

COBIApp

Una aplicación para convertir datos de monitoreo a diferentes formatos

Manual de usuario

Villaseñor-Derbez, J.C.

Contents

Propósito	2
Formatos de bases de datos	2
Formato A	2
Formato B	3
Formato C	4
Uso de la aplicación	5
Interfaz de usuario	5
Panel de control	6
Opciones de Entrada	7
Opciones de Salida	9
Ejemplos de Formatos	9
Panel de resultados	9
Barra numérica	10
Descarga de datos	10
Posibles problemas	10

Propósito

Diferentes oficinas y personal de COBI tienen diferentes preferencias sobre el formato a utilizar para realizar análisis de datos rápidos. Debido a que todas las bases de datos ahora se encuentran unificadas, convertir y modificar las bases podría tomar bastante tiempo. Por lo tanto, ésta aplicación permite convertir bases de datos entre tres formatos distintos, descritos a continuación.

Formatos de bases de datos

Por consenso, todas las bases de datos tienen el mismo formato en los campos descriptivos: Día, Mes, Año, Estado, Comunidad, Sitio, Latitud, Longitud, Habitat, Zonificacion, TipoDeProteccion, ANP, BuzoMonitor, HoraInicialBuceo, HoraFinalBuceo, ProfundidadInicial_m, ProfundidadFinal_m, Temperatura_C, Visibilidad_m, Corriente, Transecto, Genero, Especie, GeneroEspecie. Note la ausencia de caracteres especiales como acentos, espacios y paréntesis. La base de entrada puede contener ese tipo de caracteres, pero las columnas serán renombradas para evitar el mal funcionamiento de la aplicación.

Las diferencias entre las bases son más bien en la forma de reportar las tallas y abundancias de especies de peces contadas en los **transectos de peces** de cada región. A continuación se describen los tipos de formatos utilizados, junto con algunas de sus ventajas y desventajas. Las imágenes de las tablas mostradas representan **la misma información**, representada en cada uno de los formatos.

Es importante que los formatos de bases de datos sean **idénticos** a los aquí especificados. De lo contrario, la aplicación podrá mostrar algunos errores. El usuario puede descargar bases de muestra directamente desde la aplicación.

Formato A

Éste formato era utilizado por la oficina del SAM para la captura de datos de **transectos de peces**. Se caracteriza por lo siguiente:

- Las tallas se presentan en intervalos: 0-5, 6-10, 11-20, 21-30, 31-40 y >40
- Cada rango de tallas representaba una columna
- Los datos anotados en cada celda correspondían a las abundancias de la especie X en cada uno de los rangos de talla
- Hay una columna de “Talla” para indicar la talla exacta de los organismos >40 cm
- Hay una columna de “Total” para sumar las abundancias totales de la especie X

Este formato permite reducir el número total de filas y, por lo tanto, el peso del archivo. Permite una rápida generación de histogramas de talla, abundancia y biomasa en Excel sin necesidad de recurrir a tablas dinámicas. Sin embargo, el formato de datos no permite realizar análisis más complejos en softwares estadísticos como R, MiniTab, Stata, SPSS o Primer, que requieren que los datos estén organizados por *factores*. Además, este formato conlleva una pérdida de información, ya que los rangos de tallas utilizados reducen la precisión con la que se puede calcular otros parámetros, como biomasa o potencial reproductivo. La conversión a formatos B o C no puede hacerse directamente por medio de tablas dinámicas de excel, ya que estas “agregan” información, no la “separan”.

GeneroEspecie	Sexo	0a5	6a10	11a20	21a30	31a40	>40	Talla	Total
Pomacanthus paru	y		2	2					4
Chromis multilineata	y	10							10
Epinephelus striatus	y				3	2			5
Epinephelus morio	y						2	100	2
Sphyraena argentea	y					3	1	45	4

Figure 1: Ejemplo de una base de datos con el formato A.

Note que los nombres de columnas son 0a5 6a10 11a20 21a30 31a40 >40. Es importante que se use de esta manera para evitar que Excel reconozca 6-10 como 6 de Octubre, o 11-20 como Noviembre 20. Al utilizar separar los valores con “a” en lugar de un guión “-”, evitamos que el sistema tome decisiones por nosotros.

Formato B

El formato B es utilizado por las oficinas de PBC y RGI. Se caracteriza por lo siguiente:

- Las tallas se ordenan en una sola columna, permitiendo utilizar esta columna como *factor* para análisis en softwares estadísticos.
- Cada intervalo de talla (columna *PromedioDeTalla*) representa un registro (fila) con su respectiva abundancia, indicada en la columna *Abundancia*.

Este formato reduce elimina 6 columnas en total, pero extiende el número de filas. Para bases con mucha información, pueden extenderse bastante. En excel se pueden usar tablas dinámicas para convertir a formato A, pero no a formato C.

Genero	Especie	GeneroEspecie	Sexo	Talla	PromedioDeTalla	Abundancia
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	10
Pomacanthus	paru	Pomacanthus paru	y	6a10	8.5	2
Pomacanthus	paru	Pomacanthus paru	y	11a20	15.5	2
Epinephelus	striatus	Epinephelus striatus	y	21a30	25.5	3
Epinephelus	striatus	Epinephelus striatus	y	31a40	35.5	2
Sphyraena	argentea	Sphyraena argentea	y	31a40	35.5	3
Sphyraena	argentea	Sphyraena argentea	y	>40	45	1
Epinephelus	morio	Epinephelus morio	y	>40	100	2

Figure 2: Ejemplo de una base de datos con el formato B.

Note que en la columna de TALLA los intervalos de talla se especifican como 0a5 6a10 11a20 21a30 31a40 >40. Es importante que se use de esta manera para evitar que Excel reconozca 6-10 como 6 de Octubre, o 11-20 como Noviembre 20. Al utilizar separar los valores con “a” en lugar de un guión “-”, evitamos que el sistema tome decisiones por nosotros. El programa buscará específicamente estos rangos de tallas, especificados en las instrucciones de la base “*Peces_Transecto.xlsx*”.

Formato C

El formato C ha sido utilizado, en ocasiones, por las oficinas de PBC y RGI. Se caracteriza por lo siguiente:

- Tiene los mismos nombres de columnas que el formato B (PromedioDeTalla y Abundancia), pero
- Cada registro (fila) representa un organismo, no un rango de talla

Este formato es significativamente más extenso, y puede que bases con mucha información no puedan ser abiertas por Excel, que tiene baja capacidad computacional. Sin embargo, es la mejor opción para realizar análisis estadísticos complejos. Permite manipular el formato más fácilmente, pues los *factores* de talla pueden “agregarse” para convertirse en columnas (como en el formato A) o simplemente sumar por rango de tallas (Formato B).

Genero	Especie	GeneroEspecie	Sexo	Talla	PromedioDeTalla	Abundancia
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	1
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	1
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	1
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	1
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	1
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	1
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	1
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	1
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	1
Chromis	multilineata	Chromis multilineata	y	0a5	2.5	1
Pomacanthus	paru	Pomacanthus paru	y	6a10	8.5	1
Pomacanthus	paru	Pomacanthus paru	y	6a10	8.5	1
Pomacanthus	paru	Pomacanthus paru	y	11a20	15.5	1
Pomacanthus	paru	Pomacanthus paru	y	11a20	15.5	1
Epinephelus	striatus	Epinephelus striatus	y	21a30	25.5	1
Epinephelus	striatus	Epinephelus striatus	y	21a30	25.5	1
Epinephelus	striatus	Epinephelus striatus	y	21a30	25.5	1
Epinephelus	striatus	Epinephelus striatus	y	31a40	35.5	1
Epinephelus	striatus	Epinephelus striatus	y	31a40	35.5	1
Sphyraena	argentea	Sphyraena argentea	y	31a40	35.5	1
Sphyraena	argentea	Sphyraena argentea	y	31a40	35.5	1
Sphyraena	argentea	Sphyraena argentea	y	31a40	35.5	1
Sphyraena	argentea	Sphyraena argentea	y	>40	45	1
Epinephelus	morio	Epinephelus morio	y	>40	100	1
Epinephelus	morio	Epinephelus morio	y	>40	100	1

Figure 3: Ejemplo de una base de datos con el formato C.

Note que en la columna de **TALLA** los intervalos de talla se especifican como **0a5 6a10 11a20 21a30 31a40 >40**. Es importante que se use de esta manera para evitar que Excel reconozca 6-10 como 6 de Octubre, o 11-20 como Noviembre 20. Al utilizar separar los valores con “a” en lugar de un guión “-”, evitamos que el sistema tome decisiones por nosotros. El programa buscará específicamente estos rangos de tallas, especificados en las instrucciones de la base “*Peces_Transecto.xlsx*”.

Uso de la aplicación

La aplicación se desarrolló con el paquete **shiny** de **R**. Así obtenemos una aplicación que “vive” en el internet. Esto es especialmente útil ya que no hay necesidad de instalar software específico en las computadoras. Además, la aplicación puede utilizarse desde cualquier dispositivo con conexión a Internet, incluyendo tablets y celulares. Para acceder a ésta aplicación, el link es <https://jcvd.shinyapps.io/COBIApp/>.

Después de 15 minutos de inactividad, la página se vuelve inactiva para ahorrar espacio computacional. Para reiniciar la aplicación, uno puede usar la tecla **F5**, refrescar la página, o dar click en el botón de *Reload*. La página se volverá a cargar y la aplicación estará activa de nuevo.

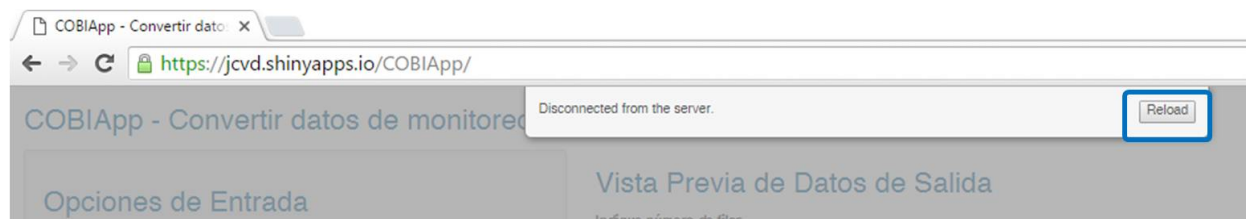


Figure 4: Mensaje de la página indicando que han transcurrido más de 15 minutos y que la aplicación se ha desconectado del servidor. De click en el botón encerrado en azul para volver a activar la aplicación.

Interfaz de usuario

La interfaz de usuario es el medio con que el usuario puede comunicarse con el código detrás de la aplicación. En este caso, provee los espacios necesarios para manipular la entrada y salida de los datos. En este caso, la interfaz se divide en dos secciones principales, el *Panel de control* y el *Panel de resultados*.

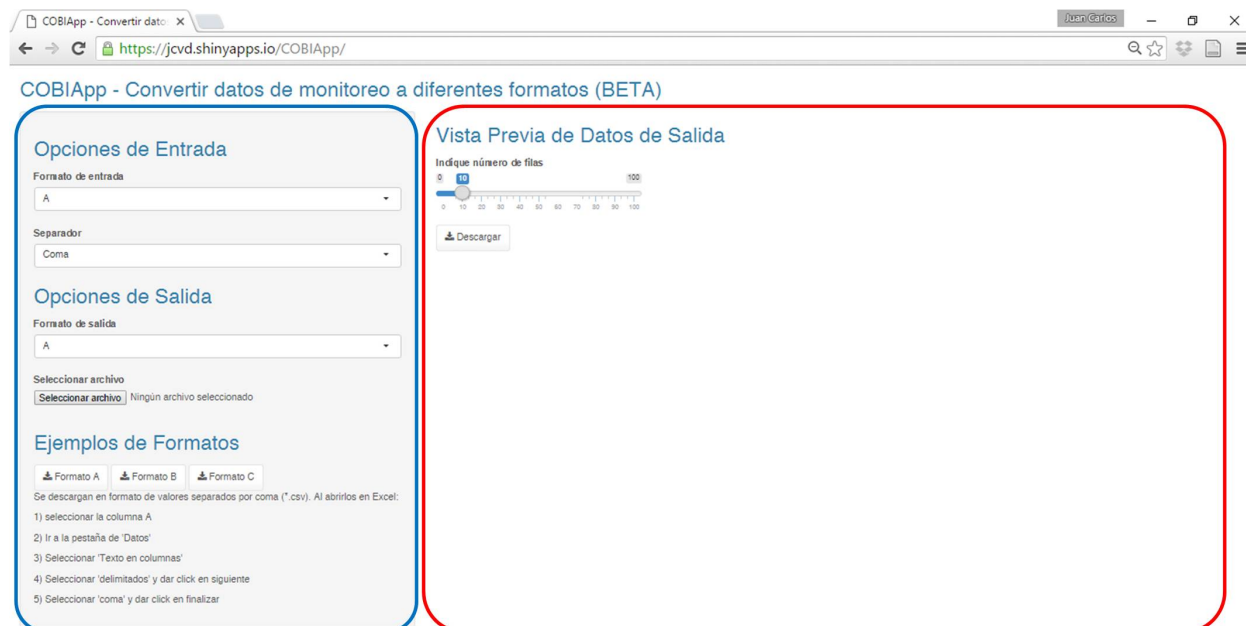


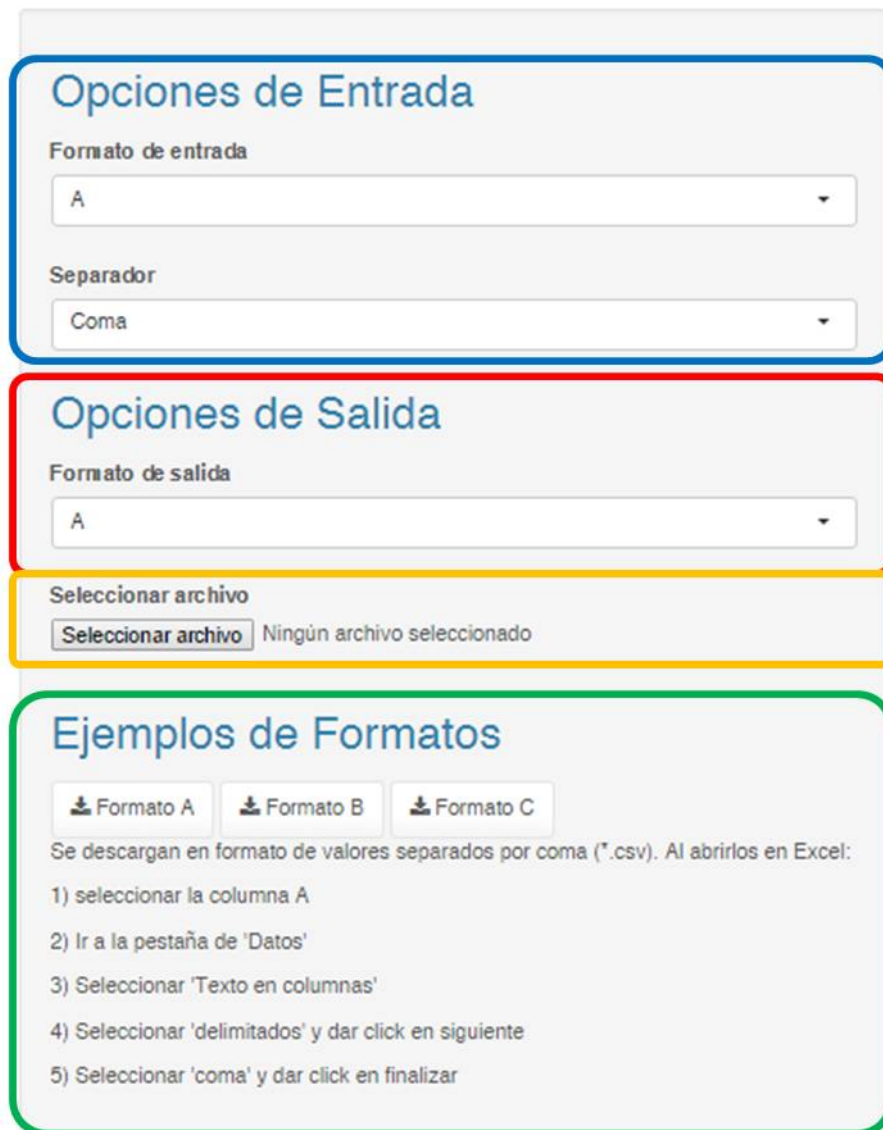
Figure 5: El rectángulo azul indica la sección de *Panel de control*, el rectángulo rojo indica la sección de *Panel de resultados*

La interfaz de usuario está diseñada con la siguiente lógica en mente:

1. El usuario indica el formato de datos que va a cargar (Opciones de Entrada)
2. El usuario indica el formato de datos al que quiere convertir sus datos de entrada (Opciones de Salida)
3. El usuario provee al programa la base de datos
4. El usuario observa una vista previa de los datos convertidos
5. El usuario descarga los datos convertidos

Adicionalmente, la sección de *Ejemplos de Formatos* provee al usuario con la opción de descargar bases de datos con cada uno de los formatos.

Panel de control



The image shows a web interface for data conversion. It is divided into four main sections, each highlighted with a colored border:

- Opciones de Entrada (Blue border):** Contains two dropdown menus. The first is labeled 'Formato de entrada' and has 'A' selected. The second is labeled 'Separador' and has 'Coma' selected.
- Opciones de Salida (Red border):** Contains one dropdown menu labeled 'Formato de salida' with 'A' selected.
- Cargar datos (Orange border):** Contains a button labeled 'Seleccionar archivo' and the text 'Ningún archivo seleccionado'.
- Ejemplos de Formatos (Green border):** Contains three buttons labeled 'Formato A', 'Formato B', and 'Formato C'. Below these buttons is a paragraph of text: 'Se descargan en formato de valores separados por coma (*.csv). Al abrirlos en Excel:' followed by a numbered list of five steps: 1) seleccionar la columna A, 2) Ir a la pestaña de 'Datos', 3) Seleccionar 'Texto en columnas', 4) Seleccionar 'delimitados' y dar click en siguiente, 5) Seleccionar 'coma' y dar click en finalizar.

Figure 6: El *Panel de control* se divide en tres secciones: *Opciones de Entrada* (azul), *Opciones de Salida* (rojo), *Cargar datos* (naranja) y *Ejemplos de Formatos* (verde).

Opciones de Entrada

Uno de los formatos más utilizados para el manejo de bases de datos es el de valores separados por coma o *.csv. Este formato puede ser abierto en Excel, y la mayoría de los softwares para análisis de datos. Entre sus ventajas destacan un tamaño de archivo reducido, elimina funciones en celdas, y estandariza el formato de la información contenida.

Debido a que ésta aplicación “vive” en el Internet, es necesario usar archivos de bajo peso que permitan que los procesos computacionales se realicen rápidamente. Utilizar formatos *.csv en lugar de *.xls permite que la información se cargue entre 8 y 10 veces más rápido. Para guardar una base de Excel como *.csv dar click en “Guardar como” y seleccionar la opción que indica “valores separados por coma (CSV)”.

En esta sección permite al usuario especificar el formato de entrada y los separadores de valores del archivo que carga. En el campo de *Formato de Entrada*, el usuario especifica si los datos están en formato A, B o C (ver sección *Formatos de bases de datos*).

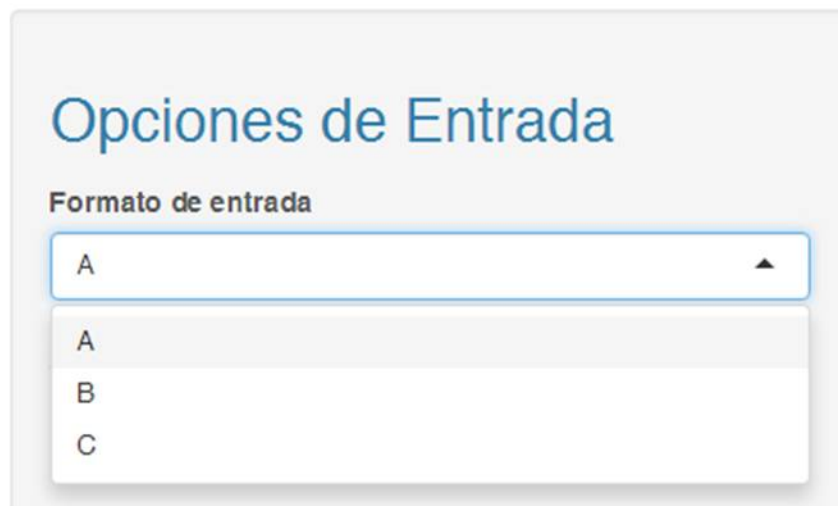


Figure 7: Opciones para formato de entrada

El usuario también especifica el separador utilizado en su archivo *.csv en el campo de *Separador*. Diferentes computadoras pueden tener defaults distintos para usar comas, espacios, tabulaciones, o punto y coma. Esto brinda flexibilidad al usuario. Las computadoras con lenguaje predeterminado a Español (Latinoamérica) suelen utilizar el “;” como separador. Las MACs suelen utilizar la “,”.

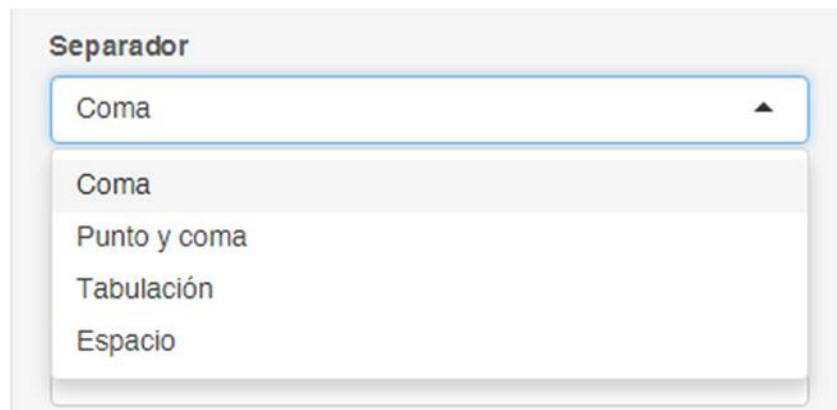


Figure 8: Opciones para el tipo de separador de valores del archivo de entrada

Si el formato incorrecto es especificado, se mostrará un mensaje de error o la conversión se verá extraña.

En el ejemplo a continuación, observe las diferencias en los campos de la sección de *Opciones de Entrada* del *Panel de Control*.

The screenshot shows the COBIApp interface with the following settings and results:

- Opciones de Entrada:**
 - Formato de entrada: A
 - Separador: Coma
- Opciones de Salida:**
 - Formato de salida: A
- Vista Previa de Datos de Salida:**
 - Indique número de filas: 100
 - Preview table with 1 column (all data is concatenated into a single cell):

Dia	Mes	Ano	Estado	Comunidad	Sitio	Latitud	Longitud	Habitat	Zonificacion	TipoDeProteccion	ANP	BuzoMonitor	HorainiciaBuceo	HorafinalBuceo	Profundidad	Inicio	Fin
1	a	b	c	d	1	1	g	h	coral	j	k	l	m	n	o	p	q
2	a	b	c	d	1	1	g	h	coral	j	k	l	m	n	o	p	q
3	a	b	c	d	1	2	g	h	coral	j	k	l	m	n	o	p	q
4	a	b	c	d	1	2	g	h	alga	j	k	l	m	n	o	p	q
5	a	b	c	d	2	1	g	h	coral	j	k	l	m	n	o	p	q

Figure 9: Los datos cargados aquí están separados por “;”, pero el usuario indicó que estaban separados por “,”. Por lo tanto, el resultado es una tabla de una sola columna (el programa no reconoce las separaciones entre columnas).

The screenshot shows the COBIApp interface with the following settings and results:

- Opciones de Entrada:**
 - Formato de entrada: A
 - Separador: Punto y coma
- Opciones de Salida:**
 - Formato de salida: A
- Vista Previa de Datos de Salida:**
 - Indique número de filas: 100
 - Preview table with 18 columns (data is correctly separated):

Dia	Mes	Ano	Estado	Comunidad	Sitio	Latitud	Longitud	Habitat	Zonificacion	TipoDeProteccion	ANP	BuzoMonitor	HorainiciaBuceo	HorafinalBuceo	Profundidad	Inicio	Fin
1	a	b	c	d	1	1	g	h	coral	j	k	l	m	n	o	p	q
2	a	b	c	d	1	1	g	h	coral	j	k	l	m	n	o	p	q
3	a	b	c	d	1	2	g	h	coral	j	k	l	m	n	o	p	q
4	a	b	c	d	1	2	g	h	alga	j	k	l	m	n	o	p	q
5	a	b	c	d	2	1	g	h	coral	j	k	l	m	n	o	p	q

Figure 10: Cuando el usuario indica el separador correcto, el resultado se presenta de forma correcta. Las columnas siguen hacia la derecha.

Opciones de Salida

En el campo de *Formato de Salida*, el usuario especifica el formato A, B o C (ver sección *Formatos de bases de datos*) al cual quiere convertir los datos.



Figure 11: Opciones para el *Formato de Salida* indicado por el usuario.

Ejemplos de Formatos

Los formatos de ejemplo se descargan en formato de valores separados por coma (*.csv). Al abrirlos en Excel:

1. Seleccionar la columna *A*
2. Ir a la pestaña de *Datos*
3. Seleccionar *Texto en columnas*
4. Seleccionar *delimitados* y dar click en siguiente
5. Seleccionar *coma* y dar click en finalizar

Panel de resultados

El *Panel de resultados* se compone de una barra numérica, la tabla de datos a descargar (unicamente visible al producir la conversión) y el botón de descarga de datos.



Figure 12: Panel de resultados antes de producir un archivo para descargar

Vista Previa de Datos de Salida

Indique número de filas



	Día	Mes	Año	Estado	Comunidad	Sitio	Latitud	Longitud	Habitat	Zonificacion	TipoDeProteccion	ANP	BuzoMonitor	HorainicialBuceo	HorafinalBuceo	Profundidad
1	a	b	c	d	1	1	g	h	coral	j	k	l	m	n	o	p
2	a	b	c	d	1	1	g	h	coral	j	k	l	m	n	o	p
3	a	b	c	d	1	2	g	h	coral	j	k	l	m	n	o	p
4	a	b	c	d	1	2	g	h	alga	j	k	l	m	n	o	p
5	a	b	c	d	2	1	g	h	coral	j	k	l	m	n	o	p

Descargar

Figure 13: Panel de resultados después de producir un archivo para descargar. Las columnas siguen hacia la derecha.

Barra numérica

La barra numérica permite al usuario indicar el número de filas a previsualizar. El valor de default es de 10, pero el usuario puede ver hasta las primeras 100 filas antes de descargar el archivo.

Descarga de datos

El usuario da click en el botón de *Descargar*, que aparece al final de la tabla de previsualización. Se abrirá una ventana donde el usuario podrá indicar el directorio donde quiere descargar los datos. Los datos se decargan con formato *.csv delimitado por comas. Según el sistema operativo y la configuración de lenguaje de cada usuario, podrá ser necesario seguir los siguientes pasos para visualizar los datos correctamente. Al abrirlos en Excel:

1. Seleccionar la columna *A*
2. Ir a la pestaña de *Datos*
3. Seleccionar *Texto en columnas*
4. Seleccionar *delimitados* y dar click en siguiente
5. Seleccionar *coma* y dar click en finalizar

Posibles problemas

Por ahora no se han encontrado problemas específicos. De ocurrir algo, comunicarse con Juan Carlos Villaseñor a juancarlos.villader@gmail.com. Si se provee la siguiente información, será más sencillo rastrear el error y corregirlo:

- Captura de pantalla de la configuración utilizada y los mensajes de error (de existir)
- Bases de datos cargadas y descargadas al/del programa
- Cualquier otra información relevante como conversiones anteriores realizadas