
Informe final

Contrato N° IC/2015/092

“Consultor de apoyo para el diseño lógico y físico de bases de datos y su implementación para el desarrollo del Sistema Integrado de Cuentas Ambientales y Económicas en el MMA”

Consultor: Sergio Marchant Acevedo
16 de diciembre de 2015

Contenido

Introducción.....	3
1. Diseño de bases de datos para cuentas ambientales	4
1.1. Diseño modelo de datos de bosques y uso de suelos	4
1.2. Diseño modelo de datos de emisiones al aire.	6
1.3. Modelo de datos para gastos de protección ambiental.....	7
1.4. Diseño base de datos de las cuentas económicas.....	8
1.5. Diseño modelo de datos de directorio de actividades	9
1.6. Diseño base de datos de experiencias internacionales	10
1.7. Diseño base de datos documental	11
1.8. Diseño base de datos para cuentas de borrador o provisionales	11
1.9. Diseño base de datos de la Matriz Insumo Producto	13
2. Consultas a las bases de datos para cuentas ambientales	14
3. Procesos automatizados para captura de insumos informativos de cuentas ambientales	15
4. Nueva versión y carga de base de datos del directorio RETC y base documental	22
4.1. Directorio RETC.....	22
4.2. Base documental	23
Anexo 1. Matriz de entidades de los modelos de datos para cuentas ambientales	25

Indice de gráficos

Gráfico 1. Modelo de datos de cobertura y uso de suelos	4
Gráfico 2. Modelo de datos de bosques.....	5
Gráfico 3. Modelo de datos de emisiones.....	6
Gráfico 4. Modelo de datos de gastos de protección ambiental	7
Gráfico 5. Modelo de datos de base económica	8
Gráfico 6. Modelo de datos del directorio.....	9
Gráfico 7. Modelo de datos de experiencias internacionales	10
Gráfico 8. Modelo de datos base documental.....	11
Gráfico 9. Modelo de datos del RETC	12
Gráfico 10. Modelo de datos de la Matriz de Insumo-Producto	13

Indice de cuadros

Cuadro 1. Consultas a las bases de datos para las cuentas ambientales. Parte 1	15
Cuadro 2. Cobertura de establecimientos del directorio RETC	23

Introducción

Este informe responde a las bases del contrato N° IC/2015/092, cargo “Consultor de apoyo para el diseño lógico y físico de bases de datos y su implementación para el desarrollo del Sistema Integrado de Cuentas Ambientales y Económicas en el MMA”, Proyecto N° 72272 “Marco de Cooperación entre el Ministerio del Medio Ambiente y el PNUD”, suscrito con fecha 16 de septiembre de 2015, entre el Programa de Naciones Unidas para el Desarrollo (PNUD), y el consultor que informa.

El presente, es el informe final de consultoría y responde a los componentes de los términos de referencia del contrato:

- i. Elaborar procesos automatizados para la captura de información económica y ambiental vinculada al SCAE y el desarrollo de otros instrumentos de documentación.
- ii. Apoyar la elaboración del diseño lógico y físico e implementación de bases de datos para la elaboración de cuentas ambientales y económicas vinculadas.
- iii. Apoyar la elaboración de consultas y cálculos, a partir de bases de datos, para la generación de ejercicios de cuentas, cuadros analíticos e indicadores borradores.
- iv. Participar como contraparte técnica en aspectos de estructuración de bases de datos en proyectos del MMA relacionados con cuentas ambientales.

En el informe de avance se cubrieron los productos del componente i: modelos de datos y bases de datos diseñadas. La documentación de esos productos se repite en la primera sección de este informe final.

Los productos del componente ii, son archivos de consultas de las bases de datos. Estas consultas forman parte de la mayoría de las bases de datos presentadas en el informe de avance y el informe final del trabajo del consultor José Venegas en el marco del Proyecto N° 72272 “Marco de Cooperación entre el Ministerio del Medio Ambiente y PNUD”. En la sección 2 de este informe se hace una relación de estas consultas con una introducción de su enfoque de diseño general.

El componente iii consideraba la generación de procesos automatizados para capturar y cargar datos económicos y ambientales en bases de datos. Los productos de estas actividades son los archivos con las consultas a base de datos elaboradas, datos económicos y ambientales capturados. A los resultados de este trabajo se refiere la sección 3 de este informe.

Las actividades y productos del componente iv no se pudieron concretar en el transcurso del período de trabajo de esta consultoría. La convocatoria y realización de reuniones de trabajo donde este consultor participaría como contraparte técnica, no dependían de su gestión personal.

El uso de las horas profesionales que deberían ser dedicadas a ese componente, se reorientaron a dos actividades y productos relevantes para la implementación de las cuentas ambientales. Por una parte se cargó la totalidad de la base de datos del directorio RETC y se amplió la cobertura de cuentas y cuadros del manual en la base documental. El diseño de ambas bases se expuso en el informe de avance de esta consultoría. Este tema es abordado en la sección 4 final de este informe.

1. Diseño de bases de datos para cuentas ambientales

1.1. Diseño modelo de datos de bosques y uso de suelos

Para efecto de las cuentas, el bosque se define como el suelo con extensión superior a 0,5 hectáreas, con árboles de más de 5 metros de altura y un dosel de más del 10%, o árboles capaces de alcanzar in situ esos umbrales. El ámbito de la cuenta de bosques y otros suelos boscosos sigue una perspectiva de uso del suelo. Por tanto, no comprende la superficie predominantemente dedicada a usos agrícolas o urbanos, y no se define estrictamente sobre la base de los cambios de la superficie de árboles. A partir de esta situación, para fines analíticos se deben considerar dos modelos. El de uso de suelos que se muestra en el Gráfico 1 y el de bosques en el Gráfico 2.

Gráfico 1. Modelo de datos de cobertura y uso de suelos

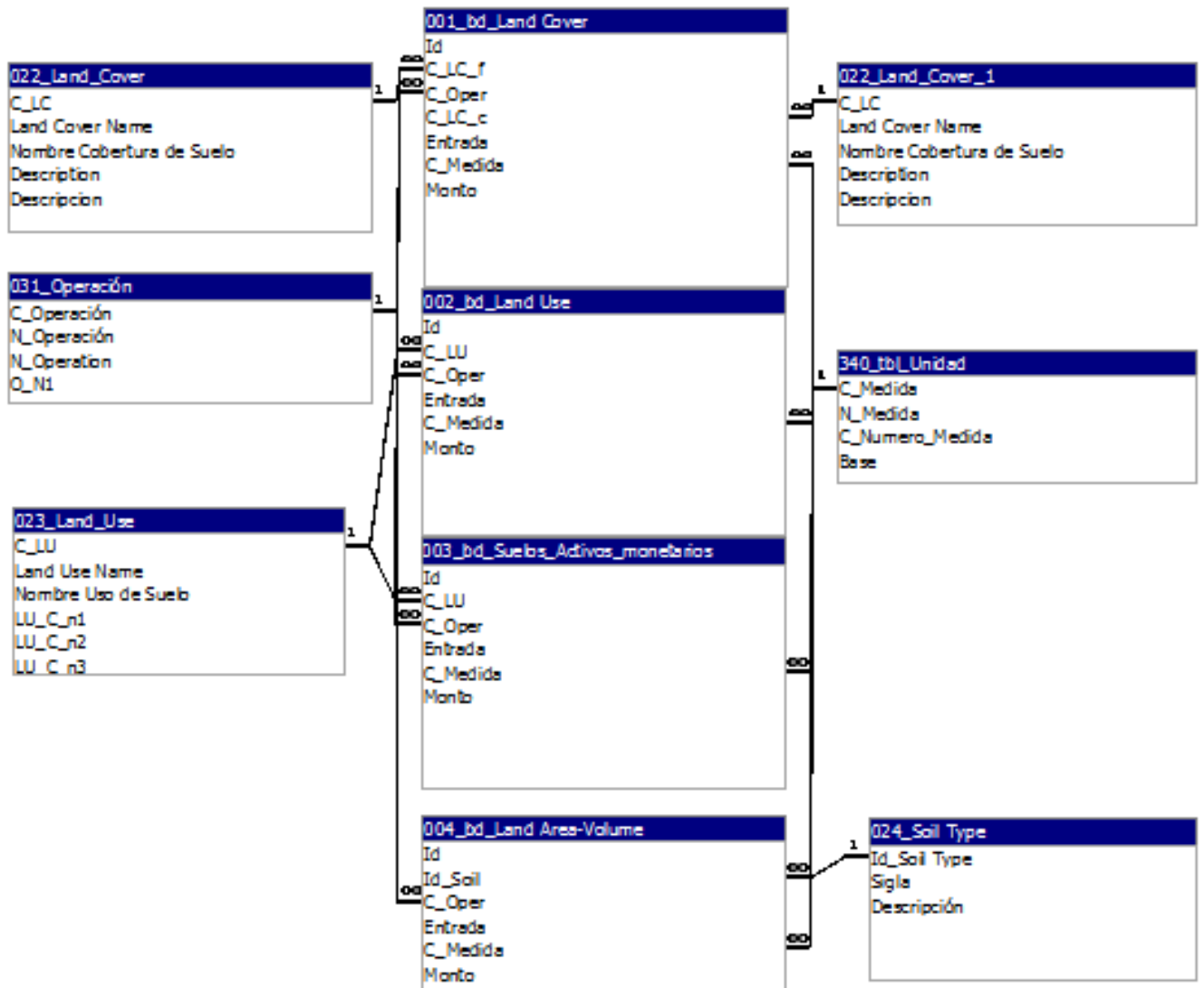
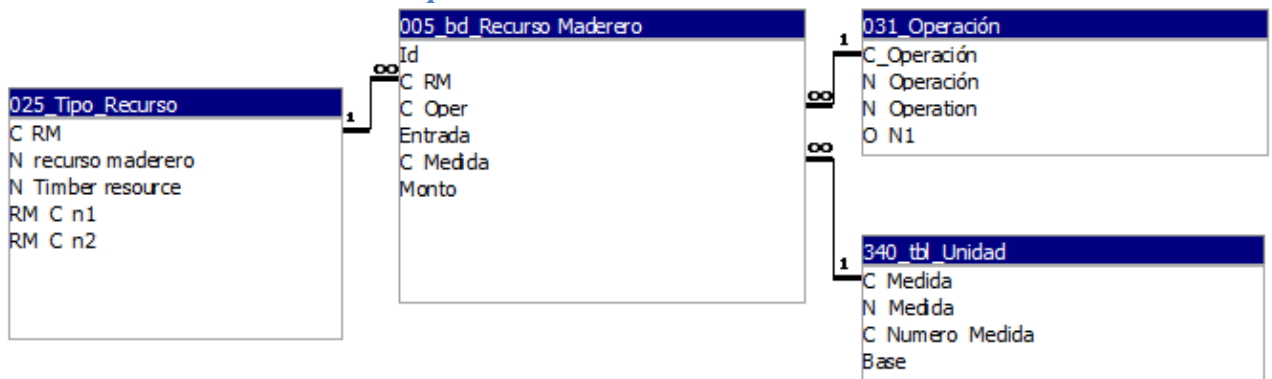


Gráfico 2. Modelo de datos de bosques



La tabla de uso de suelos y bosques, sólo tienen en común las clases de cobertura y uso de suelos relacionados con la silvicultura. El suelo se puede examinar de varias perspectivas, lo que define las tablas de hechos económicos y ambientales (HEA): cobertura

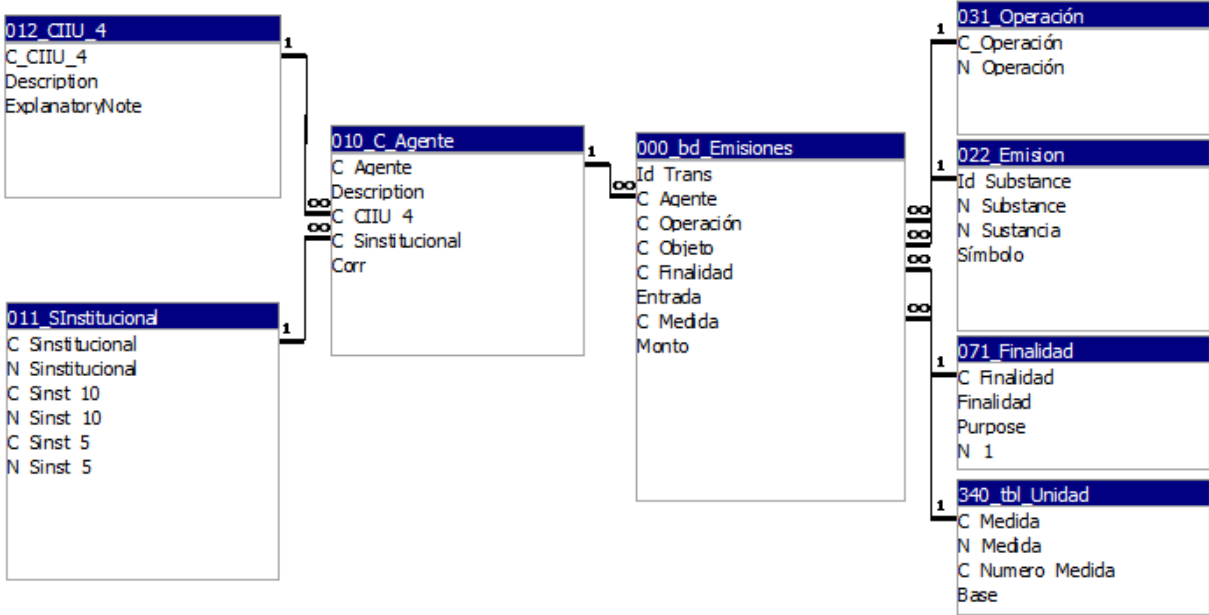
(001_bd_Land Cover), uso (002_bd_Land Use), medición en unidades monetarias (003_bd_Suelos_Activos monetarios) y medida en unidades físicas (004_bd_Land_Area-Volume). Estas tablas se complementan con las tablas que definen los registros de stocks y flujos (031_Operación) y las unidades de medida (340_tbl_Unidad) que se usan en la cuenta.

La base de bosques, en cambio es bastante más simple, con una sola base HEA, donde se registran las operaciones (031_Operación) de stocks y flujos de los recursos madereros (025_Tipo_Recurso) como objetos. En esta tabla se codifican los recursos madereros cultivados, naturales, naturales disponibles y no disponibles para la oferta de madera. Las unidades de medición (340_tbl_Unidad) es la misma tabla común a todas las cuentas ambientales.

1.2. Diseño modelo de datos de emisiones al aire.

El diseño de la base de datos se presenta en el gráfico 3.

Gráfico 3. Modelo de datos de emisiones



Las emisiones al aire son sustancias gaseosas y en partículas descargadas al aire por los establecimientos y los hogares como consecuencia de procesos de producción, consumo y acumulación. En el SCAE la contabilidad de las emisiones al aire registra la generación de esas emisiones por unidades económicas residentes y por tipo de sustancia.

Como se puede observar, el modelo considera básicamente dos tablas relevantes que determinan los hechos económicos y ambientales (000_bd_Emissiones). Una tabla de

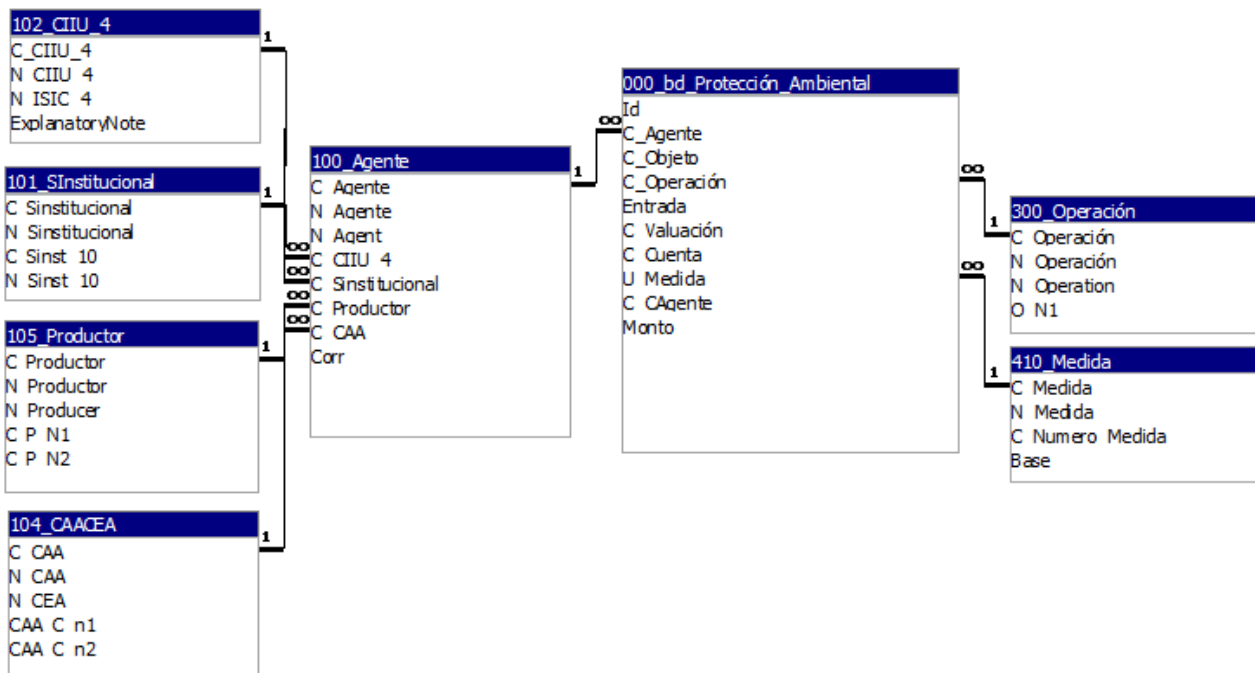
objetos, como son las sustancias (022_Emision) y una tabla de agentes (010_C_Agente) que considera los sujetos contaminantes.

La tabla de agentes considera tanto las actividades intermedias (productores) clasificadas por CIIU (012_CIIU_4) y hogares que es un sector institucional (011_SInstitucional). Como se trata de identificar la fuente de emisión, entonces para los hogares es necesario una clasificación de funcionalidad del gasto (071_Finalidad). La cuenta de emisiones es una cuenta de flujo (031_Operación) y que considera unidades físicas (340_tbl_Unidad).

1.3. Modelo de datos para gastos de protección ambiental

Las CGPA facilitan información sobre la producción de servicios específicos de protección ambiental producidos en toda la economía, y sobre el gasto en bienes y servicios con propósitos ambientales de las unidades residentes. Las actividades de protección ambiental son aquellas cuya finalidad principal es la prevención, la reducción y la eliminación de la contaminación y otras formas de degradación del ambiente. En el gráfico 4, se muestra el modelo de datos de gastos de protección ambiental.

Gráfico 4. Modelo de datos de gastos de protección ambiental



Esta es una cuenta en valores monetarios propio de cualquier sistema contable. Ese es el registro básico del hecho económico y ambiental (000_bd_Protección_Ambiental) de esta cuenta. El elemento ambiental lo proporciona la clasificación de actividades ambientales (104_CAA&CEA), que se emplea en el análisis del gasto. Además se consideran otras clasificaciones del agente, como la actividad económica (102_CIIU_4) que realiza el gasto

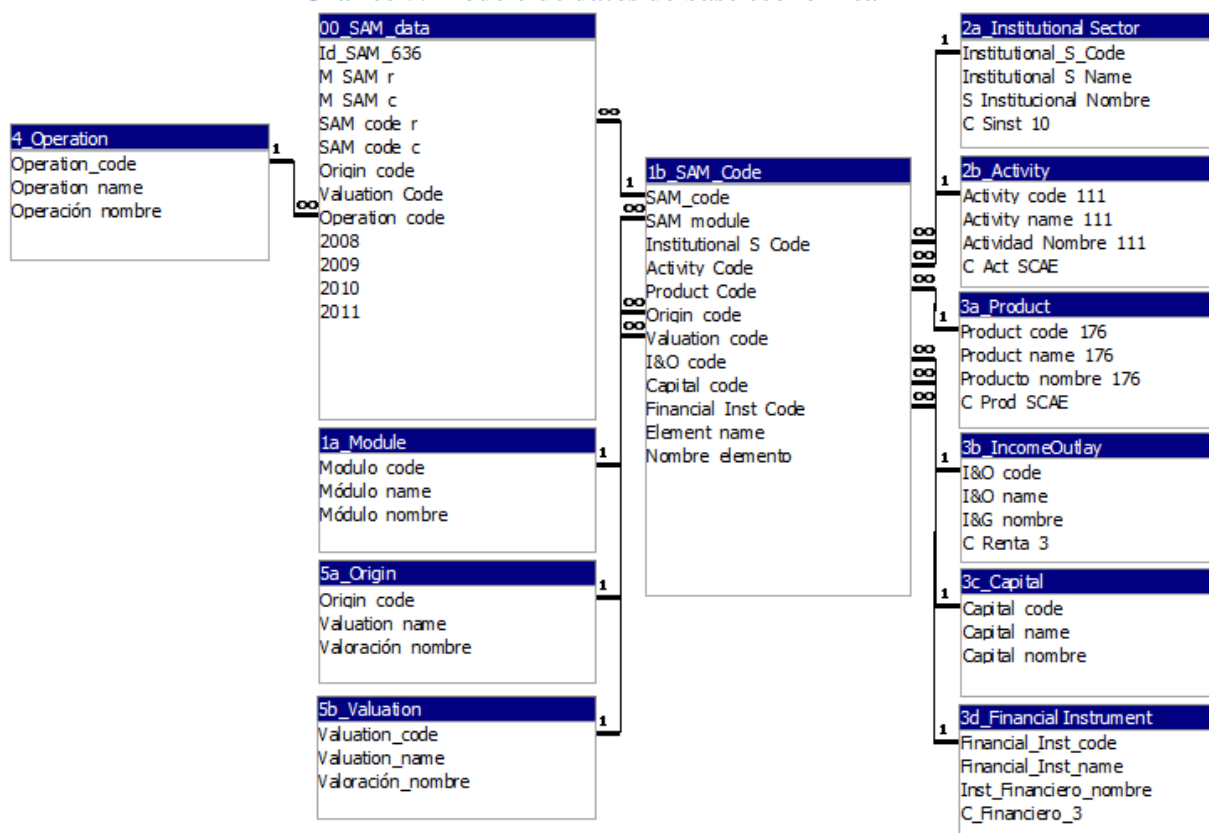
y el sector institucional (101_SInstitucional) que considera hogares y gobierno que realizan el gasto ambiental. Para efectos de clasificar al productor de servicios ambientales por su grado especialización se emplea una clasificación definida en el manual del SCAE (105_Productor).

La cuenta de gastos de protección ambiental por definición es una cuenta de flujos (300_Operación), y la medición se realiza en unidades monetarias (410_Medida).

1.4. Diseño base de datos de las cuentas económicas

Esta base de datos considera el conjunto de las cuentas nacionales, como organización sistemática y articulada de la información de un país. El diseño se organiza en torno a la matriz de cuentas social (Social Accounting Matrix, SAM), como herramienta que indexa y localiza el sistema completa de las cuentas nacionales de Chile. En el gráfico 5, se muestra el diseño de la SAM para Chile con registros correspondientes a las cuentas nacionales anuales de la serie 2008-2012.

Gráfico 5. Modelo de datos de base económica



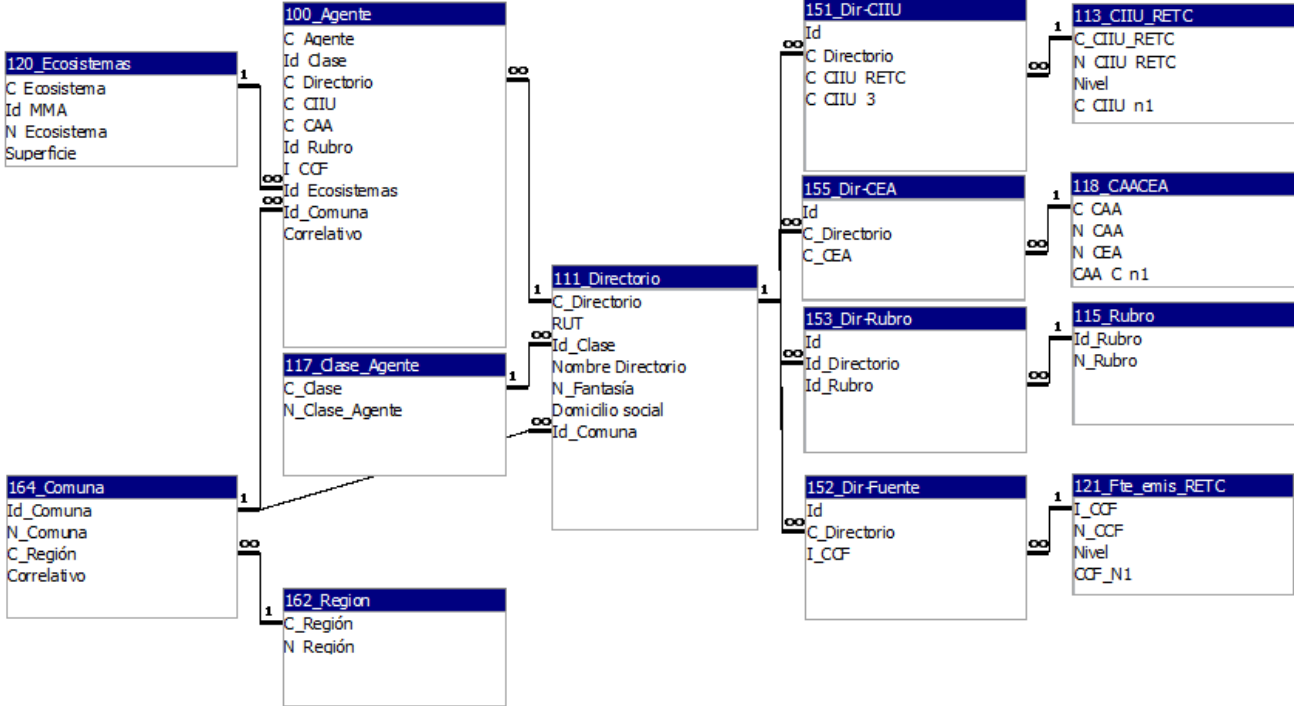
La base de hechos económicos (00_SAM_data) que registran las cuentas nacionales, comprenden una serie de agentes y objetos, que deben ser indexados en una tabla central

(1b_SAM_Code). A esa tabla convergen los agentes que en cuentas nacionales están definidos como sectores institucionales (2a_Institutional Sector) y ramas de actividad económica (2b_Activity). Por su parte los objetos están constituidos por los tipos de rentas (3b_Income&Outlay), de instrumentos de capital real (3c_Capital) y financiero (3d_Financial Instrument) Los bienes y servicios (3a_Product) deben distinguir su origen (5a_Origin) nacional o extranjero y la forma de valorización (5b_Valuation) que tienen en los registros.

1.5. Diseño modelo de datos de directorio de actividades

El directorio es un instrumento fundamental para identificar al agente específico o genérico en un hecho económico y ambiental. En el gráfico 6, se muestra la gran cantidad de elementos que definen a un agente en la práctica contable y estadística moderna.

Gráfico 6. Modelo de datos del directorio



Ante todo los agentes (100_Agente) pueden ser genéricos o individuales. Los agentes individuales se registran en un directorio (111_Directorio) de empresas y personas naturales.

Los agentes genéricos pueden ser actividades económicas (112_CIIU_Rev3 o 113_CIIU_RETC), pueden ser rubros (115_Rubro) usados en el RETC, actividades de protección y gestión ambiental (118_CAA&CEA). También se debe considerar un agente a una fuente de emisión (121_Fte_emis_RETC). En algunos casos el agente corresponde a una localización geográfica (162_Region o 164_Comuna).

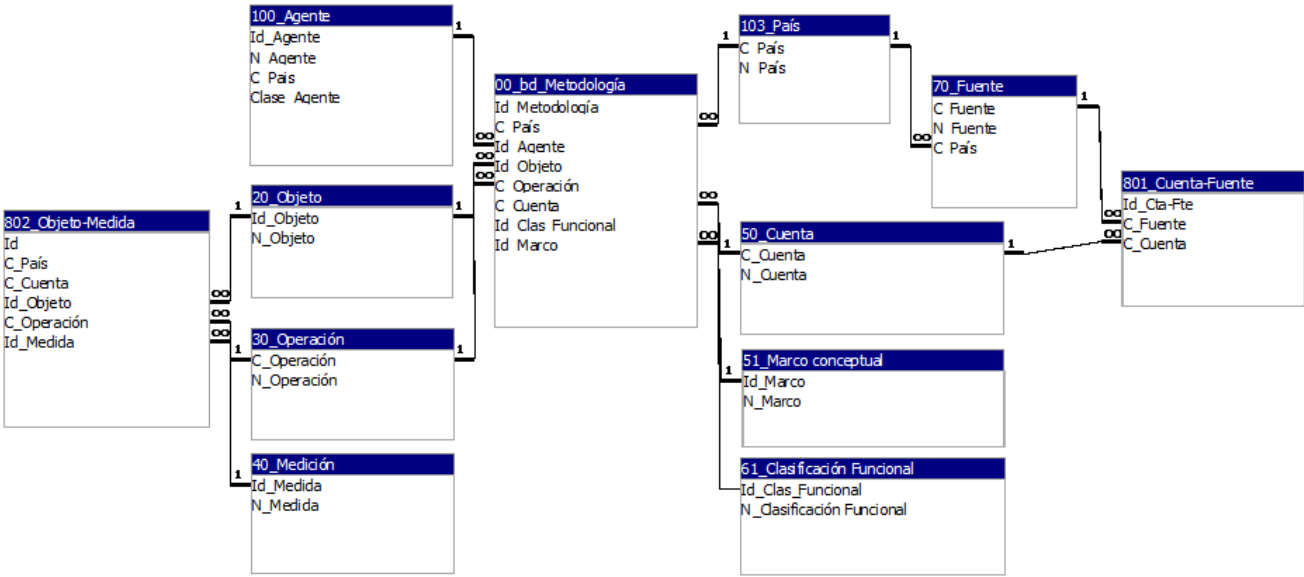
A su vez los agentes genéricos juegan el rol de clasificadores de los agentes individuales. El modelo de datos define esta pertenencia mediante tablas puente del tipo (151_Dir-CIIU). Esta tabla vincula el agente del directorio con la clasificación respectiva. Habrá tantas tablas puentes como agentes genéricos se definan.

Finalmente se debe dejar constancia de la futura incorporación de otros agentes genéricos que forman parte de bases de datos que deberán incorporarse al modelo de procesos de las cuentas ambientales. Es el caso del Clasificador de Actividades Económicas (CAE) que el Banco Central emplea en las cuentas nacionales, o del Sistema de Clasificación del Servicio de Impuestos Internos.

1.6. Diseño base de datos de experiencias internacionales

Esta base de datos procesa las fichas de datos de experiencias internacionales de metodologías de cuentas ambientales. En el Gráfico 7, se muestra el modelo diseñado que fue usado en una primera práctica para el análisis comparativo de las cuentas definidas como piloto para Canadá, Australia, México y Colombia

Gráfico 7. Modelo de datos de experiencias internacionales



Este modelo considera la metodología (00_bd_Metodología) de cada cuenta ambiental (50_Cuenta) en cada país (103_País) como tabla descriptora central. Al identificar la cuenta se debe individualizar al agente (100_Agente) y objeto (20_Objeto) que definen sus hechos económicos y ambientales.

Las cuentas registran stocks y/o flujos (30_Operación), pudiendo ser en unidades monetarias o físicas (40_Medición). Eventualmente pueden incluir clasificaciones funcionales del gasto o inversión (61_Clasificación Funcional).

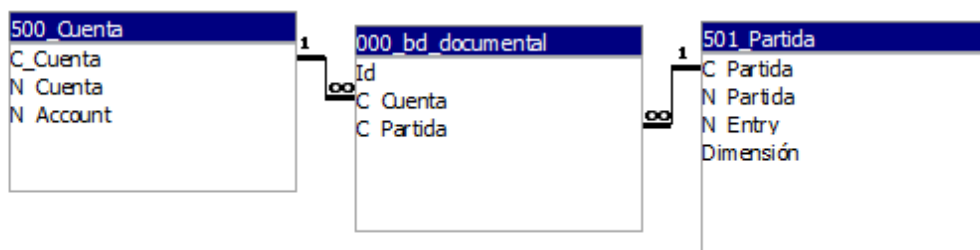
La elaboración de las cuentas está sujeta a un marco conceptual (51_Marco conceptual) que puede ser el SCAE o un manual anterior o especializado o una adaptación de cualquiera de ellos. Para su elaboración se recurre a información básica proveniente de instituciones o informes (70_Fuente) que deben ser identificados.

1.7. Diseño base de datos documental

La base documental contiene elementos de clasificaciones y tablas del SCAE-SEEA con sus términos en español e inglés. El objetivo es disponer de un apoyo metodológico para el trabajo de documentación de las cuentas ambientales a implementar, considerando que la literatura técnica está principalmente en inglés. En esta versión de la base sólo se han cargado términos español-inglés de un par de cuentas. En el próximo informe la base incorporará otras cuentas.

El modelo que se expone en el gráfico 8, es muy simple. Los registros de la base transaccional (000_bd_documental), consideran tres elementos básicos de clasificación:

Gráfico 8. Modelo de datos base documental



- Cuenta (500_Cuenta): Considera el código o numeración de la cuenta en el manual.
- Partida (501_Partida): Es el ítem o componente de la cuenta.
- Dimensión (componente de la tabla 501_Partida): Clasifica la partida por la fila o columna de la cuenta

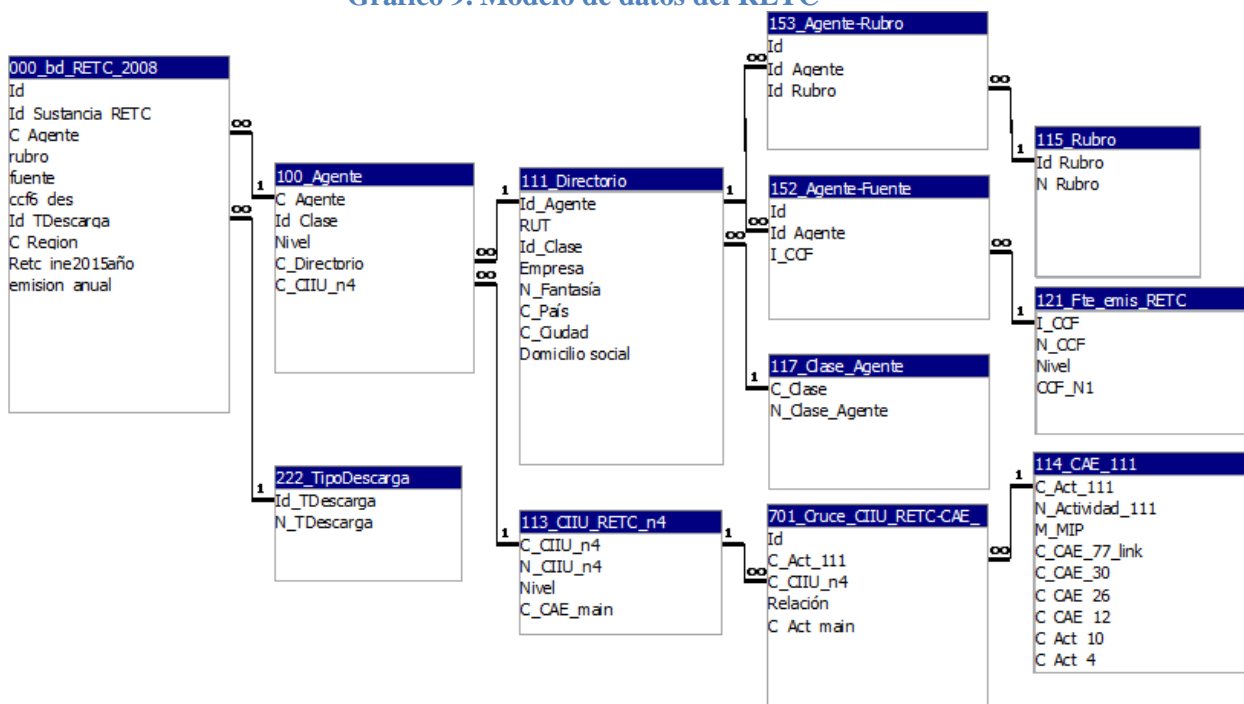
Así por ejemplo, en la tabla de emisiones al aire, las partidas son sustancias que van en la fila de la cuenta, en tanto la actividad o sector que provoca la emisión va en la columna de cuenta.

1.8. Diseño base de datos para cuentas de borrador o provisionales

Los elementos para el diseño de estas cuentas deben ser proporcionados por las consultorías en curso de las cuentas de bosques y uso de suelos, emisiones al aire y gastos de protección ambiental. Estas consultorías no han entregado los informes finales, que deben contener los elementos mencionados. No obstante, en paralelo a la ejecución de las consultorías, se han producido avances en el análisis de la información del RETC.

Los elementos del RETC constituirán el insumo básico para las cuentas provisionales de emisiones y gastos de protección ambiental. A continuación se expone el modelo de datos de la información del RETC para emisiones al aire. El diseño y la carga de la base de datos del RETC, se hizo sobre la base del archivo Access proporcionado por la unidad del RETC, agregado al nivel de CIUU. La base definitiva deberá considerar la información al nivel de empresa informante. La incorporación futura de la entidad empresa, en todo caso, no alterará la base del modelo que se expone en el gráfico 9.

Gráfico 9. Modelo de datos del RETC



La base de datos del RETC (000_bd_RETC), contiene los registros de agentes emisores (100_Agente) y sustancias emitidas (221_Sustancia) como elementos básicos. Los agentes contaminantes son las empresas y hogares. Las empresas se registran en el directorio (111_Directorio) del RETC.

Para efectos analíticos las empresas están sujetas a varias clasificaciones del agente. La principal es la clasificación de actividades (112_CIIU_Rev3). Las actividades se encuentran geográficamente localizadas en regiones (162_Region).

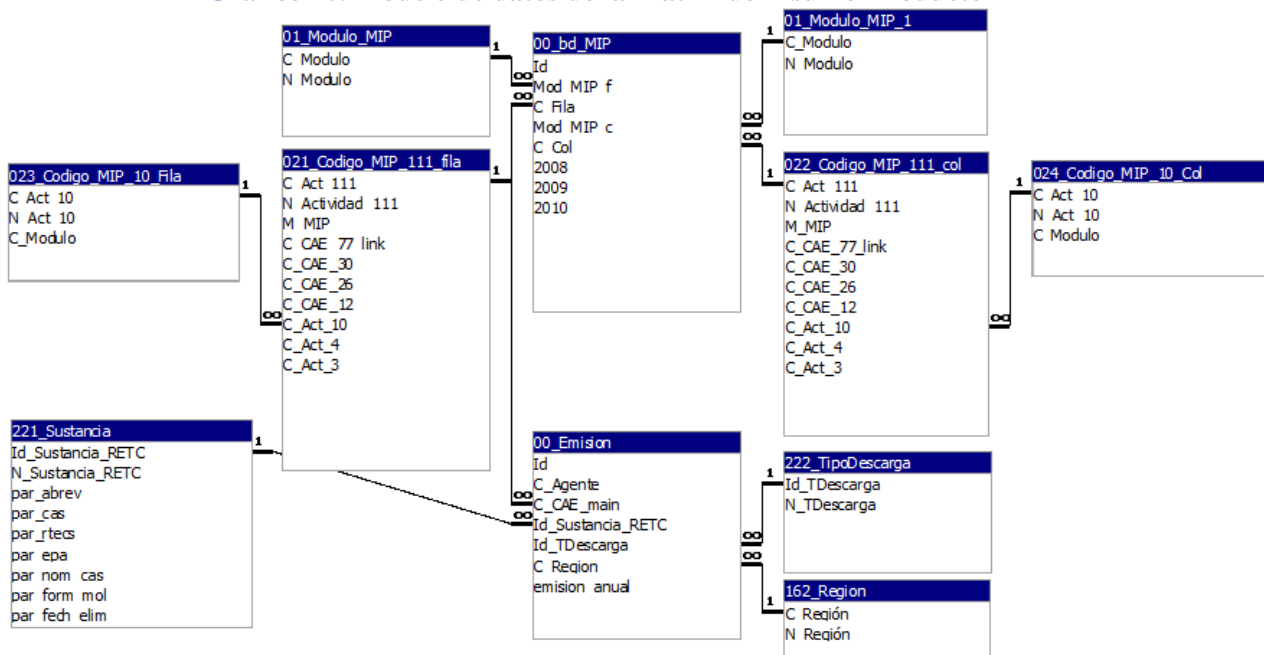
Como una forma de especificar las actividades en una visión más ambiental, se utiliza la clasificación de rubros (115_Rubro). Para una visión funcional de las emisiones, se utiliza la clasificación de fuentes de emisión (121_Fte_emis_RETC).

Como ya se dijo la clasificación del objeto de las emisiones es la sustancia que se emite. Otra clasificación de objeto muy importante en esta base de datos es el tipo de descarga (222_TipoDescarga).

1.9. Diseño base de datos de la Matriz Insumo Producto

La matriz de insumo-producto (MIP) es un cuadro de doble entrada donde la producción de un país se registra simétricamente, considerando los costos de producción en las columnas y utilización de bienes y servicios en las filas. En el gráfico 10 se muestra el modelo de datos de la MIP.

Gráfico 10. Modelo de datos de la Matriz de Insumo-Producto



La MIP es de gran importancia para las cuentas ambientales, ya que permite articular información ambiental con la información económica de la producción. Para ello se “extienden” los datos de las columnas a mediciones físicas. Esta matriz extendida se denomina matriz de cuentas nacionales y cuentas ambientales (NAMEA por su nombre en inglés).

Este modelo fue la base para cargar los datos de la Matriz de Insumo-Producto del Banco Central que 111x111 actividades y productos del año 2008, 2009 y 2010. Posteriormente estos datos fueron cruzados con la información de Emisiones del RETC, obteniéndose un formato de MIP extendida a emisiones año 2008. Las tablas del modelo relevantes para la MIP extendida se explican a continuación.

La base de hechos económicos o base transaccional (00_bd_MIP) contiene la información de las filas y columnas de la MIP. Por su parte la información ambiental se registra en la base transaccional de emisiones (00_Emision) del RETC. A su vez el objeto de la base de emisiones es la tabla de sustancias (221_Sustancia).

La información de la MIP del Banco Central considera 111 (021_Codigo_MIP_111) o 10 (021_Codigo_MIP_10) actividades. Esa desagregación es válida para el dato fila o el dato columna de la matriz.

Para fines de homogeneizar la clasificación las actividades económicas registradas en el RETC, se agregaron a la clasificación de 10 actividades de la MIP.

2. Consultas a las bases de datos para cuentas ambientales

El diseño de consultas en las bases de datos Access descritas en el primer capítulo de este informe tienen un enfoque común, consistente en:

- a. Una consulta Microsoft Access que filtra los datos relevantes para el análisis específico que quiera darle el usuario de la base de datos.
- b. El traspaso de la consulta a un libro Microsoft Excel, que toma la forma de una base de datos Excel en la primera hoja del libro.
- c. La creación de una o varias tablas dinámicas Excel que diseña el usuario para los distintos usos analíticos de la información filtrada.
- d. La transformación de la tabla en un cuadro que puede editarse para ser incluido en un informe.

Cada una de las fases de ese proceso requiere la intervención del analista de la información. No obstante por las funcionalidades de Microsoft Access y Excel, es posible transformar esas acciones en una consulta estructurada. Esto implica que ante la ocurrencia repetida del proceso en el futuro, ya no será necesario desarrollar cada uno de sus pasos.

Dados esos antecedentes en el Cuadro 1 se hace una relación de las consultas que fueron desarrolladas por esta consultoría en colaboración con José Venegas, consultor para el desarrollo de antecedentes técnicos y metodológicos para la implementación de un Plan Nacional de Cuentas Ambientales, del mismo proyecto PNUD 72272.

Cuadro 1. Consultas a las bases de datos para las cuentas ambientales. Parte 1

Base Access	Base de datos	Consulta	Producto
03_SCAEmc_Land	Cobertura y uso de suelos	T_5_13_Stocks_LC	Tabla 5.13 del SCAE
		T_5_14_Matriz_Cambios_LC	Tabla 5.14 del SCAE
		T_5_15_Forest	Tabla 5.15 del SCAE
		T_5_16_Suelo_Monetario	Tabla 5.16 del SCAE
		T_5_17&18_TipoSuelo	Tabla 5.17 del SCAE
03_SCAEmc_Timber	Bosques y áreas boscosas	T_5_19&20_Recursos madereros	Tablas 5.19 y 5.20 del SCAE
03_SCAE_emissions	Emisiones	Cta_OU	Tabla 3.7 del SCAE
03_SCAEmc_PA	Gastos de protección ambiental	Cuentas de Protección Ambiental	Tabla 4.2 del SCAE
		Cuentas de Protección Ambiental	Tabla 4.3 del SCAE
		Cuentas de Protección Ambiental	Tabla 4.4 del SCAE
		Cuentas de Protección Ambiental	Tabla 4.5 del SCAE
01_BDatos_CtasNac_Chile	Base económica	C_Activ-SCAE	Cuentas de Producción
		OOUU_SCN-SCAE	Cuentas de oferta y uso
08_Directorio_RETC_151215	Directorio	CIU-Establecimiento	Directorio RETC
05_bd_Experiencias_modelo	Experiencias internacionales	Varias tablas descriptivas	Varios
13_Base_documental	Base documental	SCAE_Cuenta-Partida	Tablas del SCAE
10_bd_MIP_2008-2010_Chile	Matriz de Insumo-Producto	bd_MIP_10x10	Matriz de Insumo-Producto
		Emisión_10x10	Emisiones por actividad

Cuadro 1. Continuación. Parte 2

Base Access	Consulta	Descripción
03_SCAEmc_Land	T_5_13_Stocks_LC	Cobertura del suelo en unidades físicas (uf)
	T_5_14_Matriz_Cambios_LC	Matriz de cambios en el uso del suelo (uf)
	T_5_15_Forest	Stocks y variaciones de cobertura de suelo (uf)
	T_5_16_Suelo_Monetario	Stocks y variaciones de cobertura de suelo (um)
	T_5_17&18_TipoSuelo	Activos de recursos del suelo (uf)
03_SCAEmc_Timber	T_5_19&20_Recursos madereros	Stocks y variaciones de recursos madereros (uf y um)
03_SCAE_emissions	Cta_OU	Cuenta de emisiones al aire (uf)
03_SCAEmc_PA	Cuentas de Protección Ambiental	Producción de servicios de protección ambiental (um)
	Cuentas de Protección Ambiental	Oferta y utilización de servicios de protección ambiental (um)
	Cuentas de Protección Ambiental	Gasto nacional en protección ambiental (um)
	Cuentas de Protección Ambiental	Financiamiento del gasto nacional en protección ambiental (um)
01_BDatos_CtasNac_Chile	C_Activ-SCAE	Cuenta de producción actividades base SCAE
	OOUU_SCN-SCAE	Cuentas de oferta y uso productos base SCAE
08_Directorio_RETC_151215	CIU-Establecimiento	Establecimientos RETC por CIU
05_bd_Experiencias_modelo	Varias tablas descriptivas	Agentes, cuentas, objetos, operaciones y fuentes
13_Base_documental	SCAE_Cuenta-Partida	Oferta y uso, stocks, gasto de protección ambiental
10_bd_MIP_2008-2010_Chile	bd_MIP_10x10	MIP 10 x 10 actividades
	Emisión_10x10	Sustancias de emisión para 10 actividades

El detalle de estas consultas se encuentra en las correspondientes bases de datos anexas a este informe.

3. Procesos automatizados para captura de insumos informativos de cuentas ambientales

Los insumos informativos de cuentas ambientales cuando son cuadros o tablas estadísticas de doble entrada, es posible capturarlos automáticamente para formatos de bases de datos que constituyen la plataforma del procesamiento de cuentas ambientales.

En este trabajo se han desarrollado tres procesos.

- Proceso específico de captura de información de las matrices de insumo-producto publicadas anualmente por el Departamento de Cuentas Nacionales del Banco Central de Chile (DCN-BCCH) “ver BD MIP 2012”.
- Proceso específico de captura de información de la matriz de producción que constituye la oferta de productos nacionales de los cuadros de oferta y uso (fuente DCN-BCCH) “ver BD Matriz prod 2012 (CCNN)”
- Proceso genérico de captura de información en formato de cuadro o tabla de doble entrada de cualquier fuente estadística “ver BD Matriz genérica”.

Los dos primeros procesos fueron desarrollados en forma experimental, para arribar finalmente al proceso genérico.

La gran ventaja del proceso genérico, es que constituye una solución en cualquier arreglo matricial para transformarlo en arreglo vectorial funcional a las tablas de bases de datos.

Los archivos Microsoft Access con los tres procesos se encuentran anexos a este informe.

A continuación se desarrollará el procedimiento para ejecutar la solución genérica. Se ha definido que el caso genérico considere una matriz de 200x200. Al nivel de los arreglos matriciales más conocidos, la matriz de producción de cuentas nacionales tiene una dimensión de 175 productos (filas) por 111 actividades (columnas)

Se debe tener presente que por problemas de performance la ejecución del programa SQL debe dividirse. Como la matriz máxima es de 200x200 se consideran sub matrices de 200x40, como se muestra en el Gráfico 11.

Gráfico 11. Dominio del arreglo matricial para carga en base de datos

	1	...	40	41	...	80	81	...	120	121	...	160	161	...	200
1															
2															
3															
⋮															
⋮															
198															
200															

I.- Ambiente Excel.

La información original debe cargarse en un archivo Excel

1. Crear archivo para ambiente Access. Esto significa que el archivo original debe quedar con una sola fila y una sola columna de encabezados con códigos y/o títulos, y con los valores correspondientes a cada cruce fila/columna.
2. Generar instrucciones SQL. Es más rápido escribirlas en Excel (por la gran cantidad) y luego pasarlas a cada consulta Access

II.- Ambiente Access.

4. Importar desde Access la matriz creada en el punto 1.
5. Crear consultas de fusión datos
Consulta1 PROCESO VECTOR Nro. 1 (COL 1-40)
... hasta...
Consulta5 PROCESO VECTOR Nro. 5 (COL 161-200)
6. Ejecutar consultas de creación y almacenamiento de archivo vector
Consulta11 CREAR VECTOR Nro. 1
... hasta...
Consulta55 CREAR VECTOR Nro. 5
7. Como producto del proceso anterior resultan los 5 vectores en forma de tablas
Convertir matriz en vector fila columna en la forma de 5 tablas
VECTOR Nro. 1
... hasta...
VECTOR Nro. 5
8. Anexar los archivos creados por cada consulta
Consulta111 ANEXAR VECTOR 1
... hasta...
Consulta555 ANEXAR VECTOR 5
9. Una vez creadas las consultas de fusión (proceso matriz a vector), se puede ejecutar la macro existente en Access, esta ejecuta todas las consultas y crea un vector único.
VECTOR 2012

III.- Consultas Access.

10. Matriz. Resultado de la fase 4 del proceso

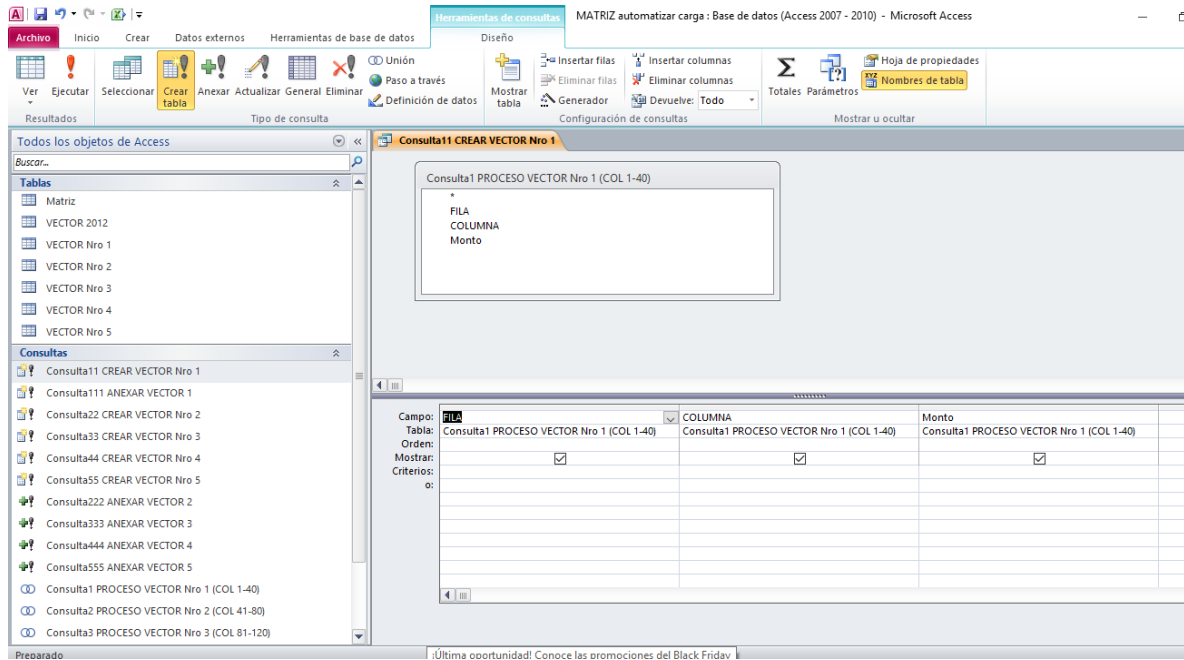
	A	B	C	D	E	F	G	H	I	J	DK	DL	DM	DN	DO	DP	DQ	DR	DS	DT	DU	DV	DW	DX	DY	EA	EB	EC	ED	EE	EF	EG	EH	EI	EJ		
1	FILA	1	2	3	4	5	6	7	8	9	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	
2	1	53.599	1.763	7	47	192.087	1.257	42.191	24.326	156		24.396	17.204	135.824	1.168.089	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
3	2	1.580	41.174	1.190	7.980	4.261	28	931	538	358		538	980	92.291	626.425	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
4	3	1	79	2	15	0			0	1		46.577		383.255	774.405	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	
5	4	5	535	18	106	0			11	1	5			766.566	1.420.622	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
6	5	3				10.308	978	5.631	2					-3.108	-4.502	954	823.775	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	
7	6													10.065	1.669	871	618.048	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	
8	7	14				8	4.285	24.708	9					894	5.795	15.651	832.887	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	
9	8	15				9	4.691	26.998	10					3.098	103	4.539	88.016	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	
10	9	105.140	4.741	42.610	46.627									1	169	0	16	425.724	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	
11	10	129.144	8.229	52.449	57.834	918	603	942	130	22				78.281	85.336	16.588	1.448.085	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	
12	11	0	0	0	0	0								3	395	17.056	4.479	1.591.083	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	
13	12	0	0	0	0									4	560	-4	39.394	698.874	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	
14	13		407											14	426	36	86	20.879	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	
195	194	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	76	
196	195	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	77	
197	196	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78	78
198	197	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79	79
199	198	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80	80
200	199	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81	81
201	200	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82	82

11. Vector, resultado de consulta indicado en fase 7.

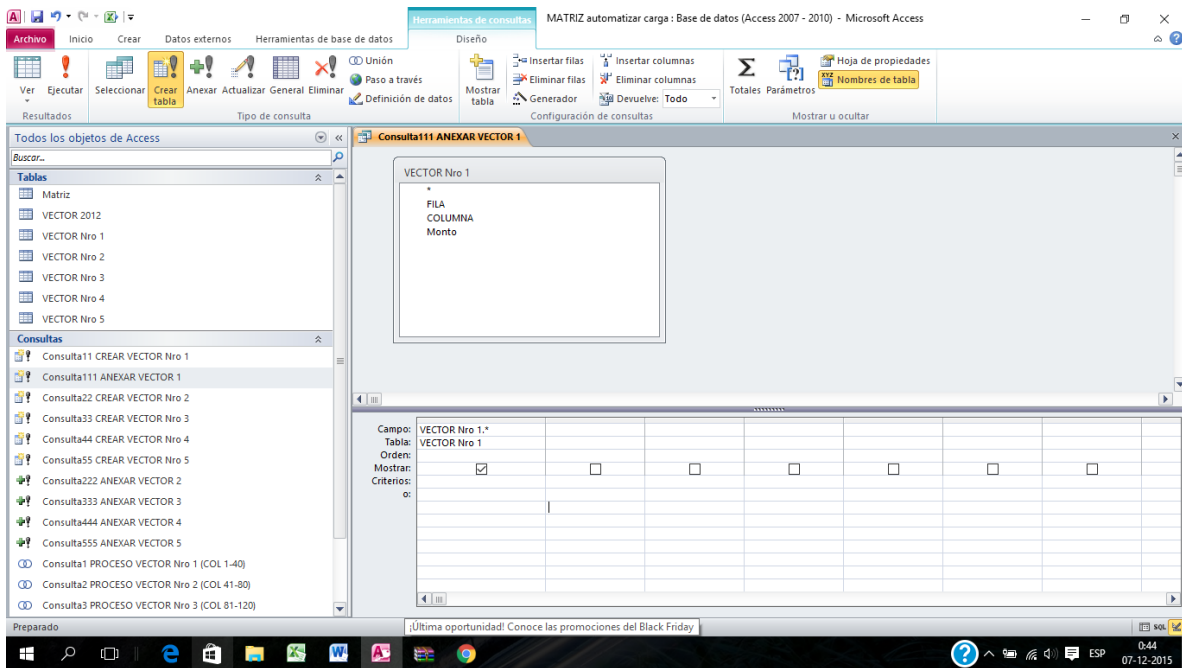
Fila	Columna	Monto
1	1	53.599,29
1	2	1.762,94
1	3	7,05
1	4	47,26
1	5	192.086,62
1	6	1.257,39
1	7	42.190,94
1	8	24.325,73
1	9	156,39
1	10	185,58
1	11	
1	12	345,48
1	13	
1	14	
1	15	
1	16	
1	17	
1	18	
1	19	618,92
1	20	
1	21	0,05
1	22	15.621,69
1	23	2.095,29
1	24	39,28
1	25	153.023,67
1	26	180.493,12
1	27	29.245,06
1	28	0,07
1	29	89.027,59
1	30	
1	31	
1	32	11.305,10
1	33	
1	34	8.186,90

1	35	2.276,22
1	36	
1	37	
1	38	0,00
1	39	0,05
1	40	

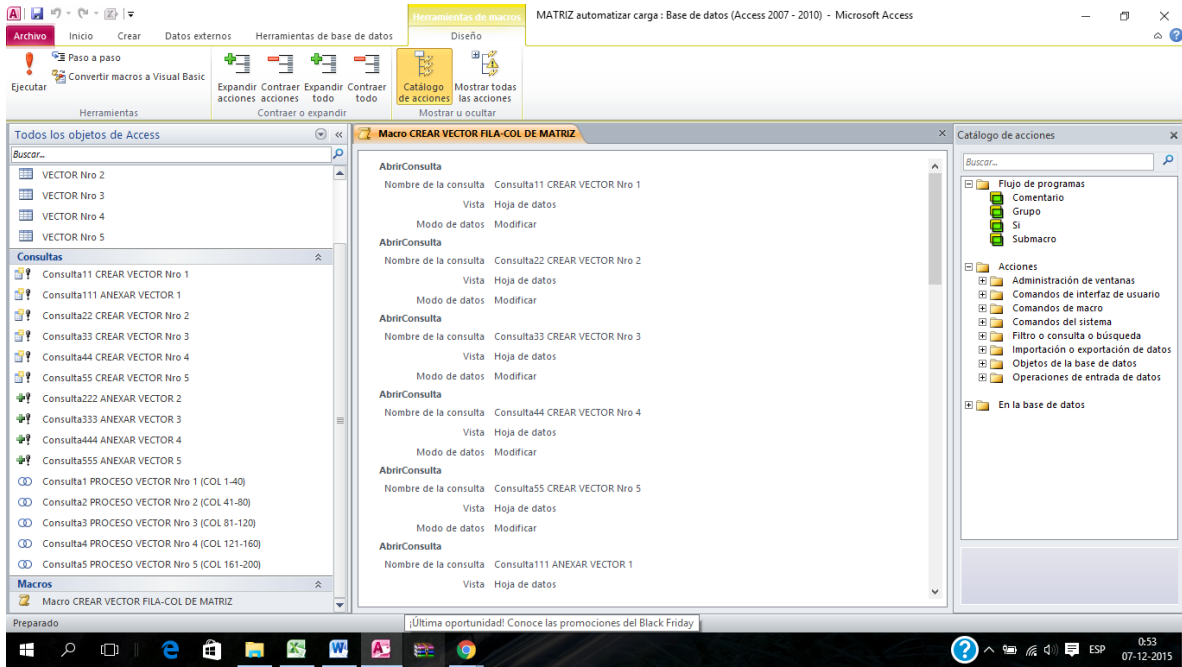
12. Consulta creación de vector (matriz 200 x 200 es necesario crear 5 consultas)



13. Consulta anexar vectores, se juntan los distintos vectores en uno solo (cada 40 col se debe generar una consulta)



13. Ejecutar Macro, entrega resultados completos hasta los vectores en uno solo anexados..



14. Instrucciones SQL, consulta fusión de datos, para convertir matriz en vector fila-columna.

```

SELECT [MATRIZ].FILA, "1" AS COLUMNA, [MATRIZ].[1] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "2" AS COLUMNA, [MATRIZ].[2] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "3" AS COLUMNA, [MATRIZ].[3] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "4" AS COLUMNA, [MATRIZ].[4] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "5" AS COLUMNA, [MATRIZ].[5] AS Monto FROM [MATRIZ];

```

UNION SELECT [MATRIZ].FILA, "6" AS COLUMNA, [MATRIZ].[6] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "7" AS COLUMNA, [MATRIZ].[7] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "8" AS COLUMNA, [MATRIZ].[8] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "9" AS COLUMNA, [MATRIZ].[9] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "10" AS COLUMNA, [MATRIZ].[10] AS Monto FROM [MATRIZ];

...

UNION SELECT [MATRIZ].FILA, "39" AS COLUMNA, [MATRIZ].[39] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "40" AS COLUMNA, [MATRIZ].[40] AS Monto FROM [MATRIZ];

FIN PRIMERA SUBMATRIZ

SELECT [MATRIZ].FILA, "41" AS COLUMNA, [MATRIZ].[41] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "42" AS COLUMNA, [MATRIZ].[42] AS Monto FROM [MATRIZ];

...

UNION SELECT [MATRIZ].FILA, "79" AS COLUMNA, [MATRIZ].[79] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "80" AS COLUMNA, [MATRIZ].[80] AS Monto FROM [MATRIZ];

FIN SEGUNDA SUBMATRIZ

SELECT [MATRIZ].FILA, "81" AS COLUMNA, [MATRIZ].[81] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "82" AS COLUMNA, [MATRIZ].[82] AS Monto FROM [MATRIZ];

...

UNION SELECT [MATRIZ].FILA, "119" AS COLUMNA, [MATRIZ].[119] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "120" AS COLUMNA, [MATRIZ].[120] AS Monto FROM [MATRIZ];

FIN TERCERA SUBMATRIZ

SELECT [MATRIZ].FILA, "121" AS COLUMNA, [MATRIZ].[121] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "122" AS COLUMNA, [MATRIZ].[122] AS Monto FROM [MATRIZ];

...

UNION SELECT [MATRIZ].FILA, "159" AS COLUMNA, [MATRIZ].[159] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "160" AS COLUMNA, [MATRIZ].[160] AS Monto FROM [MATRIZ];

FIN CUARTA SUBMATRIZ

SELECT [MATRIZ].FILA, "161" AS COLUMNA, [MATRIZ].[161] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "162" AS COLUMNA, [MATRIZ].[162] AS Monto FROM [MATRIZ];

...

UNION SELECT [MATRIZ].FILA, "199" AS COLUMNA, [MATRIZ].[199] AS Monto FROM [MATRIZ];
UNION SELECT [MATRIZ].FILA, "200" AS COLUMNA, [MATRIZ].[200] AS Monto FROM [MATRIZ];

FIN QUINTA SUBMATRIZ

Este procedimiento (o la macro) puede aplicarse a un cuadro de cualquier dimensión. Supongamos por ejemplo que se tiene un cuadro de 30 sustancias contaminantes (filas) y 50 actividades. Entonces sólo se tendrán 2 sub matrices de 30x40 y 30x10. El código desde la segunda matriz se cortará en la columna 10.

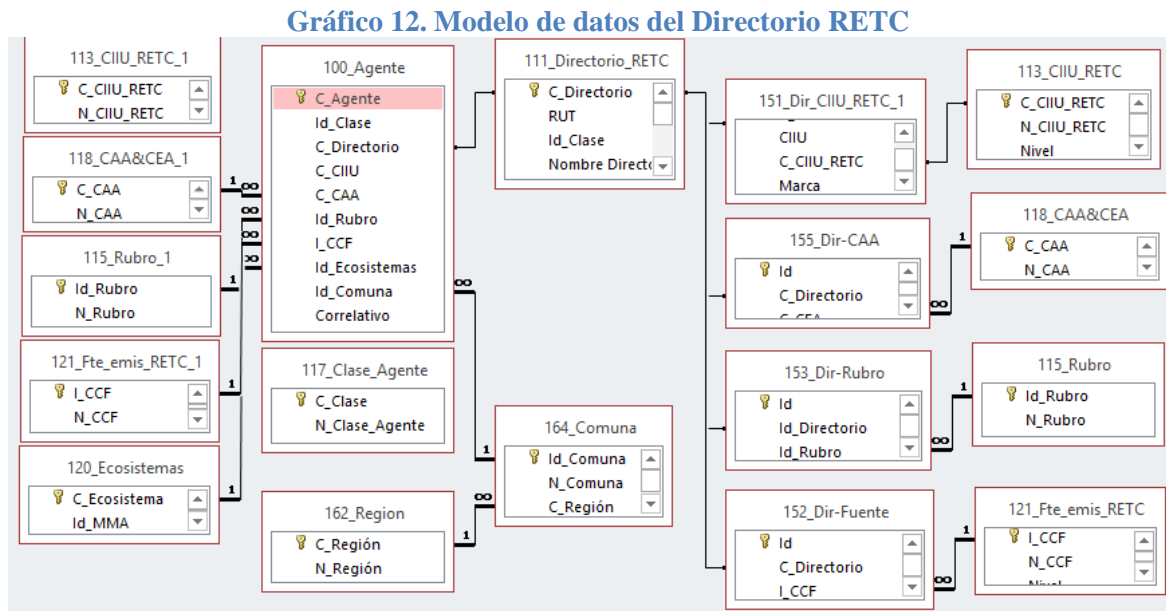
4. Nueva versión y carga de base de datos del directorio RETC y base documental

Con este informe se documenta el trabajo sobre el directorio y la base documental realizados durante el último mes de esta consultoría.

4.1. Directorio RETC

El diseño del directorio RETC no ha cambiado desde la versión previa. Se debe recordar que este directorio que reside en la base de datos Access “08_Directorio_RETC”, se creó a partir de los datos originales entregados en un archivo Excel por la unidad del RETC de la División de Información y Estudios de MMA.

El diseño de la base del directorio se reproduce en el Gráfico 12.



Este modelo ya fue explicado en el Informe Final de Consultoría de José Venegas (11-12-2015). No obstante, es necesario reiterar que:

- El agente puede ser genérico o individual (ambos tipos están representados en la tabla “100_Agente”).
- Cuando el agente es individual (unidad establecimiento) su detalle y atributos se registran en la tabla del directorio (“111_Directorio_RETC”).
- La pertenencia de un establecimiento a una clase de sector o actividad se registran en las tablas puente (*bridge tables*): “151_Dir_CIIU-RETC_1”, “155_Dir_CAA”, etc.

Los trabajos realizados este último mes consistieron en

- a) Elaboración de la tabla puente “151_Dir_CIIU-RETC_1”. El archivo Excel original mostraba sectores CIIU sin el código original CIIU Rev.3 que incluye letras, lo que dificultaba la clasificación.
- b) Carga de las 22.884 establecimientos que conforman el archivo RETC.
- c) Consulta de clasificación de los establecimientos por CIIU.

Precisamente de este último trabajo se obtuvo el Cuadro 2 siguiente:

Cuadro 2. Cobertura de establecimientos del directorio RETC

C_CIIU	N_CIIU_Rev.3. Nivel 1	Total
A	AGRICULTURA, GANADERIA CAZA Y SILVICULTURA	431
B	PESCA	359
C	EXPLOTACIÓN DE MINAS Y CANTERAS	155
D	INDUSTRIAS MANUFACTURERAS	4.087
E	SUMINISTRO DE ELECTRICIDAD, GAS Y AGUA	474
F	CONSTRUCCIÓN	628
G	COMERCIO AL POR MAYOR Y AL POR MENOR; REPARACIÓN DE VEHÍCULOS, EFECTOS Y ENSERES	3.134
H	HOTELES Y RESTAURANTES	146
I	TRANSPORTE, ALMACENAMIENTO Y COMUNICACIONES	1.441
J	INTERMEDIACIÓN FINANCIERA	771
K	ACTIVIDADES INMOBILIARIAS, EMPRESARIALES Y DE ALQUILER	761
L	ADMINISTRACIÓN PÚBLICA Y DEFENSA; PLANES DE SEGURIDAD SOCIAL DE AFILIACIÓN OBLIGATORIA	99
M	ENSEÑANZA	338
N	SERVICIOS SOCIALES Y DE SALUD	333
O	OTRAS ACTIVIDADES DE SERVICIOS COMUNITARIOS, SOCIALES Y PERSONALES	138
Q	ORGANIZACIONES Y ORGANOS EXTRATERRITORIALES	35
Z	Actividad no especificada	8.380
	Total clasificados	21.710
	Otros aún no clasificados	1.174
	TOTAL	22.884

4.2. Base documental

La base documental registra las cuentas y cuadros del SCAE 2012, en español y en inglés. Considera códigos y nombres de las cuentas y de las distintas partidas o componentes de cada cuenta.

Todas las cuentas o cuadros analíticos tienen la forma de arreglos matriciales, de modo que es fundamental la distinción de partidas fila y partidas columna.

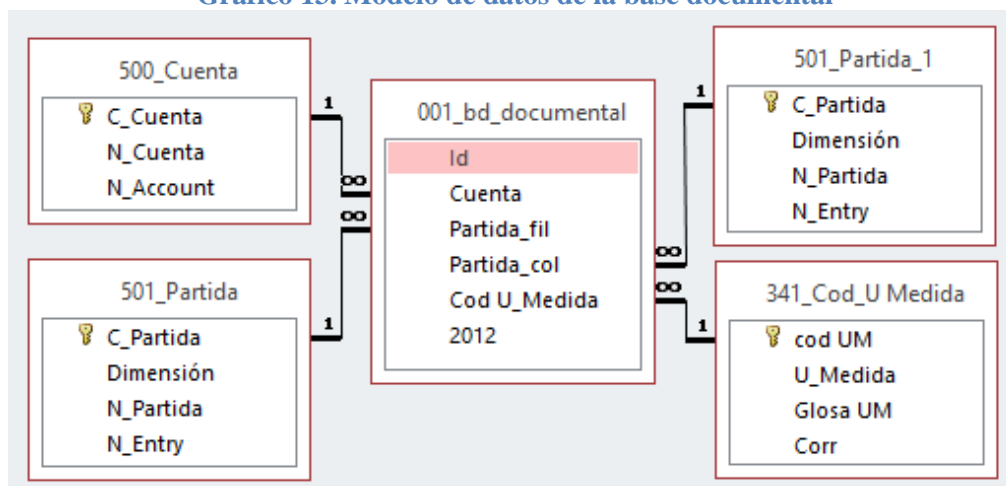
La presentación de las cuentas y cuadros es tal como se muestran en el SCAE. No existe un modelamiento de los datos como se ha expuesto en las diversas bases de datos en los informes previos. En este se trata simplemente del detalle de filas y columnas sin alterar

nombres. Incluso se registran cifras cuando en la cuenta o cuadro del SCAE se proporcionan ejemplos numéricos.

En el primer informe de avance se expuso el modelo de datos de la base documental. En esta entrega, el modelo se ha completado y se cargaron la mayoría de las cuentas del SCAE, como se verá a continuación.

El modelo de datos de la base documental se muestra en el Gráfico 13.

Gráfico 13. Modelo de datos de la base documental



Como se puede observar se trata de cuatro tablas en un enfoque estrella (*star schema*). La tabla central (“001_bd_documental”) registra todas las celdas de las distintas cuentas y cuadros. La tabla cuenta (“500_Cuenta”) detalla las cuentas y cuadros del SCAE. La tabla de partida (“501_Partida”) se repite para la fila y la columna (redundancia conceptual no física), como corresponde a una matriz. Finalmente la tabla de unidad de medida (341_Cod_U Medida”) define si la partida está en unidades monetarias o físicas. Si son físicas registra la unidad que aparece en la cuenta o cuadro del manual.

Se han cargado las partidas correspondientes a 25 cuentas o cuadros del SCAE. Pendientes por cargar se encuentran las tablas del SCAE:

- 4.3 Oferta y utilización de servicios específicos de protección ambiental
- 4.5 Financiamiento del gasto nacional en protección ambiental
- 4.6 Sector de bienes y servicios ambientales
- 4.9 Impuestos ambientales, por tipo
- 5.10 Asignación del ingreso y del agotamiento de los recursos minerales y energéticos

Anexo 1. Matriz de entidades de los modelos de datos para cuentas ambientales

Cuenta o base de datos	Tabla (entidad)	Base HEA								
		A	B	C	D	E	F	G	H	I
Uso de suelos	001_bd_Land Cover	■								
	002_bd_Land Use									
	003_bd_Suelos_Activos monetarios									
	004_bd_Land_Area-Volume	■								
	022_Land Cover			■						
	023_Land Use			■						
	024_Soil Type									
	031_Operación								■	
340_tbl_Unidad				■	■	■	■			
Bosques	005_bd_Recurso Maderero	■								
	025_Tipo_Recurso			■						
	031_Operación								■	
	340_tbl_Unidad				■	■	■	■		
Emisiones	000_bd_Emisiones	■								
	010_C_Agente		■							
	011_SInstitucional		■							
	012_CIIU_4		■							
	022_Emision			■						
	031_Operación								■	
	071_Finalidad									■
340_tbl_Unidad				■	■	■	■			
Gasto de Protección ambiental	000_bd_Protección_Ambiental	■								
	100_Agente		■							
	101_SInstitucional		■							
	102_CIIU_4		■							
	104_CAA&CEA		■							
	105_Productor		■							
	300_Operación								■	
	410_Medida				■	■	■	■		
Base de datos de las cuentas económicas	00_SAM_data	■								
	1b_SAM_Code									■
	2a_Institutional Sector		■							
	2b_Activity		■							
	3a_Product			■						
	3b_Income&Outlay			■						
	3c_Capital			■						
	3d_Financial Instrument			■						
	4_Operation								■	
5a_Origin								■		
5b_Valuation					■					

Cuenta o base de datos	Tabla (entidad)	Base HEA								
		A	B	C	D	E	F	G	H	I
Directorio	100_Agente		■							
	111_Directorio		■							
	112_CIIU_Rev3		■							
	113_CIIU_RETC		■							
	115_Rubro		■							
	118_CAA&CEA		■							
	121_Fte_emis_RETC		■							
	162_Region		■							
	164_Comuna		■							
	Experiencias internacionales	00_bd_Metodología	■							
100_Agente			■							
103_País				■						
20_Objeto				■						
30_Operación									■	
40_Medición						■				
50_Cuenta									■	■
51_Marco conceptual									■	■
61_Clasificación Funcional									■	■
70_Fuente									■	■
Base de datos documental	000_bd_documental		■							
	500_Cuenta									■
	501_Partida									■
Base de datos del RETC	000_bd_RETC_2008		■							
	100_Agente		■							
	111_Directorio		■							
	112_CIIU_Rev3		■							
	115_Rubro		■							
	121_Fte_emis_RETC		■							
	162_Region		■							
	221_Sustancia									■
	222_TipoDescarga									■
	Matriz de Insumo-Producto-RETC	00_bd_MIP		■						
00_Emision			■							
021_Codigo_MIP_111			■							
023_Codigo_MIP_10			■							
221_Sustancia										■

Base HEA: Base de datos de Hechos Económicos y Ambientales