. En las *Directrices del IPCC para los inventarios nacionales de gases de efecto invernadero, versión revisada en 1996 (Directrices del IPCC)* se especifica lo siguiente: "Cuando existe suficiente información para definir la distribución de probabilidad en que se basa el análisis estadístico convencional, se debe calcular un intervalo de confianza de 95 % como definición del rango. Los rangos de la incertidumbre se pueden estimar por medio de un análisis clásico (Robinson, 1989) o

con la técnica de Monte Carlo (Eggleston, 1993). De otro modo, el rango deberá ser estimado por expertos nacionales”.

Según ese enunciado, el intervalo de confianza se determina por los límites de confianza definidos por el percentil 2,5 y el percentil 97,5 de la función de distribución acumulativa de la cantidad estimada. En otras palabras, el rango de una cantidad incierta en un inventario debería expresarse de tal modo que: i) haya una probabilidad de 95% de que el valor verdadero de la cantidad estimada se encuentre dentro del intervalo definido por los límites de confianza, y ii) sea igualmente probable que el valor verdadero, si estuviera fuera del rango citado, se encuentre por encima o por debajo de él.

A1.2.2 Muestra individual, valor medio e intervalo de confianza

Un aspecto clave de la compilación de incertidumbres de los inventarios es la distinción entre la desviación estándar del conjunto de datos y la desviación estándar de la media muestral. La incertidumbre relacionada con la información que se está analizando (tasa de emisión, datos de actividad o factores de emisión) puede ser la desviación estándar de la población muestral o la desviación estándar de la media muestral, dependiendo del contexto (ISO 1993).

La desviación estándar de la media, conocida también como error estándar de la media, es la desviación estándar del conjunto de datos de la muestra, dividida por la raíz cuadrada del número de puntos de datos. La desviación estándar y la varianza del conjunto de datos no varían sistemáticamente con el número de observaciones; sin embargo, la desviación estándar de la media disminuye a medida que aumenta el número de observaciones. En buena parte de la literatura de estadística y ciencias físicas, la desviación estándar de la media se conoce como el error estándar de la media, pero la ISO (1993) recomienda el uso del término “desviación estándar de la media” para esa cantidad.

El uso de la desviación estándar para estimar los límites del intervalo de confianza (en este caso, el intervalo de confianza de 95%) depende directamente de la distribución de probabilidad del conjunto de datos o de la función de probabilidad elegida para representar el conjunto de datos. En algunas distribuciones de probabilidad, entre ellas las que se examinan más adelante, existen relaciones analíticas entre la desviación estándar y los intervalos de confianza necesarios. En el anexo 3 (Glosario) del presente documento y en ISO (1993) figuran algunos ejemplos. En general, se supone que la variable en estudio tiene una distribución normal; en ese caso, los límites de confianza son simétricos a ambos lados de la media. Tratándose de un intervalo de confianza de 95%, los límites de confianza son aproximadamente dos desviaciones estándar de la variable, a uno y otro lado de la media.

Es probable que en muchas circunstancias, la cuantificación de incertidumbres de las variables de entrada del inventario suponga analizar pequeñas cantidades de datos y obtener el dictamen de expertos. Por consiguiente, es importante examinar el contenido de información de pequeños conjuntos de datos. Se han hecho buenos estudios de la cantidad de información sobre incertidumbres contenida en conjuntos de datos obtenidos a partir de un pequeño número de observaciones (Manly, 1997; Cullen y Frey, 1999). El término examinado es el intervalo de confianza de 95% de la estimación de una desviación estándar. Ésa es la incertidumbre de la estimación de la desviación estándar; significa, en esencia, cómo podría variar la desviación estándar de un conjunto a otro de observaciones hechas de la misma cantidad. Cullen y Frey (1999) presentaron datos a partir de los cuales se calcularon los límites del intervalo de confianza de 95% de la desviación estándar para una variable de distribución normal cuando la muestra utilizada para calcular la desviación estándar tiene un número dado de observaciones. Los límites del intervalo de confianza de 95% para repetidas determinaciones de la desviación estándar son los siguientes:

- 7 observaciones: 0,64 y 2,2 veces la desviación estándar estimada a partir de un número muy grande de observaciones;
- 20 observaciones: 0,76 y 1,5 veces la desviación estándar estimada a partir de un número muy grande de observaciones;
- 100 observaciones: 0,88 y 1,2 veces la desviación estándar estimada a partir de un número muy grande de observaciones.

Se hizo un análisis similar de la incertidumbre en estimaciones de intervalos de confianza de muestras de datos sintéticos con distribuciones no normales, empleando la técnica *bootstrap* (Manly, 1997), con los que se obtuvieron resultados similares a los señalados más arriba. Lo que esos cálculos ponen de relieve es que se necesitan números muy grandes de observaciones para estimar con precisión la varianza, la desviación estándar

y el error estándar de la media de cualquier cantidad. En esencia, los intervalos de confianza estimados a partir de pequeños números de observaciones sobre la base de una varianza (y una distribución de probabilidad hipotética) conllevan ciertas incertidumbres y, en esos casos, incrementar el número de observaciones puede aumentar o disminuir los límites de la incertidumbre calculados. En última instancia, al hacerse un gran número de observaciones disminuirán los límites de incertidumbre de la desviación estándar.

A1.2.3 Elección de la medida apropiada de la incertidumbre

A continuación se presentan dos ejemplos hipotéticos que ilustran la elección del error estándar de la media y la desviación estándar del conjunto de datos como la incertidumbre apropiada:

En el primer caso, el factor de emisión de un gas de efecto invernadero procedente de la quema de biomasa en la sabana, que se midió en 9 ocasiones distintas, varía entre 0 y $6 \cdot 10^{-3} \text{ kg kg}^{-1}$ (masa emitida por unidad de masa de la biomasa quemada) con una media aritmética y una desviación estándar del conjunto de datos de $2 \cdot 10^{-3} \text{ kg kg}^{-1}$ y $1 \cdot 10^{-3} \text{ kg kg}^{-1}$ respectivamente, a veces expresadas como $2 \pm 1 \cdot 10^{-3} \text{ kg kg}^{-1}$. El factor de emisión utilizado para ese año en el algoritmo del IPCC para inventarios es la media aritmética, y la incertidumbre apropiada para el inventario se debe basar en el error estándar de la media, que es $1 \cdot 10^{-3} / \sqrt{9} \text{ kg kg}^{-1}$ ó $3,3 \cdot 10^{-4} \text{ kg kg}^{-1}$, tres veces menor que la desviación estándar. La media y el intervalo de confianza de 95% están entonces incluidos en $2 \pm 0,7 \cdot 10^{-3} \text{ kg kg}^{-1}$.

El segundo caso se refiere a un componente de un inventario, para el cual existe una única estimación para un año determinado, que se ha calculado en más de una ocasión. Los nuevos cálculos se hicieron al efectuarse cambios en la metodología convenida, realizarse auditorías del inventario o recogerse nuevos datos. En ese caso, conviene usar la desviación estándar de la muestra y no la desviación estándar de la media.

Ese punto se puede ilustrar por medio del conjunto de estimaciones nacionales de desechos destinados a vertederos que figura en el cuadro A1.1. Son los datos de actividad que se necesitan para calcular las emisiones de gases de efecto invernadero procedentes de desechos.

CUADRO A1.1	
ESTIMACIONES NACIONALES DE LA MASA DE DESECHOS ENVIADOS A VERTEDEROS EN EL AÑO 1990	
Fuente y año de la estimación	Masa (kilotoneladas)
Comisión de Tecnología, 1991	12.274
Consultante, 1994	11.524
Inventario nacional, 1994	14.663
Revisión del inventario nacional, 1995	16.448
Revisión del inventario nacional, 1996	12.840
Estudio académico, 1995	22.000
Media	14.958
Desviación estándar	3.883

Observamos que la media y el intervalo de confianza de 95% basado en el error estándar de la media de las seis estimaciones es 14.958 ± 3.107 . Sin embargo, en el caso de emplearse la estimación del inventario de 1996, se utiliza una única estimación y la incertidumbre que corresponde al inventario se calcula a partir de la desviación estándar del conjunto de datos.

En concreto, si sólo se tienen en cuenta los datos del cuadro A1.1, el intervalo de confianza de 95% relacionado con la estimación de 1996 debería ser de dos desviaciones estándar, a saber, 12.840 ± 7.610 . Como se trata de una sola estimación, es necesario reevaluar los datos. La razón es que la estimación de 1996 no es el valor medio de muchas determinaciones independientes.

La elección de la medida apropiada de la incertidumbre depende del contexto del análisis. Si sólo se dispone de un punto de datos por período de inventario, el rango de la incertidumbre debería basarse en la función de

densidad de probabilidad de la población si ésta se conoce o puede obtenerse de otras fuentes. Las medidas elegidas deberían examinarse en el marco del proceso de revisión de los inventarios realizado por expertos.

A1.2.4 Funciones de probabilidad

Cuando se hacen múltiples determinaciones de una cantidad que es una entrada en el inventario, se obtiene un conjunto de datos que tiene variabilidad. La cuestión está en cómo representar esa variabilidad en forma compacta. Una forma de hacerlo es determinar los siguientes estadísticos recapitulativos (ISO, 1993; Cullen y Frey, 1999):

- media aritmética;
- varianza;
- asimetría de la distribución;
- curtosis (aplanamiento de la distribución).

Sin embargo, si la determinación de los límites de la incertidumbre de los datos de entrada se considera desde el punto de vista de la frecuencia (límites de confianza de 95%) se necesita más información sobre el conjunto de datos y los estadísticos recapitulativos. Esa información suplementaria se puede obtener representando los datos como una distribución de probabilidad acumulativa o una distribución de densidad (ISO, 1993; Cullen y Frey 1999). Ése es el enfoque que se adoptó en el Capítulo 6, “La cuantificación de incertidumbres en la práctica”. Una distribución acumulativa empírica establece una relación entre los percentiles y los datos¹. El percentil es el porcentaje de valores del conjunto de datos menor o igual que un valor dado de la cantidad.

Para la tarea ulterior de calcular la propagación de errores en un sistema complejo (empleando métodos analíticos o computacionales), las distribuciones de probabilidad empíricas son engorrosas. El método común consiste en reemplazar la distribución empírica con una función analítica, ya sea la función de distribución acumulativa (FDA) o bien una función de densidad de probabilidad (FDP), que es la primera derivada de la FDA. Esas funciones son, de hecho, el primer componente de un modelo del proceso de la incertidumbre. Asimismo, son sólo una aproximación a los datos reales. Esas funciones de probabilidad son esenciales para dos aspectos del trabajo con incertidumbres. Se necesitan para i) la propagación de las incertidumbres y ii) la determinación del intervalo de confianza de la cantidad en cuestión.

En la literatura estadística se encuentran numerosas funciones de probabilidad, que a menudo representan situaciones particulares del mundo físico. He aquí algunos ejemplos de tales funciones y de las situaciones que representan:

- distribución normal – estaturas de los seres humanos;
- distribución lognormal – concentraciones de sustancias químicas en el medio ambiente.

Esas funciones también se pueden expresar en formas truncadas que representan la situación cuando se conocen los límites físicos del rango posible de los datos.

Para representar la ausencia de información sobre los procesos, se emplean otras distribuciones; por ejemplo:

- distribución uniforme – todos los valores de un rango dado tienen igual probabilidad;
- distribución triangular – están asignados los límites superior e inferior y un valor de preferencia de este rango.

Puede ser difícil determinar la función que mejor se ajuste a un conjunto de datos. Uno de los métodos es utilizar el cuadrado de la asimetría y la curtosis para definir formas funcionales que se ajusten a los datos (Cullen y Frey, 1999). La función se ajusta entonces a los datos empleando el método de mínimos cuadrados u otros medios. Existen varias pruebas para determinar la exactitud del ajuste, entre ellas la del ji-cuadrado (Cullen y Frey, 1999). En muchos casos, varias funciones se ajustarán satisfactoriamente a los datos dentro de un límite de probabilidad dado. Esas diversas funciones pueden tener distribuciones radicalmente diferentes en los extremos, donde los datos para limitarlas son escasos o inexistentes, y la función que se elija puede cambiar sistemáticamente el resultado de un análisis de incertidumbre. Cullen y Frey (1999) reiteran el consejo que

¹ Un punto fundamental con respecto a ambos conjuntos de datos y su representación como distribuciones de probabilidad acumulativas y empíricas es que no se dispone de información sobre los valores probables de la cantidad para probabilidades percentiles menores de 50/ n o mayores de (100-50/ n), en que n es el número de observaciones. De hecho, los datos de las probabilidades en las colas son muy inciertos.

dieron autores anteriores para esos casos, a saber, que *el conocimiento de los procesos físicos subyacentes debe ser lo que determine la elección de la función de probabilidad*. Las pruebas sirven de guía para saber, a la luz del conocimiento de los procesos físicos, si la función se ajusta o no a los datos de manera satisfactoria.

A1.2.5 Orientación sobre las buenas prácticas para seleccionar una función de densidad de probabilidad

Los criterios de comparabilidad, coherencia y transparencia en los inventarios de emisiones, que se definen en otra parte de este documento, se cumplen de la mejor manera cuando:

- se usa el menor número de funciones de probabilidad;
- las funciones de probabilidad se conocen bien y están bien fundamentadas.

Ésas serían las funciones de probabilidad por defecto.

Los criterios de exactitud se cumplen cuando:

- las funciones de probabilidad por defecto se ajustan bien a los datos, o bien,
- se usa una función de probabilidad más adecuada, si las funciones de probabilidad por defecto no se ajustan bien a los datos o hay razones científicas convincentes para usar otra función de probabilidad.

En la siguiente orientación sobre *buenas prácticas*, se describe de qué modo los organismos encargados de los inventarios pueden cumplir dichos criterios.

- i) Cuando se dispone de datos empíricos, lo primero que se debería suponer es que los datos tienen una distribución normal (ya sea completa o truncada para evitar valores negativos, si éstos no correspondieran a la realidad), salvo si el diagrama de dispersión de los datos sugiere que se ajustan mejor a otra distribución.
- ii) Cuando se recurre al dictamen de expertos, la función de distribución que se adopte debería ser normal o lognormal como en i), complementada por distribuciones uniformes o triangulares, según se describe en el anexo 3.
- iii) Se usarán otras distribuciones solamente cuando haya razones convincentes, basadas ya sea en observaciones empíricas o en el dictamen de expertos respaldado por argumentos teóricos.

A1.2.6 Caracterización de las funciones de densidad de probabilidad para los análisis de incertidumbre

Las características de las FDP que se refieren a la cuantificación y totalización de las incertidumbres asociadas con las cantidades incluidas en los inventarios nacionales de gases de efecto invernadero son las siguientes:

- la forma matemática de la FDP;
- los parámetros que se requieren como valores de entrada para especificar la FDP;
- las relaciones entre los parámetros que especifican la FDP y los datos disponibles acerca de la cantidad que se describe;
- la media, la varianza y el error estándar de la media, calculados a partir del conjunto de datos que se utilizan para determinar los parámetros de la FDP.

Al seleccionar los valores de entrada y la FDP, el compilador del inventario debe distinguir los casos en que la incertidumbre apropiada es la desviación estándar o los intervalos de confianza del conjunto de datos, y los casos en que la incertidumbre apropiada es el error estándar del valor medio.

Como se mencionó anteriormente, si se elige mal la medida que se emplee para estimar la incertidumbre, se podrían obtener resultados falsos.

A1.3 FUENTES DE INCERTIDUMBRE EN LOS INVENTARIOS

Algunas fuentes de incertidumbre se pueden tratar por medios estadísticos, otras caen fuera del campo de la estadística (ISO 1993).

Las incertidumbres de los inventarios obedecen por lo menos a tres procesos diferentes:

- incertidumbres que resultan de las definiciones (por ejemplo, significado incompleto o poco claro, o definición incorrecta de una emisión o absorción);
- incertidumbres generadas por la variabilidad natural del proceso que produce una emisión o absorción;
- incertidumbres que resultan de la evaluación del proceso o la cantidad; dependiendo del método que se use cabe mencionar las debidas a i) la medición; ii) el muestreo; iii) una descripción incompleta de los datos de referencia, y iv) el dictamen de expertos.

Las incertidumbres debidas a malas definiciones tienen que ver con una falta de exhaustividad y con problemas de atribución a las categorías de fuentes y, en lo posible, deben eliminarse antes de emprender el análisis.

Las incertidumbres que resultan de la variabilidad natural son inherentes al proceso de emisión y se pueden evaluar mediante el análisis estadístico de datos representativos.

Entre los defectos de medición que dan lugar a incertidumbres cabe señalar los siguientes:

- sesgos personales al medir, registrar o transmitir información;
- limitaciones de la resolución de los instrumentos o el umbral de discriminación;
- inexactitud de los valores de las normas de medición y material de referencia;
- inexactitud de los valores de constantes y otros parámetros obtenidos de fuentes externas y utilizados en el algoritmo de reducción de datos (por ejemplo, valores por defecto de las *Directrices del IPCC*);
- aproximaciones e hipótesis incluidas en el método de medición y el procedimiento de estimación;
- variaciones en la repetición de observaciones sobre la emisión, absorción o parámetros conexos, efectuadas en condiciones aparentemente idénticas.

La medición continua de las emisiones puede reducir la incertidumbre general; sin embargo, la aplicación de ese método a la evaluación de emisiones de gases de efecto invernadero suele ser limitada. Lo más frecuente es que se hagan muestreos periódicos y aleatorios, los cuales introducen otras incertidumbres, como las siguientes:

- *error de muestreo aleatorio*. Esta fuente de incertidumbre surge cuando los datos proceden de una muestra aleatoria de tamaño finito y generalmente depende de la varianza de la población de la cual se extrajo la muestra y el tamaño de ésta (número de puntos de datos).
- *falta de representatividad*. Esta fuente de incertidumbre nace cuando no existe una correspondencia completa entre las condiciones en que se recogieron los datos disponibles y las condiciones en que tienen lugar las emisiones o actividades en el mundo real. Es lo que ocurre, por ejemplo, cuando se dispone de datos sobre las emisiones producidas por una planta mientras opera a plena capacidad pero no durante su puesta en marcha o en los cambios de carga. En ese caso, los datos sólo son parcialmente aplicables a la estimación de la emisión que se desea establecer.

Por definición, las incertidumbres que plantea la opinión de expertos no se pueden evaluar por medios estadísticos ya que sólo se recurre a ese dictamen cuando no se dispone de suficientes datos empíricos. Sin embargo, el dictamen de expertos se puede combinar con datos empíricos para hacer un análisis mediante procedimientos estadísticos, siempre que se trate de conformidad con los procedimientos prácticos resumidos aquí y en el capítulo 6, “La cuantificación de incertidumbres en la práctica”.

En la evaluación de las incertidumbres de los inventarios, se deben tener en cuenta todas esas fuentes de incertidumbre.

La Organización Internacional de Normalización (ISO, 1993) subraya que, con “materiales naturales”, las incertidumbres debidas al muestreo y a la necesidad de lograr la representatividad de la muestra pueden tener más peso que las incertidumbres que resultan del método de medición. Los problemas de muestreo tienen importancia en la evaluación de las incertidumbres del inventario. La representatividad de la muestra afecta directamente la incertidumbre del inventario. El problema general de determinar la incertidumbre en los

inventarios es una combinación de un problema estadístico de análisis de errores y un problema de correspondencia entre conceptos de estadística e inventarios y los hechos del mundo real.

A1.4 EVALUACIÓN, REGISTRO Y PROPAGACIÓN DE INCERTIDUMBRES EN LOS INVENTARIOS

A1.4.1 Determinación y registro de incertidumbres en los datos de entrada

La medida de cada cantidad física que constituye un dato de entrada en los algoritmos de inventarios trae aparejada cierta incertidumbre. En algunos casos, tales como la razón de pesos moleculares, la incertidumbre es insignificante a efectos del inventario, pero en casi todos los otros casos, la incertidumbre requiere evaluación.

La *buena práctica* en la estimación de incertidumbres de los datos de entrada para inventarios se basa en varios principios. Lo ideal es tener cientos de mediciones de la cantidad de entrada y poder estimar los intervalos de confianza por métodos estadísticos clásicos. Sin embargo, en la mayoría de los casos no se dispone de datos o éstos son escasos. Según la situación, se pueden utilizar cuatro tipos de información en proporción variable:

- mediciones disponibles de la cantidad;
- información sobre los valores extremos de la cantidad;
- información sobre los procesos subyacentes que regulan la cantidad y su varianza;
- el dictamen de expertos.

La recopilación y el registro de información sobre la incertidumbre de los datos de entrada es fundamental para el éxito y la transparencia del análisis de incertidumbre. El recuadro A1.1 contiene una lista de la información que se necesita para hacer un análisis exhaustivo y transparente de conformidad con la *buena práctica*. En la práctica, quizá no se disponga de la información completa y sea necesario recurrir al dictamen de expertos.

RECUADRO A1.1**INFORMACIÓN QUE CONVIENE TENER DE CADA CANTIDAD DE ENTRADA EN UN INVENTARIO NACIONAL DE GASES DE EFECTO INVERNADERO PARA HACER UN ANÁLISIS TRANSPARENTE DE INCERTIDUMBRE**

- i) Nombre de la cantidad;
- ii) unidades;
- iii) descripción del campo espacial, temporal y de sistemas que representa esa cantidad;
- iv) valor de entrada de la cantidad;
- v) indicación de si se trata de un valor medio de un conjunto de datos o de una observación única;
- vi) indicación de si la incertidumbre en cuestión es la desviación estándar de la media muestral o la desviación estándar de la población;
- vii) tamaño de la muestra o número de estimaciones disponibles de la cantidad;
- viii) estimación de la desviación estándar de la media muestral o estimación de la desviación estándar de la población;
- ix) estimaciones de la varianza de la cantidad basadas en el conocimiento que se tenga sobre los factores y procesos de control que inciden en la cantidad;
- x) límites superiores e inferiores de los valores de la cantidad basados en análisis científicos y el dictamen de expertos;
- xi) la función de densidad de probabilidad preferida;
- xii) parámetros de entrada para especificar la función de densidad de probabilidad;
- xiii) una explicación concisa del fundamento o causa de la incertidumbre;
- xiv) referencias de las fuentes del dictamen de los expertos y los datos que se usaron en esta tabulación;
- xv) documentación de la revisión del análisis efectuada por otros expertos en la materia.

A1.4.1.1 EL DICTAMEN DE EXPERTOS

Cuando no se pueden obtener datos fiables o los datos del inventario no proporcionan suficiente información estadística, puede ser necesario recurrir a la opinión de expertos acerca de la naturaleza y propiedades de los datos de entrada. Es posible que los expertos prefieran indicar niveles relativos de incertidumbre u otras informaciones cualitativas en vez de dar información cuantitativa sobre la incertidumbre y calidad de los datos. Los protocolos de solicitud de dictamen, que se examinan en el capítulo 6, “La cuantificación de incertidumbres en la práctica”, pueden ayudar a desvanecer esas preocupaciones. Si es necesario, se debería informar a los expertos de la existencia de los rangos de incertidumbres por defecto dados por el IPCC que se usarían si no se pudieran obtener sus dictámenes.

Es aceptable recurrir a la opinión de expertos para hacer esas estimaciones cuantitativas de la incertidumbre, siempre que se tengan en cuenta todos los datos disponibles y que las opiniones sean fruto del razonamiento de personas con conocimientos o experiencia especiales sobre la cantidad en cuestión, y a condición de que el dictamen esté documentado y se pueda explicar con suficiente claridad para satisfacer un examen independiente (Cullen y Frey, 1999). El requisito fundamental al hacer estimaciones de incertidumbre basadas en el dictamen de expertos o cualquier otro método es que se tomen en cuenta todas las fuentes posibles de incertidumbre.

A menudo se cuenta con pocas observaciones para determinar datos de entrada a esos inventarios y es necesario depender en gran medida de la opinión de expertos. Debe reconocerse que los resultados de los análisis cuantitativos de incertidumbre de los inventarios proporcionan, en el mejor de los casos, una estimación de la incertidumbre, pero que también hay incertidumbres considerables relacionadas con los intervalos de confianza.

A1.4.2 Muestreo representativo, algoritmos y covarianza

Las cuestiones relativas al muestreo representativo y la elaboración de algoritmos adecuados para representar emisiones están estrechamente relacionadas. La cuestión del muestreo representativo se plantea porque el inventario debe incluir todas las emisiones (o absorciones) que tengan lugar dentro de las fronteras del país y durante el período del inventario. Sin embargo, las mediciones están limitadas por el espacio y el tiempo. Las emisiones de las diferentes actividades se calculan como el producto de los datos de actividad y el factor de emisión conexas. Los datos de ambas variables deben ser representativos de la realidad del campo espacial y temporal en estudio. Un factor de emisión se considera representativo si se calcula como el promedio ponderado de todos los factores de emisión relacionados con todas las diferentes tipologías de procesos o productos, mientras que los pesos son los porcentajes de las diferentes producciones/productos que componen el total. Los datos de actividad se pueden considerar representativos si abarcan todas las actividades del período en estudio. En muchos casos, no se dispone de datos de actividad y factores de emisión para una región o una determinada categoría de procesos, de modo que es necesario estimar las emisiones mediante factores de emisión determinados en otra región o en una categoría de procesos diferente. Se trata de un procedimiento de extrapolación. De otro modo, se podrían calcular los valores por medio de variables de sustitución. Cuando se hace una extrapolación o se emplean variables de sustitución, es necesario evaluar la representatividad de los valores elegidos. Los datos son más representativos y, por ende, más exactos si se utiliza un proceso o condición similar.

Existen métodos estadísticos para estimar la incertidumbre asociada a la extrapolación, siempre que los datos disponibles provengan de un muestreo aleatorio. Sin embargo, en el caso de los inventarios nacionales, es raro que los datos provengan de un muestreo aleatorio. Por consiguiente y en vista de la naturaleza heterogénea de las emisiones y absorciones de gases de efecto invernadero, la cuestión fundamental relativa a la extrapolación es la de la incertidumbre relacionada con el grado de representatividad del muestreo. Por ejemplo, la extrapolación de una tasa conocida de emisión de una arrocera irrigada a un campo cualquiera, incluyendo una arrocera de secano, producirá un alto grado de incertidumbre. En cambio, es posible estratificar los datos de actividad del país en arroceras irrigadas y arroceras de secano, y producir un análisis mucho más fiable. La biosfera rara vez es homogénea y la estratificación es una técnica poderosa para manejar y reducir la incertidumbre de las estimaciones de inventarios.

Si se dispone de suficientes recursos, quizá se pueda emprender una campaña de observaciones sistemáticas y diseñar una muestra estratificada de mediciones, seleccionando las variables más apropiadas para estratificar la muestra (productos, procesos, plantas, territorio, población). Se puede utilizar todo el conjunto de datos para estimar la función de densidad de probabilidad y los estadísticos recapitulativos. Luego se pueden emplear herramientas estadísticas para calcular el sesgo de la media y la varianza, los intervalos de confianza y la distribución de los errores. Cuando falten datos a nivel regional, se podrá extrapolar información de la literatura existente siempre que se ponga cuidado en elegir datos de fuentes que tengan características similares a las que se están estimando. En ese caso, es necesario recabar la opinión de expertos.

Esa actividad o proceso requiere un conjunto mínimo de procedimientos, que se esbozan en el árbol de decisiones de la figura A1.1.

En primer lugar, es necesario determinar si los datos son para una categoría principal de fuentes de emisión (de acuerdo a la descripción que figura en el capítulo 7, “Elección de la metodología y realización de nuevos cálculos”). Si no se trata de una categoría principal de fuentes, es apropiado utilizar el conjunto de datos existente, valores extrapolados, valores por defecto o el dictamen de expertos, y esos datos se pueden registrar. Si se trata de una categoría principal de fuentes ya existirá, o se podrá extrapolar, un conjunto completo de datos, o bien será necesario iniciar un programa de observaciones o reunir datos. Luego será necesario poner a prueba el conjunto de datos para determinar su representatividad, lo cual podría requerir una estratificación (a efectos de afinar o mejorar la exactitud). Por último, se deberán registrar todos los datos. La figura A1.1 muestra las vías que se siguen en el procedimiento. Es necesario tomar en cuenta el tiempo de promediación de los datos en comparación con el del inventario, y la aplicabilidad geográfica de los datos. Por ejemplo, supóngase el caso de que, para un determinado factor de emisión, los datos se basen en mediciones hechas para plazos de corta duración (p. ej., todas las horas o una vez al día) en un solo país en condiciones posiblemente propias de ese lugar, y que haya que utilizar esos datos para estimar tanto las emisiones anuales como sus incertidumbres en un país distinto. Se recomienda al analista que emplee criterios y métodos razonables para elaborar una estimación debidamente representativa de la incertidumbre que se pueda ingresar en el inventario de emisiones. Tales ajustes, aunque no siempre perfectos, son preferibles a utilizar un conjunto de datos no representativos. Por consiguiente, la carga de justificar las hipótesis empleadas en una evaluación particular recae en el analista y

En la mayoría de los casos, es imposible medir directamente una parte importante de las emisiones de una categoría de fuentes de un país durante buena parte del año. Lo que se necesita para el inventario es la suma de las emisiones y absorciones que ocurren en todo el país durante todo el año del inventario; sin embargo, lo que se mide directamente son las emisiones y absorciones que tienen lugar durante un período de tiempo mucho menor que un año y en una superficie mucho menor que la del territorio del país. La emisión observada representa sólo un subconjunto del inventario y, por consiguiente, se necesita un método de extrapolación.

El método de extrapolación se basa en los algoritmos de las *Directrices del IPCC* y la información sobre las cantidades de entrada de todo el país y todo el año de inventario. Como el interés por las emisiones de gases de efecto invernadero se ha despertado hace poco tiempo, las mediciones necesarias para cuantificar las emisiones sólo se han hecho en un número reducido de lugares y en una variedad limitada de condiciones. El algoritmo utilizado para estimar las emisiones es una aproximación que incluye sólo las principales variables que surgen de las mediciones disponibles y en general representa sólo una parte limitada de la varianza de los datos disponibles. Al mismo tiempo, muchas fuentes posiblemente importantes de covarianza de las emisiones reales desaparecen de los cálculos del inventario por falta de conocimientos sobre el mecanismo de las emisiones.

Un método eficiente para recoger otros datos representativos y mejorar al mismo tiempo la calidad de los algoritmos es realizar un programa de muestreo estratificado de las emisiones y recoger la información justificativa pertinente. El muestreo estratificado es una técnica estadística común (Cochran, 1963).

El muestreo estratificado se hace en varias etapas. La primera consiste en identificar las variables (ambientales, tecnológicas, etc.) que se sabe que tienen una influencia significativa en las emisiones en cuestión. El conocimiento sobre la influencia de las variables se puede basar en estudios de laboratorio, modelizaciones teóricas, observaciones de campo y otras fuentes de información. Una vez identificadas las variables principales, se deben estimar sus distribuciones acumulativas en el dominio del inventario. Por último, se debe verificar si las observaciones disponibles constituyen una muestra representativa de esas distribuciones. Si no es así, se puede dividir las distribuciones en estratos, y diseñar y ejecutar un programa de muestreo para conseguir datos representativos. Éstos se pueden utilizar para revisar el algoritmo de las emisiones. Un algoritmo de emisiones basado en un conjunto de datos representativos es un requisito esencial para obtener un inventario de alta calidad.

He aquí un ejemplo que ilustra esos aspectos de la representatividad de los datos. Se refiere a las emisiones de óxido nitroso (N_2O) procedentes de la aplicación de fertilizantes en cultivos de secano. La mayor parte de los datos utilizados para construir el algoritmo y el factor de emisión general por defecto que se usa actualmente en los inventarios del IPCC provienen de sistemas de cultivo de zonas templadas del hemisferio norte. Bouwman (1996) hizo un excelente análisis sistemático de los datos (disponibles en aquel momento) sobre las emisiones de N_2O procedentes de la aplicación de fertilizantes y dedujo un algoritmo basado exclusivamente en la cantidad de nitrógeno contenida en los fertilizantes aplicados y en un factor de emisión. Sin embargo, como Bouwman (1996) lo reconoció, la edafología enseña que existen otros factores importantes que pueden contribuir a la varianza de las emisiones, entre ellos la temperatura y fertilidad del suelo, la frecuencia y cantidad de lluvia, el anegamiento del suelo y la composición del fertilizante. En consecuencia, un factor de emisión que se base principalmente en sistemas de cultivo de zonas templadas del hemisferio norte puede no servir para climas tropicales cálidos, donde las variables ambientales pertinentes, tales como la temperatura del suelo y la frecuencia de la lluvia, son totalmente diferentes de las de las latitudes templadas. Cuando el algoritmo y el factor de emisión del IPCC (que se basan en los mejores datos disponibles) se aplican a regiones tropicales, las estimaciones pueden resultar involuntariamente sesgadas. El sesgo puede producirse por falta de datos adecuados sobre las emisiones en los trópicos. El problema tiene que ver con la representatividad de los datos sobre las emisiones de N_2O procedentes de la aplicación de fertilizantes. Cuando no se tienen datos representativos sobre una emisión o absorción de importancia, se deben hacer las mediciones adecuadas – en este caso, de las emisiones de N_2O procedentes de la aplicación de fertilizantes en los trópicos – y luego revisar el algoritmo y el factor de emisión. En casos como éste, los factores de emisión por defecto se deben reemplazar por factores de emisión regionales, si éstos son más adecuados. Ese proceso de examinar la representatividad de los datos y llenar las lagunas en los datos esenciales debería producir un aumento sustancial de la confianza de la estimación del inventario. Es fundamental para reducir la incertidumbre de los inventarios y constituye una *buena práctica*. Este ejemplo es sólo uno de los muchos casos en que se podría mejorar la representatividad de los datos esenciales.

Una cuestión relacionada con la incertidumbre y el examen de algoritmos es que puede haber una considerable varianza inexplicada en un algoritmo elaborado a partir de un conjunto de datos. Esa varianza inexplicada debería representarse en las estimaciones de la incertidumbre de cada parámetro del algoritmo, incluyendo los exponentes. En los análisis posteriores se deben tomar en cuenta esas incertidumbres.

El muestreo estratificado es una técnica útil en los casos de covarianza entre los datos de actividad y los factores de emisión. La covarianza se reduce estratificando los datos de actividad y los factores de emisión en conjuntos

cuidadosamente seleccionados. Este método ya se ha aplicado ampliamente en la metodología del IPCC para los inventarios.

Algunos paquetes numéricos de la propagación de errores de Monte Carlo incluyen covarianzas en sus cálculos y requieren que se entre la matriz de correlación de todas las cantidades de entrada. Por lo tanto, es importante tener métodos para estimar esas correlaciones o evitar la necesidad de hacerlas.

El problema que se plantea en la compilación de inventarios y, en particular, en esta etapa del cálculo de la incertidumbre en la estimación de emisiones es la determinación del valor probable de la covarianza o el correspondiente coeficiente de correlación entre las diversas cantidades de entrada, en este caso entre las diversas actividades y también entre las actividades y los factores de emisión conexos. Es necesario evaluar los coeficientes de correlación de una variedad de categorías de inventario: fuentes fijas de combustión, fuentes móviles, emisiones fugitivas, procesos industriales, agricultura, y cambio del uso de la tierra y silvicultura. Es necesario conocer la correlación sea cual sea el método utilizado para calcular las incertidumbres, ya sea la ecuación de propagación de errores o el método de Monte Carlo.

El aumento de la emisión que ocurre al ponerse en marcha un equipo constituye un ejemplo de una posible correlación entre actividad y factor de emisión en una sola categoría de fuentes. En este caso, una actividad local baja o períodos cortos y frecuentes de actividad (en el tiempo o el espacio) se relacionan con emisiones elevadas, y un número menor de períodos más largos de actividad local se relacionan con emisiones más bajas, lo cual constituye una correlación negativa.

De modo semejante, en el caso de las emisiones de metano (CH₄) procedentes de animales, habrá una correlación entre el número total de animales y el peso corporal medio a lo largo del año, que puede producir una covarianza que incida en las emisiones de CH₄ de animales. El efecto de esta covarianza en las emisiones se puede minimizar desagregando los cálculos según la edad de los animales y la estación del año.

A1.4.3 Propagación de incertidumbres

Existen muchos métodos que se pueden usar para la propagación de incertidumbres, entre ellos los que se describen en general como métodos analíticos, de aproximación y numéricos. Para la propagación de incertidumbres en inventarios nacionales de gases de efecto invernadero, examinaremos dos métodos generales, a saber: el método de aproximación basado en la expansión en series de Taylor de primer orden, que suele llamarse ecuación de propagación de errores, y el método numérico de Monte Carlo.

A1.4.3.1 ECUACIÓN DE PROPAGACIÓN DE ERRORES

En el primer método, una incertidumbre en una emisión se puede propagar a partir de incertidumbres en la actividad y el factor de emisión por medio de la ecuación de propagación de errores (Mandel 1984; Bevington y Robinson 1992). El empleo de este método, que figura en las actuales *Directrices del IPCC*, requiere que se cumplan las siguientes condiciones:

- las incertidumbres son relativamente pequeñas, la desviación estándar dividida por el valor medio es inferior a 0,3;
- las incertidumbres tienen distribuciones gaussianas (normales)²;
- las incertidumbres no tienen una covarianza significativa.

En esas condiciones, la incertidumbre calculada para la tasa de emisión es apropiada. El método se puede ampliar para tener en cuenta covarianzas.

La ecuación de propagación de errores es un método de combinación de varianzas y covarianzas para una variedad de funciones, incluyendo las que se utilizan en inventarios. Este método permite expandir ecuaciones no lineales por medio de la expansión de Taylor. Da una solución exacta para funciones lineales aditivas y una aproximación para productos de dos términos. La mayoría de los inventarios de emisiones son sumas de emisiones, E, que son los productos de datos de actividad, A, y factores de emisión, F. Suponiendo que ambas cantidades tengan alguna incertidumbre, esas ecuaciones de inventarios no son lineales con respecto a los cálculos de incertidumbres. Por lo tanto, la ecuación de propagación de errores da sólo una estimación

² De hecho, la condición de que las incertidumbres tengan distribuciones gaussianas (normales) no es necesaria para que el método sea aplicable.

aproximada de la incertidumbre combinada que es cada vez más inexacta para desviaciones mayores. El error sistemático que se produce cuando no se toma en cuenta esa no linealidad en los inventarios se puede evaluar caso por caso. El método es muy inexacto con respecto a funciones que contienen términos exponenciales, de potencias superiores o inversos (Cullen y Frey, 1999). Se pueden incluir términos para tomar en cuenta los efectos de la covarianza.

Cuando la actividad y el factor de emisión son mutuamente independientes, sus varianzas en una sola categoría de fuentes se pueden combinar de acuerdo a la ecuación A1.1,

$$\text{ECUACIÓN A1.1}$$

$$\sigma_E^2 = \sigma_A^2 F^2 + \sigma_F^2 A^2$$

donde σ_E^2 es la varianza de la emisión, σ_A^2 es la varianza de los datos de actividad, σ_F^2 es la varianza del factor de emisión, A es el valor esperado de los datos de actividad y F es el valor esperado del factor de emisión.

Cuando las variables están correlacionadas, pero las incertidumbres son pequeñas, el método que se indica a continuación es válido. La covarianza entre dos variables, $\text{cov}(x,y)$, se puede derivar de su coeficiente de correlación, r_{xy} , y las desviaciones estándar del modo siguiente:

$$\text{ECUACIÓN A1.2}$$

$$\text{cov}(x, y) = r_{xy} \sigma_x \sigma_y$$

La ecuación A1.1 se expande para dar:

$$\text{ECUACIÓN A1.3}$$

$$\sigma_E^2 = \sigma_A^2 F^2 + \sigma_F^2 A^2 + 2r_{AF} \sigma_A \sigma_F AF$$

La ecuación A1.3 muestra que la varianza del producto puede, en el caso extremo, duplicarse o llegar al valor cero si la correlación entre los dos componentes se aproxima a sus valores extremos de +1,0 y -1,0 y los coeficientes de variación tienen igual valor. En la práctica, la correlación entre los factores de emisión y los datos de actividad debe manejarse estratificando los datos o combinando las categorías en que hay covarianza, y esos son los métodos adoptados en la recomendación sobre *buenas prácticas* específicas de las fuentes que se formula en el capítulo 6, "La cuantificación de incertidumbres en la práctica".

Para determinar la incertidumbre de una estimación que resulta de la suma de las fuentes independientes E_1 y E_2 , en que $E = E_1 + E_2$, se puede aplicar la ecuación A1.4 de propagación de errores.

$$\text{ECUACIÓN A1.4}$$

$$\sigma_E^2 = \sigma_{E_1}^2 + \sigma_{E_2}^2$$

Si las categorías de fuentes (o sumideros) se correlacionan, la ecuación A1.4 de propagación de errores no es válida y se debe aplicar la ecuación A1.5.

$$\text{ECUACIÓN A1.5}$$

$$\sigma_E^2 = \sigma_{E_1}^2 + \sigma_{E_2}^2 + 2r_{E_1 E_2} \sigma_{E_1} \sigma_{E_2} E_1 E_2$$

Cuando la suma excede de dos términos y hay covarianza, es preferible usar el método de Monte Carlo si se dispone de recursos.

A1.4.3.2 EL MÉTODO DE MONTE CARLO

Las técnicas estadísticas numéricas, en particular la técnica de Monte Carlo, sirven para estimar la incertidumbre en las tasas de emisión (a partir de incertidumbres en las medidas de actividad y los factores de emisión) cuando:

- las incertidumbres son grandes;
- no tienen una distribución gaussiana;
- los algoritmos son funciones complejas;
- existen correlaciones entre algunos de los conjuntos de datos de actividad, factores de emisión o ambos.

Las incertidumbres en los factores de emisión y/o datos de actividad a menudo son grandes y pueden no tener distribuciones normales. En esos casos, quizá sea difícil o imposible combinar incertidumbres por medio de las reglas estadísticas convencionales. Se puede utilizar el método de Monte Carlo, que consiste en hacer muchas veces los cálculos del inventario en una computadora electrónica, cada vez eligiendo al azar (por computadora) los factores de emisión o parámetros modelo y datos de actividad inciertos a partir de la distribución de incertidumbres que el usuario especifique al comienzo. Ese procedimiento genera una distribución de incertidumbres para la estimación del inventario que es coherente con las distribuciones de incertidumbres de entrada de los factores de emisión, parámetros modelo y datos de actividad. El método exige muchos datos y tiempo de computación, pero es muy adecuado para aplicar al problema de la propagación y totalización de incertidumbres en un sistema tan vasto como un inventario nacional de gases de efecto invernadero. Se encontrarán descripciones y aplicaciones en el anexo 3, “Glosario”, y en Bevington y Robinson (1992), Manly (1997), y Cullen y Frey (1999).

A1.4.4 Propagación de incertidumbres en todo el inventario

La tarea de propagación de incertidumbres en el inventario, después de estimar las incertidumbres individuales para cada clase de emisión, es más sencilla que la tarea de propagación de incertidumbres en algoritmos, porque las emisiones y absorciones sólo se totalizan por adición y sustracción.

En la agregación de incertidumbres tienen lugar dos procesos diferentes. Uno es la totalización de las emisiones de un solo gas que obedece a las reglas de propagación de incertidumbres ya examinadas. El otro es la agregación de incertidumbres de varios gases. En este caso, las emisiones y absorciones se deben reducir a una escala común, para lo cual se emplean los potenciales de calentamiento atmosférico (PCA). Sin embargo, para los óxidos de nitrógeno (NO_x), monóxido de carbono (CO) y compuestos orgánicos volátiles (COV) no existen PCA aceptados por el IPCC. Por consiguiente, las emisiones y absorciones de esos gases no se pueden incluir las incertidumbres totalizadas para un inventario de emisiones. Además, se debe tener presente que los valores de PCA tienen asociada una incertidumbre bastante importante, que se debe tener en cuenta al hacer una evaluación científica general de la emisión total equivalente.

Como algunas de las variables para agregar no son gaussianas, tienen grandes varianzas y se correlacionan con otras variables, Monte Carlo es el método de preferencia para la totalización de las incertidumbres. La aplicación de este método a los cálculos de incertidumbres en inventarios se presenta en el capítulo 6, “La cuantificación de incertidumbres en la práctica”.

Como una aproximación de trabajo, existe la opción de estimar la incertidumbre general de un inventario por medio del teorema del límite central (Cullen y Frey 1999). Las suposiciones relativas al teorema del límite central son las siguientes:

- el número de términos de emisión y absorción es grande;
- ninguno de los términos domina la suma;
- las emisiones y absorciones son independientes.

Si eso se cumple, la suma de las varianzas de todos los términos equivale a la varianza de todo el inventario y la distribución de las emisiones totales es normal. Por lo tanto, el intervalo definido por aproximadamente dos desviaciones estándar a ambos lados de la media es el intervalo de confianza de 95% del inventario. Como se señaló más arriba, este enfoque es una aproximación grosera. Es uno de los métodos alternativos de agregación de incertidumbres que se pueden usar en el nivel 1 de un sistema de incertidumbres de un inventario. Este método se utiliza en la hoja de trabajo simplificada empleada para el análisis de incertidumbre descrito en el capítulo 6.

A1.4.5 Covarianza y autocorrelación

En el análisis que se explica a continuación, se supone que los cálculos se hacen por un procedimiento de Monte Carlo.

Las estimaciones de emisiones (o absorciones) de dos componentes del inventario se representan por las funciones $E_1(t)$ y $E_2(t)$ en que t es el año de la estimación del inventario. Las incertidumbres se representan por $\delta_1(t)$ y $\delta_2(t)$ respectivamente.

Existen por lo menos cuatro fuentes significativas de covarianza en el inventario general, que provienen de:

- la utilización de datos de actividad comunes en varias estimaciones de emisiones (como ocurre en la serie de gases procedentes de una combustión);
- las limitaciones mutuas en un grupo de estimaciones de emisiones (tales como un determinado consumo total de combustible o de producción total de estiércol que se emplean en diversos procesos);
- la evolución de actividades y factores de emisión relacionados con nuevos procesos y tecnologías, etc. que disocian las incertidumbres de un período y del siguiente;
- los factores externos (económicos, climáticos, relativos a los recursos) que afectan a una serie de emisiones o absorciones.

Para el cálculo de incertidumbres sólo interesa la covarianza entre las incertidumbres representadas por $\delta_1(t)$ y $\delta_2(t)$. Si bien existe covarianza entre $E_1(t)$ y $E_2(t)$ y esa covarianza es importante para comprender y proyectar emisiones y absorciones, no tiene una pertinencia fundamental para el problema de agregar las incertidumbres, etc. Por lo tanto, de esas cuatro fuentes de covarianza, las tres primeras son fundamentales para determinar incertidumbres. La primera fuente de covarianza, la utilización de actividades comunes en una variedad de componentes de inventario, se produce en particular cuando se emiten varios gases en un mismo proceso, tal como la quema de combustibles fósiles o de biomasa. El uso de la misma actividad en dos estimaciones diferentes resultará en una covarianza positiva entre las dos estimaciones. Una manera eficaz de eliminar esa fuente de covarianza es combinar las ecuaciones en una única fórmula, que tenga una sola actividad y la suma de varios factores de emisión (expresados en equivalentes del CO_2).

El segundo tipo de covarianza se presenta cuando existe una limitación mutua en un conjunto de actividades o factores de emisión, en que se ingresan los datos de una actividad total y se indican proporciones para cada rama de tratamiento a fin de dividir esta actividad entre varios procesos de emisiones y algoritmos. Un ejemplo es la repartición de estiércol animal entre diferentes sistemas de manejo. En ese caso, el sistema puede resultar sobreespecificado si todas las proporciones y sus incertidumbres se resuelven simultáneamente. El método apropiado para eliminar la covarianza es dejar una de las proporciones sin especificar, y determinarla por la diferencia entre las otras proporciones y la fracción total. De ese modo, se elimina la necesidad de especificar la correlación de otros términos con el componente residual. Sin embargo, si existen correlaciones entre las proporciones especificadas o entre las proporciones especificadas y la actividad total, éstas se deben cuantificar y utilizar en los cálculos de propagación de incertidumbres.

El tercer tipo de covarianza surge cuando nuevas técnicas de medición, nuevos métodos de registro de datos o nuevas tecnologías eliminan las incertidumbres existentes e introducen otras nuevas, reduciendo el grado de autocorrelación de las series a lo largo del tiempo. Las autocorrelaciones serán elevadas mientras la tecnología, las técnicas de medición y la recopilación de estadísticas no cambien, y serán bajas cuando éstas cambien. La ingeniería y las ciencias sociales pueden contribuir copiosa información a esos índices de cambio (Grübler *et al.*, 1999). Ahora que los registros de inventarios nacionales tienen casi un decenio de existencia, es necesario analizar esas covarianzas.

A1.4.6 Compilación sistemática de incertidumbres en componentes de inventarios

En las secciones anteriores se presentaron los elementos fundamentales que deben tener las *buenas prácticas* en la determinación de la incertidumbre de la emisión o absorción de un gas de efecto invernadero en un inventario. Se resumen en el recuadro A1.2.

Los cuadros estándar del IPCC que se usan en la presentación de informes se deben revisar para poder incluir información sobre incertidumbres. En los cuadros sinópticos, la información registrada podría limitarse a intervalos de confianza con límites de 2,5% y 97,5% respectivamente. Se debería registrar la información

completa que se indica en los recuadros A1.1 y A1.2. La práctica del análisis de incertidumbre en inventarios se describe en detalle en el capítulo 6, “La cuantificación de incertidumbres en la práctica”.

A1.5 APLICACIONES

A1.5.1 Importancia de las diferencias de un año a otro y las tendencias en los inventarios

Un componente importante del análisis de incertidumbre en inventarios es la determinación de diferencias en las emisiones nacionales de un año a otro y a más largo plazo.

Si se toman en cuenta dos años, t_1 y t_2 , de una serie temporal, la diferencia entre las emisiones totales de esos años se puede representar, utilizando los símbolos definidos en la sección A1.4.5, por:

ECUACIÓN A1.6

$$\Delta E(t_1 - t_2) = E(t_1) - E(t_2)$$

y la varianza de la diferencia se define por:

ECUACIÓN A1.7

$$\sigma_{\Delta E}^2 = \sigma_{E_1}^2 + \sigma_{E_2}^2 - 2 \text{cov}(\delta E_1, \delta E_2)$$

o bien,

ECUACIÓN A1.8

$$\sigma_{\Delta E}^2 = \sigma_{E_1}^2 + \sigma_{E_2}^2 - 2r_{\delta E_1, \delta E_2} \sigma_{E_1} \sigma_{E_2}$$

donde

$$E_1 = E(t_1)$$

$$E_2 = E(t_2)$$

Por lo tanto, si se conoce la autocovarianza o la función de autocorrelación de las incertidumbres estimadas del inventario, se puede determinar la importancia de las diferencias de un año a otro. (Nótese que el término autocovarianza es a la autocorrelación lo que la covarianza es a la correlación.) Para estimar la correlación de la incertidumbre entre distintos años del inventario total, se puede considerar la adición de dos series autocorrelacionadas que representen dos de los muchos componentes de la incertidumbre del inventario. La autocovarianza de las series combinadas incluye las autocovarianzas de los diferentes términos más un componente para tener en cuenta la covarianza retardada que existe entre los dos componentes del inventario. Para hacer una evaluación de más de dos términos, se recomienda emplear el análisis de Monte Carlo.

RECUADRO A1.2**PRINCIPALES ELEMENTOS DE LAS BUENAS PRÁCTICAS EN LA DETERMINACIÓN DE LA INCERTIDUMBRE DE LAS ESTIMACIONES DE EMISIONES Y ABSORCIONES**

- i) utilización de las observaciones disponibles y el dictamen de expertos para determinar la incertidumbre en las cantidades de entrada;
- ii) registro sistemático y transparente de esos datos de entrada;
- iii) examen de los datos disponibles sobre las emisiones para determinar si se ha hecho un muestreo representativo;
- iv) diseño de otros muestreos y revisión de parámetros, valores por defecto y algoritmos para categorías principales de fuentes si no se ha hecho un muestreo representativo;
- v) aplicación de *directrices sobre buenas prácticas* sectoriales al seleccionar una función de densidad de probabilidad para representar los datos;
- vi) evaluación de correlaciones significativas (covarianzas) entre las cantidades de entrada;
- vii) propagación de incertidumbres por el método de aproximación si éstas son pequeñas y tienen una distribución gaussiana; de lo contrario,
- viii) propagación de incertidumbres por el método de Monte Carlo si se dispone de recursos, y
- ix) registro de la incertidumbre.

Enting (1999) presentó un análisis similar de la incertidumbre en la tendencia de una cantidad a lo largo de un intervalo de tiempo determinado. A título de ejemplo, considérense las emisiones $E(t)$, $E(t + \Delta t)$, en dos años diferentes de una serie temporal, separados por Δt años. La varianza de la tendencia a lo largo de ese período de tiempo está dada por:

ECUACIÓN A1.9

$$\text{var}(\Delta E) = 2\sigma_E^2(1 - r_{\delta E}(\Delta t))$$

Ello demuestra que en el caso de incertidumbres estimadas que están positivamente autocorrelacionadas la incertidumbre en la tendencia de las emisiones es más pequeña que en el de las incertidumbres aleatorias de magnitud equivalente. Es necesario hacer estudios sobre autocorrelaciones de incertidumbres estimadas en inventarios, así como de correlaciones cruzadas de las incertidumbres estimadas dentro de un mismo año de inventario y entre años de inventarios subsiguientes para emisiones y absorciones conexas.

A1.5.2 Combinación ("empalme") de métodos

A medida que se haga la compilación de inventarios nacionales, en ciertos casos será necesario cambiar el algoritmo utilizado para calcular una emisión o absorción en particular, ya sea porque se conozca mejor la forma del algoritmo o porque se haya producido algún cambio en la disponibilidad de los datos de actividad. En esos casos, lo mejor es volver a calcular los inventarios de años precedentes utilizando los nuevos métodos. Si eso no es posible, habrá que combinar o "empalmar" estimaciones preparadas por métodos diferentes para formar una serie temporal coherente. La teoría estadística en que se basan las *buenas prácticas* se describe más adelante y la orientación práctica para aplicarla a los inventarios se encuentra en el capítulo 7, "Elección de la metodología y realización de nuevos cálculos". Las estimaciones de emisiones (o absorciones) hechas por cada uno de los dos métodos se representan respectivamente por las funciones $P(t)$ y $Q(t)$, en que t es el año de estimación del inventario. En cualquier año que se apliquen los dos métodos habrá una diferencia entre las estimaciones, y la tarea de "empalme" consiste en examinar esa diferencia. Hay tres posibilidades: las dos estimaciones de emisiones pueden diferir en una cantidad constante; pueden ser proporcionales entre sí, o bien, pueden estar relacionadas por una diferencia constante y por un término proporcional. En el caso que se analiza aquí, se supone que existe una diferencia casi constante. (Se puede hacer un análisis similar para los otros dos casos. En realidad, en el tercer caso conviene hacer una forma de análisis de regresión lineal.)

La incertidumbre en la diferencia entre las dos estimaciones de emisiones en el tiempo t se puede expresar como sigue:

ECUACIÓN A1.10

$$\text{Incertidumbre} = \delta \Delta_{P-Q}(t)$$

$$\text{donde } \Delta_{P-Q}(t) = P(t) - Q(t)$$

Lo ideal es determinar esa diferencia para muchos años, junto con la incertidumbre de la diferencia media teniendo en cuenta las incertidumbres en P y Q . La barra indica el promedio plurianual de la diferencia a lo largo de los años t_1 – t_2 y δ indica la incertidumbre de esta diferencia media. En este caso, se puede hacer una serie aceptable de estimaciones empalmado las series $P(t)$ y $Q(t)$, para lo cual se corrige $Q(t)$ hasta que llegue a $P(t)$, añadiendo $\overline{\Delta_{P-Q}}(t)$ promediado para el período t_1 a t_2 . Un cambio en la técnica de estimación puede producir una mejora o una disminución de la calidad de la estimación. Si se demuestra que $Q(t)$ representa una mejora, entonces se debería usar lo más posible $Q(t)$ corregido hasta llegar a $P(t)$. Es decir que $P(t)$ se debería usar hasta t_1 y, a partir de entonces, $Q(t) + \overline{\Delta_{P-Q}}(t)$. A la inversa, si se prefiere $P(t)$, se debería utilizar esa estimación hasta t_2 , etc.

En la práctica, en un inventario nacional se pueden plantear tres situaciones. Podría ocurrir que $P(t)$ y $Q(t)$ no coincidieran ningún año; pueden coincidir un número limitado de años que no sean adecuados para aplicar el proceso de afinado de la diferencia entre las dos series que se examina más arriba, o bien, pueden coincidir un número suficiente de años.

En los primeros dos casos, se necesita más información para determinar la eficacia del empalme. Se pueden usar distintos métodos, a saber:

- Identificar otros lugares (países) donde existan series temporales muy similares y utilizar esos datos para elaborar una estimación general o regional de la diferencia media $\overline{\Delta_{P-Q}}(t)$ recogiendo todos los datos disponibles hasta que $\delta \overline{\Delta_{P-Q}}(t)$ se reduzca a una incertidumbre aceptablemente pequeña o se agoten todas las fuentes de datos.
- Si, habiéndose agotado todas las fuentes de datos, $\delta \overline{\Delta_{P-Q}}(t)$ todavía supera el criterio límite, aceptar la serie temporal, teniendo presente que ésta, del principio al fin, contiene una incertidumbre suplementaria que proviene de la incertidumbre en la diferencia entre las dos series.
- Cuando los datos no se traslapen ni se disponga de datos de otras fuentes, será necesario emplear otras técnicas de empalme. Una posibilidad es utilizar técnicas de series temporales (Box y Jenkins, 1970) para hacer predicciones hacia adelante $P(t)$ y hacia atrás en el tiempo $Q(t)$ y determinar si, en los años que siguen o preceden inmediatamente al empalme, esas predicciones concuerdan con el otro conjunto de datos dentro del intervalo de confianza de 95%. Si así es, el empalme se puede aceptar; en caso contrario, habría que registrar una discontinuidad en las estimaciones de las emisiones (o absorciones). En ambos casos, la incertidumbre aplicada a lo largo de toda la serie temporal sería, como mínimo, la incertidumbre combinada que resulte de cada una de las estimaciones $P(t)$ y $Q(t)$.

En el capítulo 7, “Elección de la metodología y realización de nuevos cálculos”, se examinan métodos prácticos de empalme.

A1.5.3 Análisis de sensibilidad y establecimiento de prioridades de investigación para los inventarios nacionales

Teniendo en cuenta el objetivo de reducir las incertidumbres de los inventarios, las prioridades para los próximos trabajos de investigación se deberían establecer basándose en tres características principales:

- la importancia de la categoría de fuentes o sumidero;
- la magnitud de la incertidumbre en la emisión o absorción;

- el costo de los trabajos de investigación y el logro previsto, medido como una reducción general de la incertidumbre del inventario.

La importancia de la categoría de fuentes debería establecerse empleando los criterios descritos en el capítulo 7, “Elección de la metodología y realización de nuevos cálculos”. Entre las categorías de fuentes de igual importancia, se debería dar prioridad a las que tienen incertidumbres más grandes o un mayor efecto en la tendencia.

Para cada categoría de fuentes, las opciones de investigación dependerán de las causas de la incertidumbre. En la mayoría de los casos, hay un número de variables que determinan la actividad y el factor de emisión. Se debería dar prioridad a las cantidades que tienen mayor incidencia en la incertidumbre general. Entre las opciones de investigación, una mayor estratificación de las emisiones y absorciones puede resultar muy ventajosa. De hecho, muchos valores por defecto que se usan actualmente están definidos para una amplia variedad de condiciones, lo cual se traduce necesariamente en grandes intervalos de confianza.

En el contexto actual, los costos de investigación comprenden costos financieros, insumo de tiempo y otros elementos que no siempre se pueden cuantificar.

Existen técnicas computacionales complejas para determinar la sensibilidad de un modelo (como puede ser un inventario) en cuanto a la relación entre las cantidades de salida y de entrada. Esos métodos se basan en la determinación de un coeficiente de sensibilidad, λ , que relaciona las emisiones agregadas E_T con una cantidad (o parámetro) de entrada, que en este caso se representa por a . Esos métodos determinan el coeficiente como sigue:

Ecuación A1.11

$$\lambda = \partial E_T / \partial a$$

Algunos programas informáticos del método de Monte Carlo tienen una opción para hacer ese análisis. Se ha aplicado a sistemas químicos de la atmósfera que entrañan decenas a cientos de reacciones químicas (NAS, 1979; Seinfeld y Pandis, 1998). Sin embargo, entre esos modelos químicos y los inventarios de gases de efecto invernadero hay una diferencia, que es el estado de los conocimientos. Los modelos químicos generalmente representan un sistema cerrado con conservación de masa, relaciones bien definidas y una serie de constantes de velocidad de reacción que en su mayoría han sido bien cuantificadas. La magnitud de las interacciones y los valores de las cantidades y parámetros de los inventarios de gases de efecto invernadero se conocen mucho menos.

Hay otros enfoques que pueden servir para determinar las prioridades de investigación y medición en la preparación de inventarios. Es posible elaborar métodos más simples basados en hipótesis generales que den pautas sobre las prioridades de investigación. La ventaja de esos sistemas más simples es que pueden ser utilizados por todos los compiladores de inventarios. La información sobre prioridades en materia de investigación y medición se basa en las evaluaciones de muestreos representativos que se examinan en la sección A1.4.2, “Muestreo representativo, algoritmos y covarianza”, el análisis de incertidumbre del capítulo 6, “La cuantificación de incertidumbres en la práctica” y el capítulo 7, “Elección de la metodología y realización de nuevos cálculos”, así como en la *orientación sobre buenas prácticas* que se dan para cada sector (véanse los capítulos 2 a 5). Esas diversas informaciones, combinadas con la opinión especializada de los compiladores de los inventarios son la mejor guía para determinar las prioridades en cuanto a la elaboración de inventarios.

A1.6 ASPECTOS QUE REQUIEREN INVESTIGACIÓN

Aunque algunos de los supuestos en que se basan los inventarios del IPCC son evidentes y ya han sido examinados, la investigación sistemática del conjunto de esos supuestos facilitaría la elaboración de un enfoque estructurado de la identificación de incertidumbres y la concepción de experimentos destinados a poner a prueba y afinar dichos supuestos. Esa labor comprende el examen de cuestiones de definición y las bases teóricas de los algoritmos de emisiones. Facilitaría la reunión de conocimientos y el intercambio de información entre los inventarios del IPCC y los estudios sobre los ciclos mundiales de gases en estado de trazas que forman parte de la tarea del Grupo de Trabajo 1 del IPCC, lo cual sería ventajoso para ambas actividades.

Un aspecto de la preparación de informes sobre emisiones y absorciones que aún falta resolver es el número de cifras significativas que se registran (precisión numérica). Según la ISO (1993), los valores numéricos de la estimación y su desviación estándar no deberían darse con un número excesivo de dígitos. Para el Inventario

Canadiense de Gases de Efecto Invernadero se ha adoptado la práctica de presentar los datos sólo con un número de dígitos acorde con la incertidumbre de las estimaciones de los inventarios. Si se pone cuidado en mantener esa relación en todo el inventario, es posible visualizar con claridad la incertidumbre de los valores y la diferencia entre las incertidumbres asociadas con las emisiones de cada categoría de fuentes. El otro enfoque es definir la unidad mínima que se debe presentar en los informes como una cantidad fija, de modo que los inventarios de todos los países y todos los componentes de esos inventarios se presenten con la misma unidad numérica. Desde un punto de vista práctico, es probable que este método ofrezca la ventaja de facilitar la auditoría de los cuadros; sin embargo, este asunto requiere un debate más detallado.

REFERENCIAS

- Bevington, P. R. y Robinson, D. K. (1992). *Data Reduction and Error Analysis for the Physical Sciences*. WCB/McGraw-Hill Boston, Estados Unidos, pág. 328.
- Bouwman, A.F. (1996). "Direct emission of nitrous oxide from agricultural soils". *Nutrient Cycling in Agroecosystems*, 46, págs. 53-70.
- Box, G.E.P. y Jenkins, G.M. (1970). *Time Series Analysis forecasting and control*. Holden-Day, San Francisco, Estados Unidos, pág. 553.
- Cochran, W.G. (1963). *Sampling Techniques*. 2ª edición, John Wiley & Sons Inc., Nueva York, pág. 411.
- Cullen, A.C. y H.C. Frey, H.C. (1999). *Probabilistic Techniques in Exposure Assessment*, Plenum Publishing Corp., Nueva York, Estados Unidos, pág. 335.
- Eggleston, S. (1993). Citado en IPCC (1996) *Revised Guidelines for National Greenhouse Gas Inventories*, op. cit.
- Enting, I.G. (1999). *Characterising the Temporal Variability of the Global Carbon Cycle*. CSIRO Technical Paper No 40, CSIRO Aspendale, Australia, pág. 60.
- Grübler, A., N. Nakićenović, N. y Victor D.G. (1999). "Dynamics of energy technologies and global change", *Energy Policy*, 27, págs. 247-280.
- IPCC (1996). *Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories: Volumes 1, 2 and 3*. J.T. Houghton et al., IPCC/OECD/IEA, París, Francia.
- ISO (1993). *Guide to the Expression of Uncertainty in Measurement* [Guía para la expresión de la incertidumbre en las mediciones]. International Organisation for Standardization, ISBN 92-67-10188-9, ISO, Ginebra, Suiza, pág.101.
- Mandel, J. (1984). *The Statistical Analysis of Experimental Data*. Dover Publications New York, Estados Unidos, pág. 410.
- Manly, B.F.J. (1997). *Randomization, Bootstrap and Monte Carlo Methods in Biology*. 2ª edición, Chapman & Hall, pág. 399.
- NAS (1979). *Stratospheric Ozone Depletion by Halocarbons: Chemistry and Transport*. Panel on Stratospheric Chemistry and Transport, National Academy of Sciences, Washington D.C., Estados Unidos, pág. 238.
- Robinson, J.R. (1989). "On Uncertainty in the Computation of Global Emissions for Biomass Burning". *Climatic Change*, 14, págs. 243-262.
- Seinfeld, J.H. y Pandis, S.N. (1998). *Atmospheric Chemistry and Physics*. John Wiley and Sons, Nueva York, Estados Unidos, pág. 1326.
- Subsidiary Body for Scientific and Technological Advice (SBSTA), United Nations Framework Convention on Climate Change (1999). National Communications from Parties included in Annex 1 to the Convention, Guidelines for the Preparation of National Communications, Draft conclusions by the Chairman. FCCC/SBSTA/1999/L.5, pág. 17.