ANEXO I. FORMATO DE LAS TABLAS

Los datos representativos del modelo de enfermedad de transmisión están estructurados en tablas con formato **CSV**. Dichas tablas pueden ser generadas con el editor interno o usando cualquier herramienta externa siempre que cumpla con un formato y etiquetado compatible. En este apartado se describen tanto las etiquetas como las reglas del formato que han de aplicarse para un correcto funcionamiento e los módulos de la aplicación (formato de datos interno) y/o importación de datos (formato de datos externos).

Las etiquetas pueden ser <u>simples o compuestas</u>. Las simples contienen unicamente un operador mientras que las compuestas están conformadas por un operador seguido por uno o dos nombres de grupos de población dependiendo del significado del operador separados por espacios.

A continuación están descritos los formatos generales de datos tanto internos como externos y finalmente se enumeran las etiquetas con su descripción particular.

ESTRUCTUA Y FORMATO DE DATOS INTERNOS

Este apartado contiene el formato propio de los módulos internos de la aplicación cuyos datos están agrupados por tablas {HST,PAL,DEF,PRJ,MAP,DEF}, por lo tanto a excepción parcial de las tablas MAP, HST y PRJ, deberá observar tres preceptos generales:

- 1. No hay un orden del etiquetado estricto pero debe respetarse el mostrado en las etiquetas obligatorias cuando no se indique otra cosa.
- 2. Todas las etiquetas descritas en este apartado son obligatorias, no así los datos asociados que pueden estar presentes o no.

- La primera fila de cada tabla coloreada en marrón está reservada a la descripción de cada columna y no forman parte del etiquetado de los datos. También debe respetarse dicha nomenclatura.
- 4. La primera columna no es intercambiable en posición con otras columnas.
- 5. En las etiquetas, columnas y datos asociados donde una parte de su etiquetado referencia a grupos de población también referenciados en otras tablas se debe mantener la coherencia de dichas referencias.
- 6. Los datos numéricos decimales no deben estar en formato científico, dicho formato provoca en algunos casos perdidas de precisión o incluso lecturas incorrectas.
- 7. Los datos numéricos decimales usan el punto como separador "." es decir, formato europeo.

Tabla PRJ

La Figura 30 muestra el formato del módulo que define un modelo de enfermedad de transmisión o proyecto de la aplicación. La columna "*Tipo*" contiene las etiquetas obligatorias {*NAME,AUTHOR,DESCRIPTION,DATE0,DATE1,VERSION,NG*} y las etiquetas opcionales {*map,def,rel,hst,pal*} mientras que la columna "*Dato*" contiene los valores asociados a dichas etiquetas.

Las restantes etiquetas hacen referencia a los ficheros que contienen la información referente a los otros módulos del modelo. Serán obligatorias cuando se introduzca una referencia de un archivo en la columna "*Dato*".

Data
Dato
Modelo B.csv
Modelo de enfermedad de transmisión.
1.0
4
Modelo B.map
Modelo B.def
Modelo B.rel
Modelo B.hst
Modelo B.pal

Figura 30: Tabla PRJ

Tabla DEF

La Figura 31 muestra el formato de la tabla para la definición de los atributos de una enfermedad de trasmisión. Todas las etiquetas {PTE,DME,DMI,IP,IT,FT} de la columna "Tipo" son obligatorias y sus datos deben estar agregados antes de cualquier simulación o generación de evolución (histórico) mientras que los valores asociados para dichas etiquetas están en la columna "Dato".



Figura 31: Tabla DEF

Tabla PAL

La Figura 32 representa la escala de colores por defecto de la aplicación con los niveles de contagio por cada 100000 habitantes.

La Figura 33 muestra los 10 niveles de contagio representados en esta aplicación por la columna "Level" mediante las etiquetas {L0,L1,L2,L3,L4,L5,L6,L7,L8,L9}. Estas etiquetas son obligatorias y debe respetarse el orden. Las columnas {Red,Green,Blue} almacenan los valores asociados a cada etiqueta de la columna "Level" en formato numérico entero, cuyos valores desde 0 hasta 255, representan el grado de color en formato **RGB** para cada nivel de contagio.

Editando modulo: Paleta 🕒 📗						
Level	Red	Red Green				
LO	82	190	128			
L1	40	180	99			
L2	174	214	241			
L3	46	134	193			
L4	247	220	111			
L5	243	156	18			
L6	210	180	222			
L7	118	68	138			
L8	230	126	34			
L9	231	76	60			

Figura 33: Tabla PAL

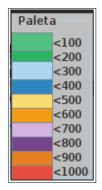


Figura 32: Escala de colores por defecto

Tabla REL

La Figura 34 muestra el formato de la matriz de contactos o módulo de relaciones de grupos de población. La nomenclatura mostrada a modo de ejemplo es particular para los datos del modelo cargado a modo de ejemplo y difiere de cualquier otro que pueda generarse si bien, muestra la estructura donde la primera columna contiene la etiqueta "*Grupos*" la cual puede variar a placer. Los nombres de las restantes columnas debe coincidir con el nombre de las etiquetas de la primera columna así como su orden. Dichos nombres o tanto de las columnas como de las filas de la primera columna se corresponderán con el nombre asignado a cada grupo de población definido. Los valores contenidos en las celdas de la matriz deben ser numéricos y valores enteros.

Los valores asignados no deberían quedar vacíos, y tendrá efecto diferente un valor de "0" que el almacenado en la celda simétrica a la diagonal. Esto cobra especial importancia cuando se procede a realizar los cálculos de un modelo mediante el uso de un generador de datos basados en el modelo dinámico de transmisión de enfermedad SIR, no así en la importación de datos donde dichos cálculos ya han sido realizados y no se pretende generarlos ni modificarlos.

En el ejemplo mostrado los nombres de los grupos de población son los mostrados en las columnas $\{G1,G2,G3,G4\}$.

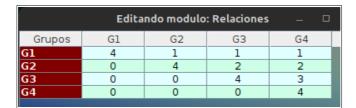


Figura 34: Tabla REL

Tabla MAP

La Figura 35 muestra la estructura y el etiquetado de la tabla correspondiente al módulo de grupos de población, en esta versión de la aplicación también a los datos representativos gráficos de los grupos de población. El orden de las columnas es estricto y no debe variarse aunque si es posible en esta versión usar otra nomenclatura para las cabeceras. La primera columna muestra los identificadores de cada grupo de población, su orden, si bien no es estricto es conveniente su inserción en orden. La definición de las

etiquetas {*ID,NAME,PEOPLE,AREA,S,I,R,P,C100K*} son obligatorias y las referentes a la definición de los puntos "*Px;yn*" dependerá de la necesidad del usuario.

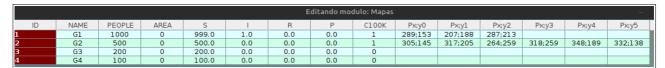


Figura 35: Tabla MAP

Tabla HST

La tabla del módulo histórico es el resultado de la generación de una simulación en función de las variables iniciales que modelan la enfermedad de transmisión en el contexto de los grupos de población y sus relaciones. Es por ello que las etiquetas pueden variar en función de la fuente generadora de los datos y de si son generados por la propia aplicación o por otras herramientas externas.

En todo caso hay ciertas etiquetas que van a aparecer y una estructura general que en este caso y a diferencia de las anteriores tablas tienen un orden en las etiquetas laxo con algunas restricciones al tipo de datos y estructura:

- 1. La cabecera es exclusiva para la definición de tiempos siendo:
 - 1.1. La columna uno debe tener una etiqueta referida a la palabra tiempo.
 - 1.2. Los nombres de las restantes columnas deben tener el formato de grupo fecha hora siendo este "dd/MM/yyyy hh:mm", otro formato provocará lectura incorrecta en el momento de reproducción de los datos.
- 2. Las filas de la primera columna están reservadas al etiquetado de los diferentes operadores y grupos de población.
- No hay restricción sobre el orden de las etiquetas. El sistema lee y asigna los datos asociados a las etiquetas según operador, grupo de población y la marca de tiempo en la cabecera de cada columna.
- 4. Una etiqueta compuesta debe contener los valores asociados para todo el intervalo de tiempo definido en la cabecera. De otro modo pueden producirse errores durante la reproducción del histórico.

Si bien no hay otras restricciones, este módulo tiene como finalidad la reproducción de la evolución de una enfermedad de transmisión a través de grupos de población. Por ello hay ciertas etiquetas básicas que deberían estar presentes para una correcta integración con el resto de la aplicación a pesar de no ser obligatorias. Estas etiquetas son {*S,I,R,C100K*}.

La Figura 36 muestra un ejemplo de las posibles etiquetas que pueden aparecer al importar un ejemplo de datos generados con una herramienta externa para cuatro grupos de población con una fuente de datos con estructura de **ModeloA**.

La Figura 36 también muestra dos errores posibles. El primero marcado en violeta refleja un uso incorrecto del formato numérico decimal con el uso de notación científica y el segundo en el par de etiquetas complejas "*CAB*" finales que no tienen definidos los valores para toda la secuencia del intervalo de tiempos.

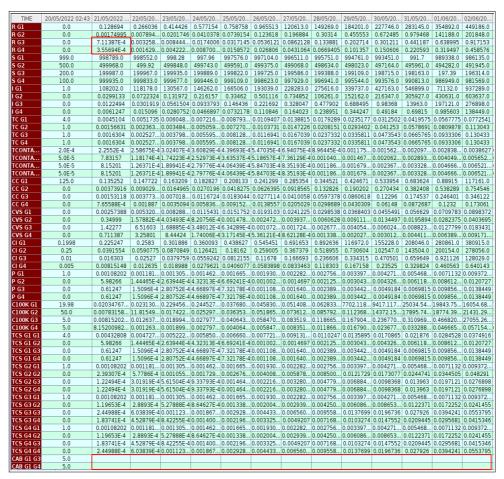


Figura 36: Ejemplo de histórico al importar un ModeloA de 4 grupos de población.

La Figura 37 muestra un ejemplo similar donde se ven variaciones en el número y algunos tipos de etiqueta al importar un ejemplo de datos generados también con una herramienta externa para cuatro grupos de población pero con una estructura de datos de **ModeloB**.

TIME	20/05/2022 02:40	21/05/2022 02:40	22/05/2022 02:40	23/05/20	24/05/20	25/05/20	26/05/20
R G1	0.0	0.128694	0.266036	0.414426	0.577154	0.758758	0.965513
R G2	0.0	0.00174995	0.00789409	0.0201746	0.0410378	0.0739154	0.123618
R G3	0.0	7.11387E-4	0.00325801	0.008444	0.0174006	0.0317145	0.0536121
R G4	0.0	3.55694E-4	0.00162901	0.004222	0.008700	0.0158572	0.026806
S G1	999.0	998.789	998.552	998.28	997.96	997.576	997.104
S G2	500.0	499.968	499.92	499.848	499.743	499.591	499.375
S G3	200.0	199.987	199.967	199.935	199.889	199.822	199.725
S G4	100.0	99.9935	99.9833	99.9677	99.9446	99.9109	99.8623
l G1	1.0	1.08202	1.18178	1.30567	1.46262	1.66506	1.93039
I G2	0.0	0.0299133	0.0722324	0.131972	0.216157	0.33462	0.501116
I G3	0.0	0.0122494	0.0301919	0.0561504	0.0933793	0.146436	0.221692
I G4	0.0	0.0061247	0.015096	0.0280752	0.0466897	0.0732178	0.110846
CC G1	0.125	0.135252	0.147722	0.163209	0.182827	0.208133	0.241299
CC G2	0.0	0.00373916	0.00902906	0.0164965	0.0270196	0.0418275	0.0626395
CC G3	0.0	0.00153118	0.00377399	0.007018	0.0116724	0.0183044	0.0277114
CC G4	0.0	7.65588E-4	0.001887	0.0035094	0.005836	0.009152	0.0138557
CVS G1	0.0	0.00257388	0.00532072	0.008288	0.0115431	0.0151752	0.0193103
CVS G2	0.0	3.4999E-5	1.57882E-4	4.03493E-4	8.20756E-4	0.001478	0.002472
CVS G3	0.0	1.42277E-5	6.51603E-5	1.68885E-4	3.48012E-4	6.34289E-4	0.001072
CVS G4	0.0	7.11387E-6	3.25801E-5	8.44424E-5	1.74006E-4	3.17145E-4	5.36121E-4
CI G1	0.1998	0.225247	0.2583	0.301886	0.360093	0.438627	0.545451
CI G2	0.025	0.0391554	0.0590775	0.0870849	0.126421	0.18162	0.259005
CI G3	0.01	0.016303	0.02527	0.0379759	0.0559242	0.0812155	0.11678
CI G4	0.005	0.00815148	0.012635	0.018988	0.0279621	0.0406077	0.0583898
P G1	0.001	0.001082020309	0.00118178021	0.001305	0.001462	0.001665	0.001930
P G2	0.0	5.982664029324	1.44464763453	2.639439	4.323138	6.692406	0.001002
P G3	0.0	6.124701200839	1.50959160409	2.807525	4.668970	7.321794	0.001108
P G4	0.0	6.124701200808	1.50959962244	2.807520	4.668970	7.321781	0.001108
C100K G1	19.98	22.52470644206	25.8300047527	30.18859	36.00930	43.86270	54.54510
C100K G2	5.0	7.831085274235	11.8154970109	17.41697	25.28419	36.32403	51.80102
C100K G3	5.0	8.151501598224	12.6349715769	18.98798	27.96213	40.60771	58.38991
C100K G4	5.0	8.151481598179	12.6349968399	18.98800	27.96210	40.60771	58.38982

Figura 37: Ejemplo de histórico al importar un ModeloB de 4 grupos de población.

Algunos ejemplos de etiquetas mostrados en las tablas:

- R G1, indicaría que es un dato del grupo de población G1 cuyo valor hace referencia al número de personas recuperadas.
- C100K G2, indicaría el nivel de contagio para el grupo de población G2.
- **TCS** G1 G2, indica la tasa de contactos del grupo de población G1 con sintomáticos del grupo de población G2.

ESTRUCTURA Y FORMATO DE DATOS EXTERNOS

La aplicación es capaz de importar datos desde aplicaciones externas usando etiquetados para leer los diferentes atributos que intervienen para cada módulo que define el modelo de enfermedad de transmisión con dos posibles disposiciones de etiquetas llamados **Modelos**, por columnas (*vertical*) al que se le llama "*Modelo A*" como muestra la Figura 38, como por filas (*horizontal*) llamado "*Modelo B*" como muestra la Figura 39.

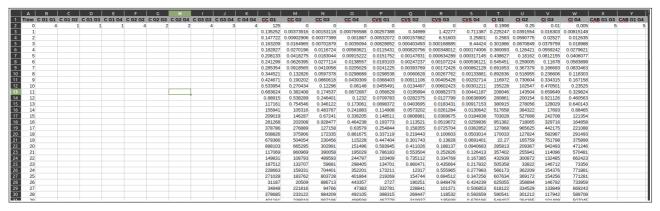


Figura 38: Ejemplo de estructura de datos Modelo A.

	Α	В	С	D	E	F	G	н	1	J	K	L	٨
1	Time	. 0	1	2	3	4	5	6	7	8	9	10	
2	ÇÇ G1	0,125	0,135252	0,147722	0,163209	0,182827	0,208133	0,241299	0,285354	0,344521	0,424671	0,533954	0,6
3	CC G2	0	0,00373916	0,00902906	0,0164965	0,0270196	0,0418275	0,0626395	0,0918565	0,132826	0,190202	0,270434	0,3
	ÇÇ G3	0	0,00153118	0,00377399	0,00701879	0,0116724	0,0183044	0,0277114	0,0410058	0,0597378	0,0860618	0,12296	
	ÇÇ G4	0	0,000765588	0,001887	0,0035094	0,00583621		0,0138557	0,0205029	0,0298689	0,0430309	0,06148	
	CVS G1	0	0,00257388	0,00532072	0,00828852	0,0115431	0,0151752	0,0193103	0,0241225	0,0298538	0,0368403	0,0455491	0,0
	CXS G2	0	0,000034999	0,000157882		0,000820756		0,00247237	0,00393769		0,00911106		
	CXS G3	0	0,0000142277			0,000348012		0,00107224	0,00172426		0,00405428		
	CVS G4		0,00000711387	0,0000325801	0,0000844424			0,000536121				0,00301211	
	ÇĮ G1	0,1998	0,225247	0,2583	0,301886	0,360093	0,438627	0,545451	0,691653	0,892636	1,16972	1,55228	2,
	ÇI G2	0,025	0,0391554	0,0590775	0,0870849	0,126421	0,18162	0,259005	0,367379	0,518955	0,730604	1,02547	1,
	ÇI G3	0,01	0,016303	0,02527	0,0379759	0,0559242	0,0812155	0,11678	0,166693	0,236606	0,334315		0,6
	ÇI G4	0,005	0,00815148	0,012635	0,018988	0,0279621	0,0406077	0,0583898	0,0833463	0,118303	0,167158	0,23525	
14	I G1	1	1,08202	1,18178	1,30567	1,46262		1,93039	2,28283	2,75616	3,39737	4,27163	
15	I G2	0	0,0299133	0,0722324	0,131972	0,216157	0,33462	0,501116	0,734852	1,06261	1,52162	2,16347	3,
16	I G3	0	0,0122494	0,0301919	0,0561504	0,0933793	0,146436	0,221692	0,328047	0,477902	0,688495		1
17	I G4	0	0,0061247	0,015096	0,0280752	0,0466897	0,0732178	0,110846	0,164023	0,238951	0,344247	0,49184	0,
	R G1	0	0,128694	0,266036	0,414426	0,577154	0,758758	0,965513	1,20613	1,49269	1,84201	2,27746	
	R G2	0	0,00174995	0,00789409	0,0201746	0,0410378	0,0739154	0,123618	0,196884	0,30314	0,455553	0,672485	
	R G3	0	0,000711387	0,00325801	0,00844424	0,0174006	0,0317145	0,0536121	0,0862128	0,133881	0,202714	0,301211	0,4
	R G4	0	0,000355694	0,00162901	0,00422212	0,00870031	0,0158572	0,026806	0,0431064	0,0669405	0,101357	0,150606	0,2
	S G1	999	998,789	998,552	998,28	997,96		997,104	996,511	995,751	994,761	993,451	
	S G2	500	499,968	499,92	499,848	499,743	499,591	499,375	499,068	498,634	498,023	497,164	49
	S G3	200	199,987	199,967	199,935	199,889	199,822	199,725	199,586	199,388	199,109	198,715	
	S G4	100	99,9935	99,9833	99,9677	99,9446	99,9109	99,8623	99,7929	99,6941	99,5544	99,3576	99
	C G1 G1	4											
	C G1 G2	1											
	C G1 G3	1											
	C G1 G4	1											
	C G2 G2	4											
	C G2 G3	2											
	C G2 G4	2											
	C G3 G3	4											
	C G3 G4	3											
	C G4 G4	0.05											
	PIE	0,05											
	DME	50											
	DMI IP	0											
39	J.C.	U											
40													

Figura 39: Ejemplo de estructura de datos Modelo B.

Siendo su fin último la reproducción de una secuencia, no son necesarias y por tanto no son leídas todas las etiquetas asignadas en todas las tablas tales como fecha de modificación o nombre del proyecto entre otras, lo cual complicaría enormemente la importación de datos desde otras aplicaciones.

Ambos sistemas de adquisición de datos comparten unas reglas comunes dependientes del tipo de dato al que va asociada la etiqueta. Por ejemplo, hay etiquetas que definen un atributo de la enfermedad, por tanto van al módulo **DEF** y su valor es definido una única vez, en cambio otros datos varían a lo largo del tiempo y dependen de otros valores, por tanto es necesario asignar valores para todo el intervalo.

Regla de datos asociados a las etiquetas

Las etiquetas:

- {S0,I0,R0,PT0,C,PTE,DME,DMI,IP,IT,FT} definen sus valores una única vez.
- {S,I,R,CAB,CC,CVS,CI,C100K,P,TC,TCS,TCONTAGIO,TR,TVS} requieren tener valores asociados para todo el intervalo de tiempo importado.

Regla agrupación etiqueta "C"

Los elementos con la etiqueta "C" en común deben estar agrupados, pues de ella depende el cálculo correcto del número de grupos de población y la lectura de los nombres que se asignará a cada grupo.

Regla intervalo de tiempo

La etiqueta mostrada "time" no es leída y unicamente ofrece una referencia al lector para el análisis de los datos en forma de tabla, la columna o fila asociada a los datos del tiempo no es leída tampoco por tanto pueden contener cualquier valor que genere el usuario o la aplicación en su progreso de tiempo, dichos datos serán sustituidos por la fecha y hora en curso en el momento de importar los datos. Esta política de importación facilita al usuario el uso de herramientas externas, aunque es recomendable ajustarse a un formato grupo fecha-hora compatible con la aplicación para en posteriores versiones estos datos sean leídos.

ETIQUETAS USADAS EN LAS TABLAS

La relación de etiquetas usadas por la aplicación en sus diferentes tablas son detalladas en la Tabla 4 indicando el tipo de dato que usa, si es obligatorio, dependiente u opcional y el significado de la propia etiqueta. Aquellas etiquetas que sean de uso común estarán definidas a partir de la fila "Generales". Al final de la tabla está disponible una leyenda detallando significados.

Tabla 4: Etiquetas usadas en las tablas de cada módulo.

Etiqueta	•		Descripción
PRJ,	{NAME,AU7	THOR,DESC	RIPTION, DATE0, DATE1, VERSION, NG} Etiquetas simples.
NAME	Texto	Si	Indica el nombre del modelo, puede coincidir con el nombre del fichero en el dispositivo de almacenamiento.
AUTHOR	Texto	No	Específica el nombre del autor del modelo.
DESCRIPTION	Texto	No	Contiene una breve descripción de la enfermedad o característica del modelo.
DATE0	Fecha	No	Fecha de creación del modelo respetando la estructura "dd/MM/yyyy".
DATE1	Fecha	No	Fecha de última modificación del modelo respetando la estructura "dd/MM/yyyy".
VERSION	Texto	No	Versión del modelo que puede combinar tanto texto como números.
NG	Entero	Si	Número de grupos de población contenidos en el modelo.
			F,DME,DMI,IP,IT,FT} Etiquetas simples.
_	•		EOPLE,AREA,S,I,R,P,C100K} Etiquetas simples.
ID	Entero	Si	Identificador del grupo de población, no debe repetirse.
NAME	Texto	Si	Nombre del grupo de población, debe ser coherente con las referencias en otras tablas.
PEOPLE	Entero	Si	Población inicial del grupo de población.
AREA	Entero	Si	Superficie geográfica al que pertenece el grupo de población.
Px;yn	Texto	No	etiquetas opcionales que definen los puntos de coordenadas con valores enteros. Las coordenadas están separadas por el símbolo punto y coma ";" y "n" representa a la posición de ese punto dentro de la cadena de puntos que conforman la figura.
Generales.	Se describe	n como etiqu	letas simples aunque su uso general conlleva el uso de etiquetas compuestas (<i>Ej. S Madrid</i>).
S	Decimal	Si	Número de personas susceptibles.
1	Decimal	Si	Número de personas incidentes.
R	Decimal	Si	Número de personas recuperadas.
S0	Decimal	No	Número de susceptibles inicial.
10	Decimal	No	Número de incidentes inicial.
R0	Decimal	No	número de recuperados inicial.
P	Decimal	D	Indica la prevalencia.
PT0	Entero	No	Establece la población inicial para un grupo de población.

С	Texto	Si	Referente a una estructura matricial indica una relación entre dos grupos de población siendo el primer grupo de población el correspondiente a la fila en la matriz de contactos y el segundo a la columna.
тс	Decimal	No	Tasa de contactos.
TR	Decimal	No	Tasa de recuperados o curación.
TVS	Decimal	No	Tasa de vuelta a la susceptibilidad.
TCONTAGIO	Decimal	Si	Tasa de contagio.
CC	Decimal	No	Casos de curación.
cvs	Decimal	No	Casos de vuelta a la susceptibilidad.
CI	Decimal	D	Casos de incidentes.
CAB	Decimal	No	Número de contactos del grupo A con el grupo B.
TCS	Decimal	No	Tasa de contactos del grupo A con sintomáticos del grupo B.
C100K	Decimal	D	Número de casos por cada 100 mil habitantes.
PTE	Decimal	Si	Indica la probabilidad de trasmisión de la enfermedad
DME	Entero	Si	Indica la duración media d ella enfermedad en días.
DMI	Entero	D	Indica la duración media de la inmunidad en días.
IP	Booleano {0,1}	D	Indica si el modelo de enfermedad contempla la inmunidad permanente como atributo o no. O indicaría que no la contempla.
ΙΤ	Entero	No	Tiempo inicial de la simulación dentro del intervalo de tiempo simulado medido en días. Normalmente el valor inicial es 0 para referirse al primer día.
FT	Entero	No	Tiempo final de la simulación dentro del intervalo de tiempo simulado. Este dato suele coincidir con el total de tiempos en el intervalo simulado.

Leyenda apartado "Generales"

Columna obligatoria:

- **Si**, etiqueta requerida para el correcto funcionamiento.
- **No**, no es requerida para el correcto funcionamiento, pero si está presente optimiza cálculos o procesos. Cuando no está incluida sus valores no serán calculados.
- **D**, no es obligatoria cuando existen las etiquetas a partir de las cuales pueden ser calculadas. Mayor detalle en la sección Estructura y formato de datos externos.