

El formato HDF: Un modelo de datos para el almacenamiento y gestión de información espacial de carácter ambiental.

Palomo, M.*
Ormeño, S.* ; Rincón, J. A.**

email: marcos.palomo@upm.es

* Universidad Politécnica de Madrid; ** CONACYT, Colegio e Postgraduados, México

1. Introducción:

En el ámbito científico, se genera gran cantidad de información en formato digital, que es necesario distribuir. Dada la diferente naturaleza de ésta, es necesario utilizar distintos formatos especialmente adaptados a cada uno de los tipos de datos que se vayan a almacenar. El principal problema, radica en la estandarización y la interoperabilidad, ya que, pese a que existen numerosos estándares de almacenamiento de información, no todos son soportados por las distintas aplicaciones informáticas utilizadas, lo que puede presentar problemas en el momento de la difusión y reutilización de dicha información.

El formato HDF (Hierarchical Data Format), representa una alternativa eficaz al ser adoptado como estructura de almacenamiento de datos, cuyas características más representativas son las siguientes:

- Permite obtener información acerca de los datos de un archivo desde dentro de ese archivo, sin necesidad de recurrir a fuentes externas.
- Permite almacenar datos de distinta naturaleza en un mismo archivo y relacionarlos entre ellos.
- Estandariza los formatos y las descripciones de los tipos de datos más comúnmente empleados.
- Se trata de un formato abierto, con sus especificaciones publicadas, lo que permite su implementación en diversas aplicaciones informáticas, facilitando la portabilidad, así como permitiendo al usuario desarrollar sus propias aplicaciones específicas.
- Es flexible y puede ser adaptado para almacenar cualquier tipo de dato.

2. El formato HDF:

Historia:

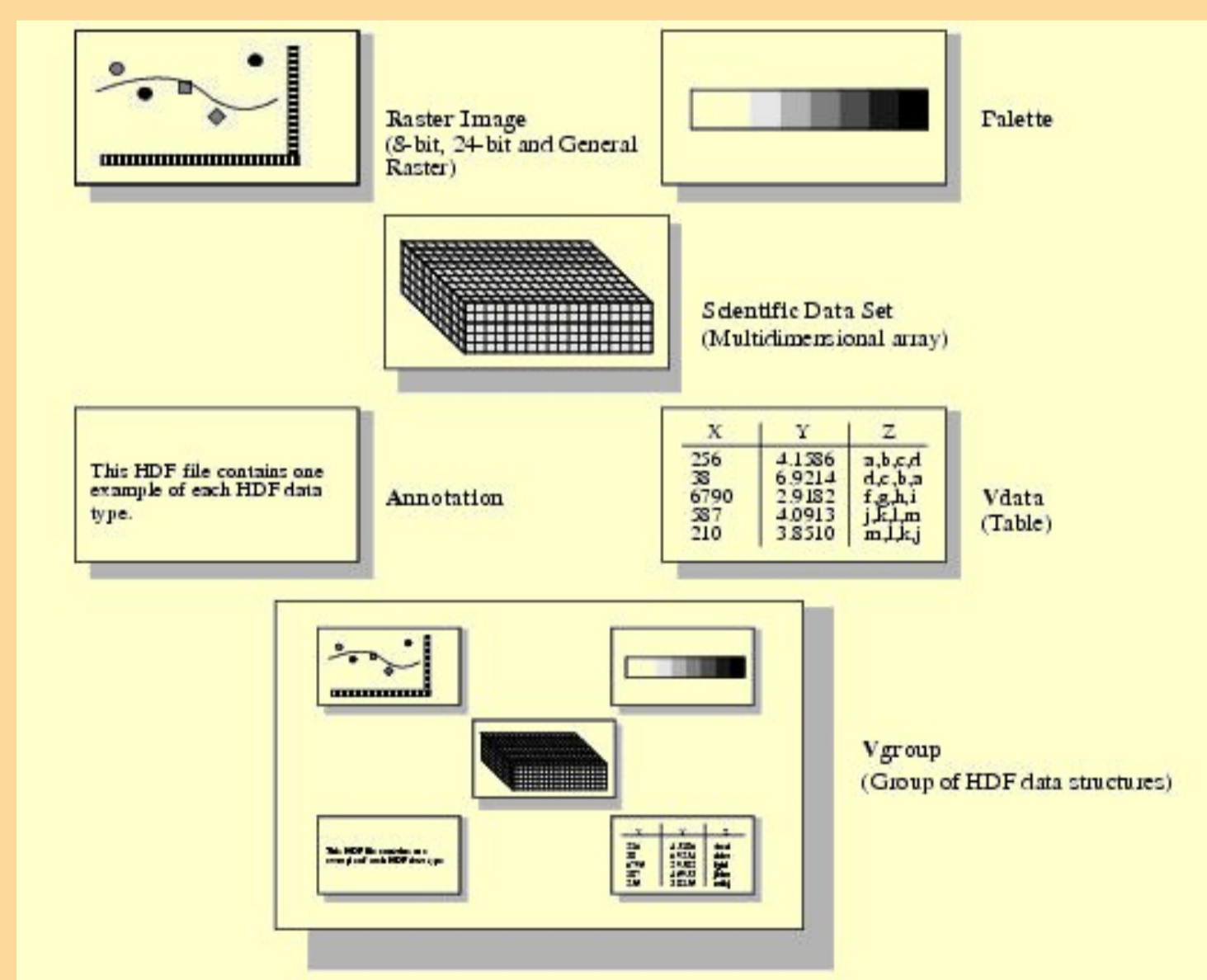
Con el fin de dar respuesta a las necesidades de almacenamiento y difusión de datos científicos de diversa naturaleza, en un formato flexible e independiente de la plataforma, comenzó a desarrollarse el formato HDF en los laboratorios de la NCSA (National Center for Supercomputing Applications) a partir del año 1988, siendo soportado en la actualidad por el HDF Group, dependiente de la Universidad de Illinois.

Actualmente, este formato se utiliza por múltiples organismos, tanto públicos como privados, para la difusión de sus datos y resultados, entre las que se encuentran la NASA, la Agencia Europea del Espacio (ESA), que los aplican en datos procedentes de los sensores MODIS, MERIS o ETM+, entre otros.

Estructura:

La estructura de los archivos HDF permite el almacenamiento de diversos tipos de datos, tales como:

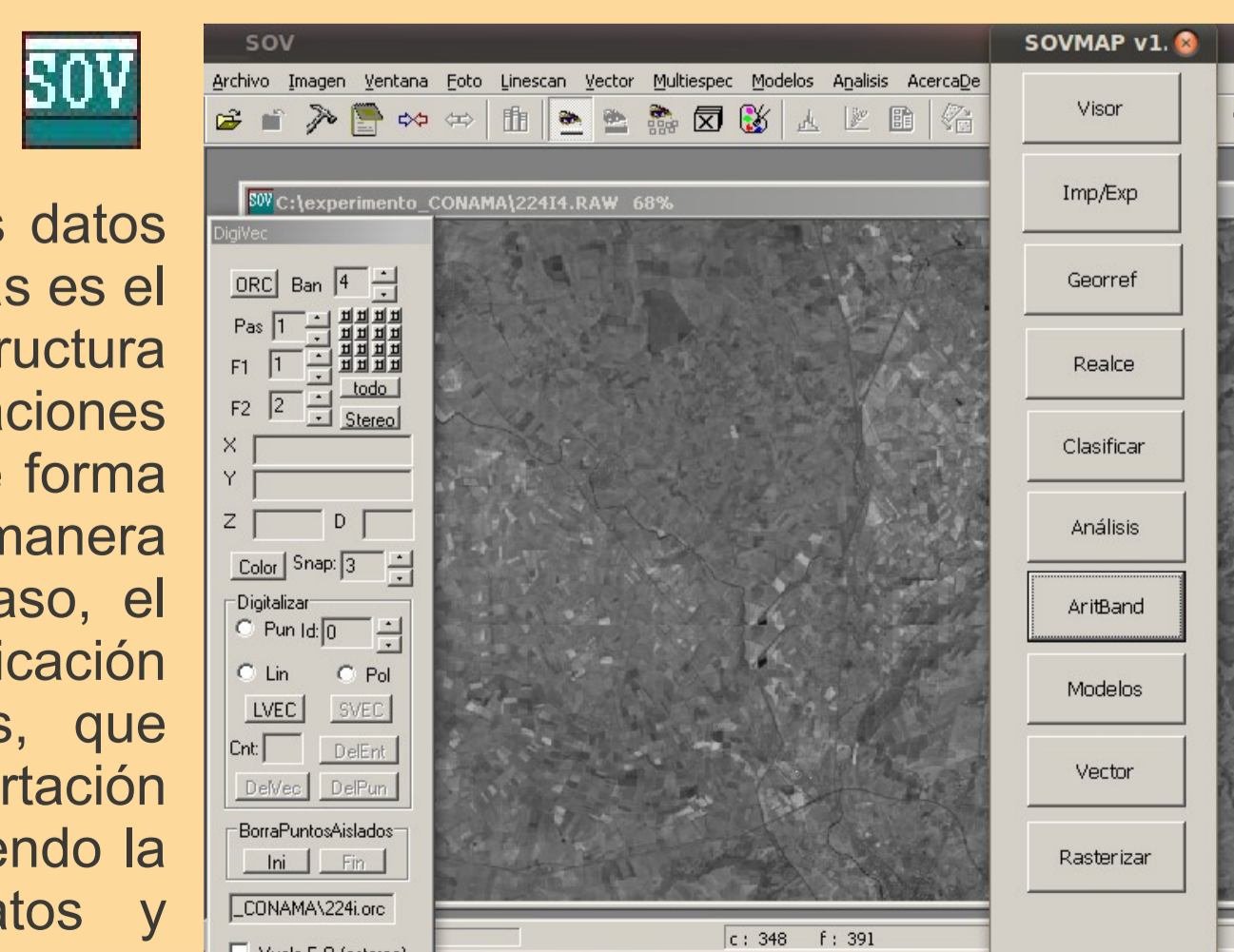
- **Scientific Data Sets (SD)**, utilizados para almacenar matrices n-dimensionales de datos enteros (8, 16 y 32 bits, con o sin signo) o reales (32 o 64 bits) en su formato estándar, o por medio de las APIs, crear datos en otros rangos de valores (1 a 32 bits), junto con sus metadatos (dimensiones de la matriz, atributos, etc.)
- **Raster Images (RI)**, que permiten almacenar imágenes de 8 bits (0-255 niveles de gris) o 24 bits (RGB), o bien utilizar nuevamente las librerías de programación para guardar imágenes en otros formatos (16 a 32 bits enteros, 32 a 64 bits coma flotante). Además, se permite almacenar información sobre las dimensiones de la imagen, así como la paleta de color asociada a ella. También es posible utilizar compresión (RLE, JPEG, GZIP y adaptativa Huffman) para reducir el tamaño de los archivos resultantes.
- **Text Annotations (TA)**, para almacenar cualquier tipo de información textual: Etiquetas, descripciones o información de archivo.
- **VData (VD)**, para almacenar datos vectoriales sin topología.
- **Vgroups (VG)**, que nos permiten asociar datos relacionados dentro de un archivo.



Estructura de un archivo HDF (fuente: The HDF Group)

Acceso a los datos:

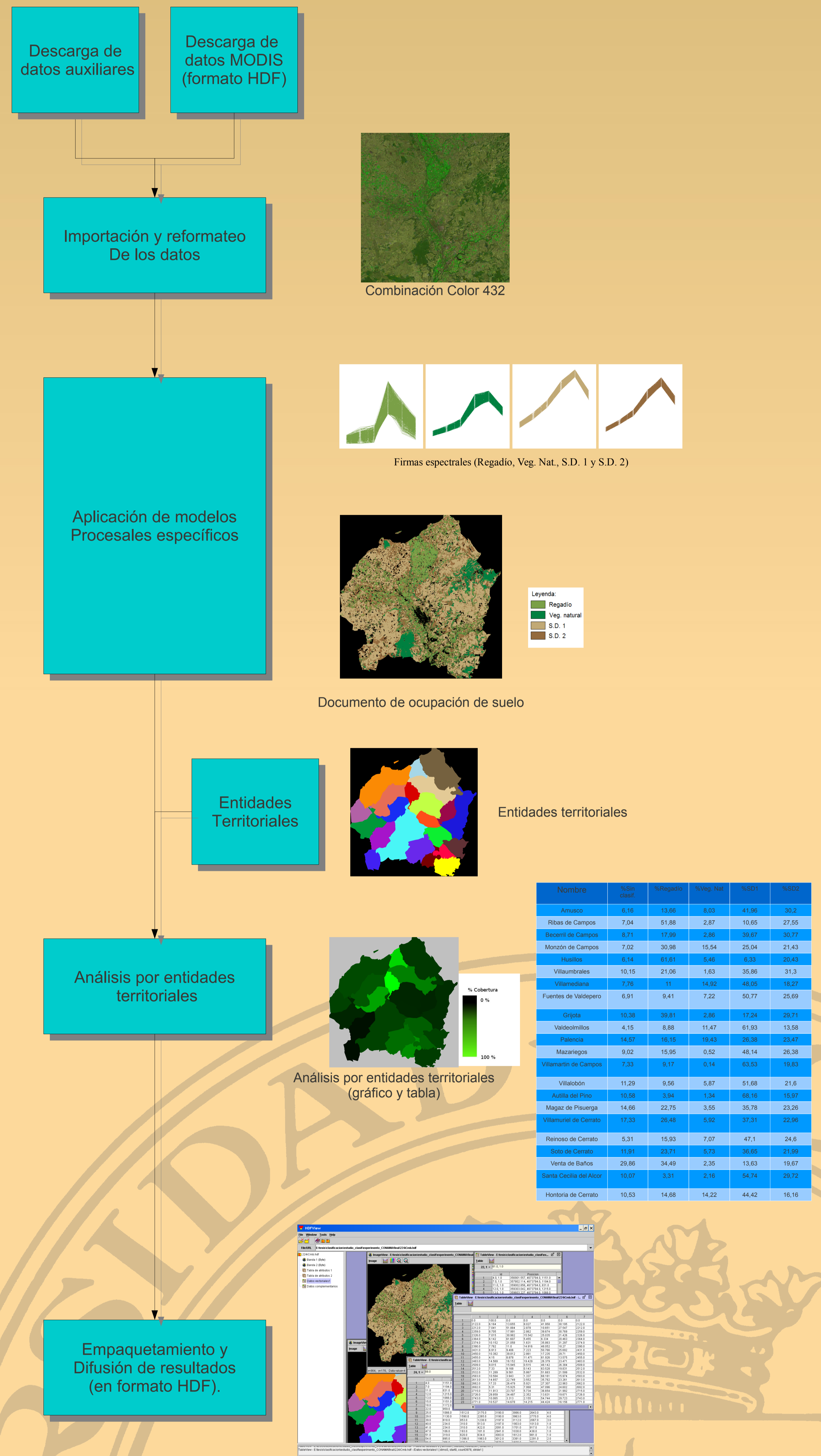
Existen diversas formas de acceder a los datos contenidos en un archivo HDF. Una de ellas es el acceso a nivel binario, conociendo la estructura del archivo, o bien a través de aplicaciones disponibles que soporten este formato, de forma que el proceso se realice de manera transparente para el usuario. En este caso, el acceso se realiza a través de la aplicación SOVMAP, desarrollada por los autores, que incluye un módulo de importación/exportación con soporte para el formato HDF, permitiendo la extracción y empaquetamiento de datos y productos generados en este formato.



3. Obtención de parámetros ambientales a partir de MODIS:

Ocupación del suelo, por entidades territoriales.

* Descripción del proceso:



4. Conclusiones:

En los trabajos referidos en la presente comunicación se ha demostrado la validez del modelo de datos de los archivos HDF para el almacenamiento y distribución de imágenes y otros datos de interés ambiental, por su versatilidad y posibilidades de adaptación, en este caso, referidos a un documento cualitativo de ocupación del suelo así como diversos documentos temáticos, tablas y datos vectoriales.

El formato HDF se convierte en una estructura fundamental para los fines referidos, al tratarse de un formato nativo adoptado por diversas agencias, así como organismos distribuidores y productores de información, que es utilizado por un número cada vez mayor de aplicaciones software.

Referencias bibliográficas:

- Ormeño, S., Arozarena, A., Martínez, M., Palomo, M., Villa, G., Peces, J. J., Pérez, L. (2008): "Los satélites de media y baja resolución espacial como fuente de datos para la obtención de indicadores ambientales". IX Congreso Nacional de Medio Ambiente. Madrid.
- Ormeño, S (2006): "Teledetección Fundamental". ETSIGC. Departamento de Ingeniería Cartográfica y Topografía. Universidad Politécnica de Madrid.
- Palomo, M (2006): "El formato HDF para el almacenamiento de información relativa a imágenes de satélite. Importación y exportación en SOV de ficheros HDF". Trabajo del Curso de doctorado "Cartografía ambiental". Universidad Politécnica de Madrid.
- The HDF Group (2010): "HDF4 User's Guide. HDF4 Release 2.5" [en línea] <ftp://ftp.hdfgroup.org/HDF/Documentation/HDF4.2.5/HDF425_UserGd.pdf>
- The HDF Group (2010): "HDF4 Reference Manual. HDF4 Release 2.5" [en línea] <ftp://ftp.hdfgroup.org/HDF/Documentation/HDF4.2.5/HDF425_RefMan.pdf>
- The HDF Group (2008): "HDF Specification and Developer's Guide. HDF4 Release 2.3" [en línea] <ftp://ftp.hdfgroup.org/HDF/Documentation/HDF423_SpecDG.pdf>