

# Copying Datasets on the ABSDB for Testing SEASABS/TSUpdate

Frank Masci, 4/10/2006

## 1. Purpose

Below I outline a procedure for creating structurally independent datasets on the ABSDB (on either USER or the TEST UAC4 DB) for use in testing new TSUpdate features. Note that simply copying the series knowledge (metadata) using the 'HouseKeeping' utility is not enough. This is because a copy of the series knowledge for an existing group will still point to the same output products table on the ABSDB. An execution of TSUpdate using a copy of the series knowledge will update (and hence inadvertently corrupt) existing products intended for publication. This will make clients and your director very cranky.

A schematic of the interoperability between components on the ABSDB with SEASABS is in Figure 1. A description of the components is given in section 2. If one creates a copy of the series metadata for a group, i.e. "Series Metadata 2", together with a new Time Series link, "TSLink 2", the same "Output products" table is referenced and thus is potentially updated. This products table 'hangs' off the "DatasetTable" containing the 'raw' original data.

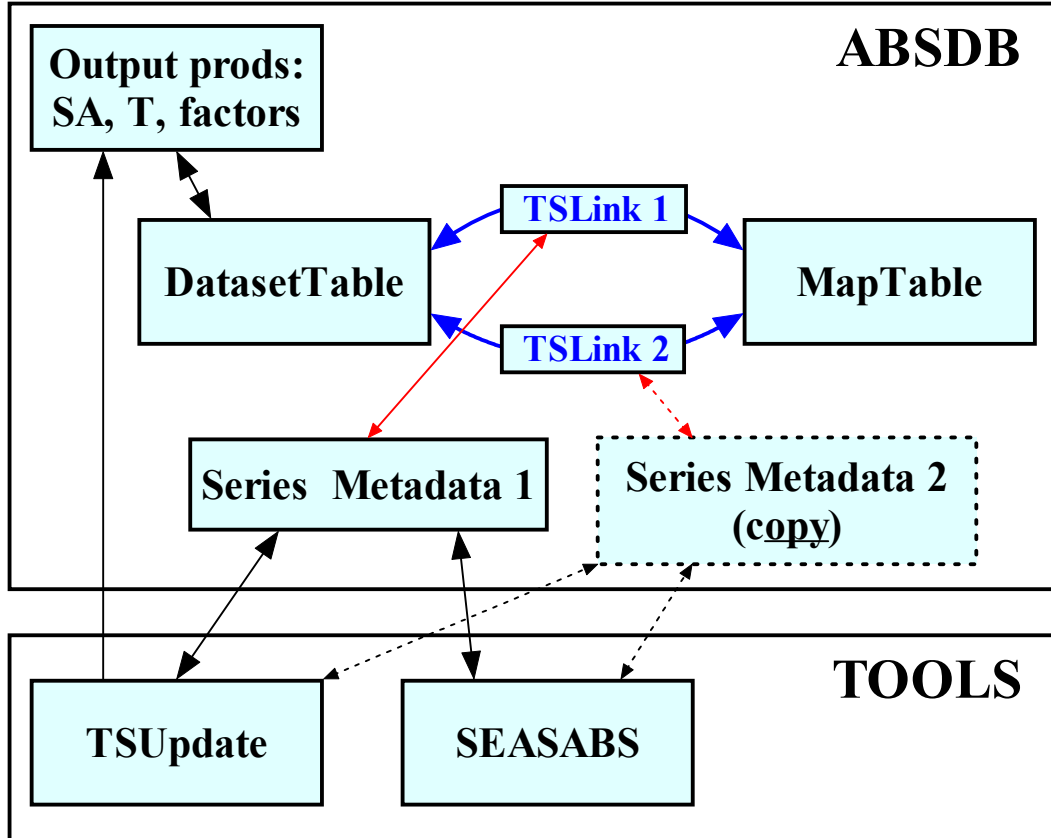


Figure 1: Dataset tables on ABSDB and some time series tools that access them.

## **2. Solution**

So, how do you avoid updating products intended for publication in the regular production cycle during testing? The simple answer is to create an independent copy of the original "DatasetTable". Overall, physically distinct copies of the following four data tables/attributes will need to be made on the appropriate database.

1. The entire DatasetTable containing 'raw' data tables, i.e. the annotation and macrodata tables.
  2. The MapTable listing actual SEASABS series names that map into original SMA-defined names.
  3. The time series link (TSLink) that links the DatasetTable to the respective MapTable. Incidentally, this is equal to the actual SEASABS group name stored in the series metadata.
  4. The series knowledge, or metadata table referenced by the TSLink. This contains processing parameters (e.g. priors) and settings updated/used by SEASABS/TUpdate to create seasonally adjusted series estimates.
- for collections stored on the ABSDB, copies of items 1, 2 and 3 can be created using tools on the ABS Information Warehouse (ABSIW).
  - for collections stored on a FAME database, equivalent tables/attributes to items 1 and 3 can be created using FAME scripts. There is no such equivalent to the MapTable (item 2) in FAME that needs copying.
  - for all series knowledge copies (i.e. item 4), use the SEASABS "housekeeping" facility. You will need the target "Owner Group" name and "TSLink" (=SEASABS group) name copy that you made on the ABSIW (item 3).

## **3. Which Database?**

There are two options:

1. If you intend to retain your test group(s) for future reference, then create your dataset copies directly on the main USER DB. The data management team copies the contents of USER to the user-acceptance test database UAC4 about once every 6 months. The existing contents of UAC4 get overwritten.
2. If you are not concerned about losing your test groups (within ~modulo 6 months), then create your dataset copies directly on the user-acceptance test database "UAC4". It is a good idea to check with data management on when the next USER to UAC4 replication will be. At the time of writing, the next replication will be on 30/10/2006 and then again ~6 months after this.

It is strongly recommended that all 'intensive' user-acceptance testing be carried out on the UAC4 test database. This is to reduce the load on the USER DB. If you go with option 1, the test group copies you created on USER are for storage only. That is, so they won't be lost when UAC4 gets overwritten.

If you plan to keep your test groups for eternity (or until deleted; i.e. option 1) and are planning on undergoing intensive/continuous testing, then you must perform the procedure below twice: first in the USER DB and then in UAC4.

#### **4. Owner Group and Naming Convention?**

It is recommended that all test groups be placed in the existing SEASABS owner group: "**ACC\_TEST\_SEAS**".

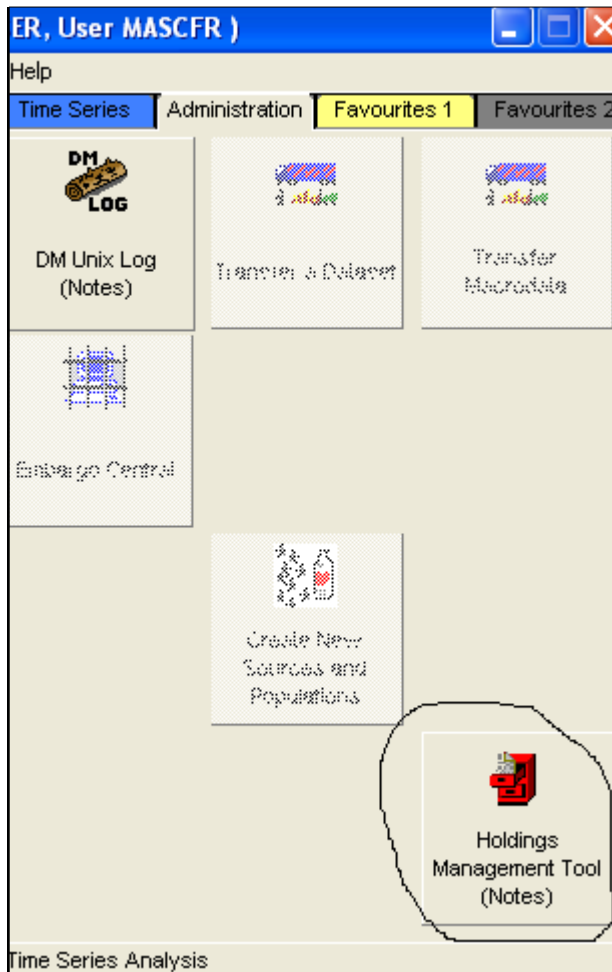
For consistency in the naming of test groups and to enable them to be easily identified, please use the following convention: ***original\_group\_name\_test1*** , ***original\_group\_name\_test2*** , etc.. where you simply append the suffix "***\_test#*** " to the original group name. This enables one to have multiple copies of the same test group.

#### **5. Procedure**

The procedure below assumes you are copying datasets and series knowledge entries directly within the USER database. To perform this in another database (e.g. UAC4), ensure you log into the desired database when using a specific tool in any of these steps.

**(1)**. Bring up the ABSIW utility.

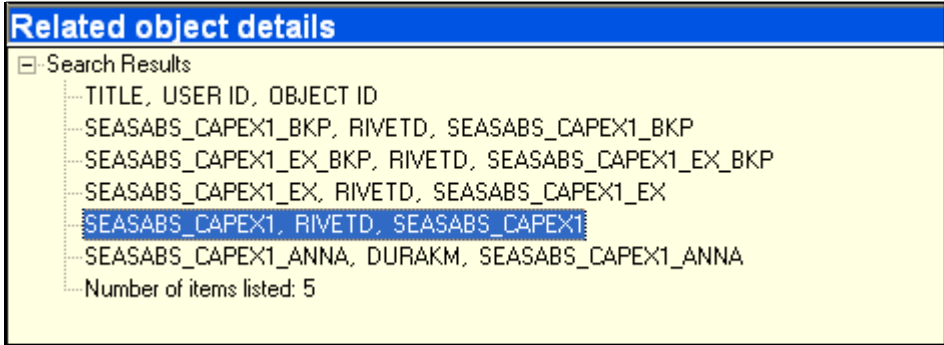
**(2)**. Use the "Holdings Management Tool" to identify your original dataset ID and MapTable name:



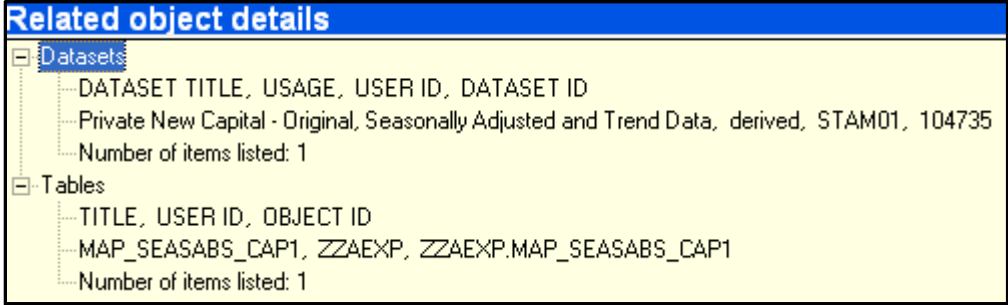
After logging in, click on "Search for: Time Series Links" and then enter the original TSLink name. This is equal to the original "group name" in SEASABS, e.g:

Keyword	User ID	Owner Group
Please enter the Id and appropriate keywords (or parts of them) that you wish to search by.		
Comments	<input type="text"/>	
Time Series Link	<input type="text" value="SEASABS_CAPEX1"/>	

Then you select and double click on the desired group name. Here it's shown highlighted:

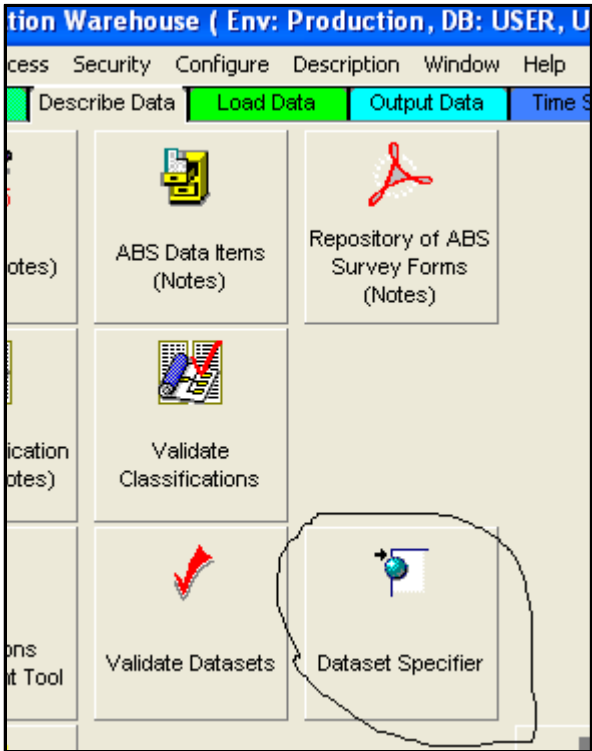


You will then be presented with the dataset identifiers and other table names:

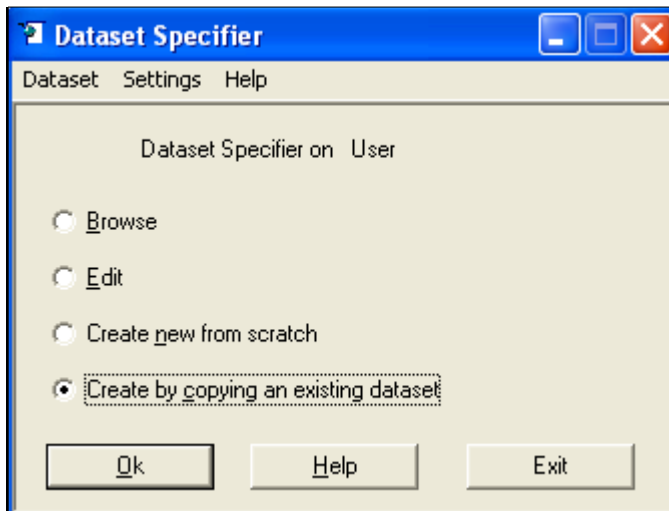


Here, the original dataset ID is 104735 and the map table name is: "MAP\_SEASABS\_CAP1". The user ID to which this belongs to is "ZZAEXP". We will need these three items below.

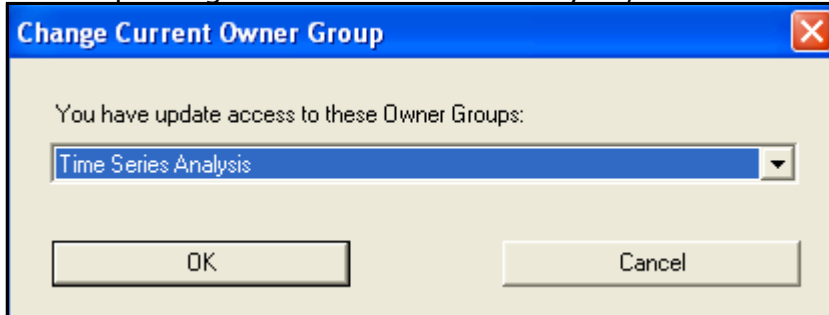
**(3).** Using the dataset ID above, create a copy of this dataset using the "Dataset Specifier" tool on the ABSIW:



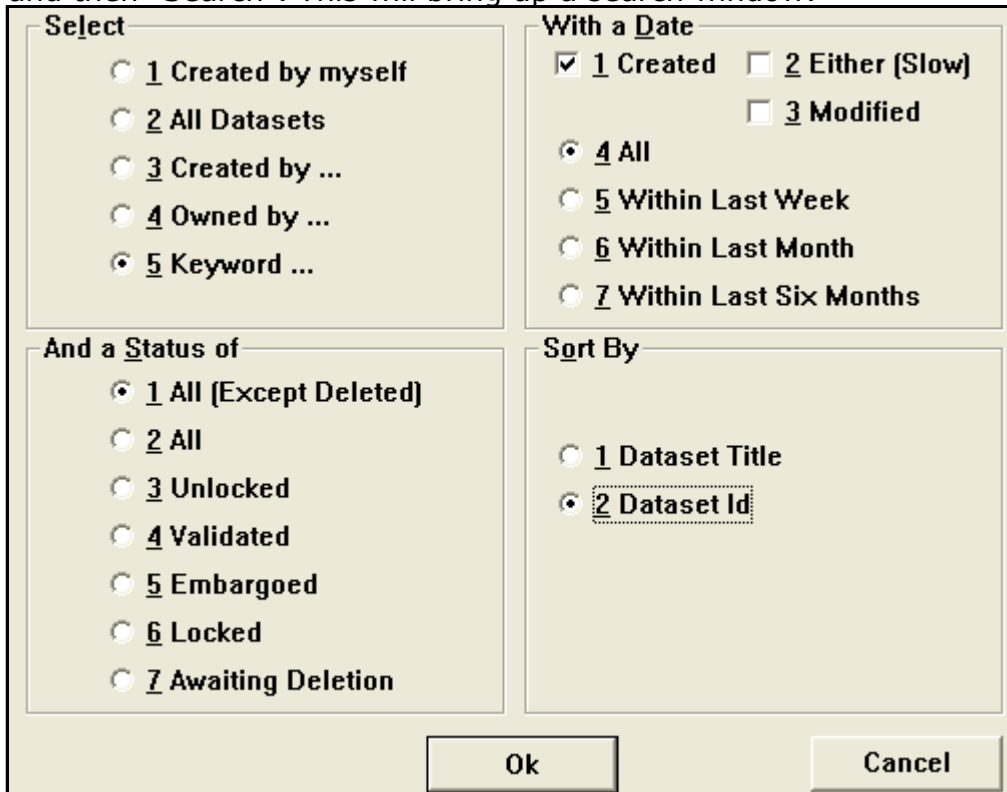
After double clicking, select the "Create by copying an existing dataset" button:



Ensure you have your default "Current Owner Group" set to one that you have access privileges. For time series analysts, this is usually "Time Series Analysis":



After clicking OK, the "Create Dataset Wizard" will pop up. Click on "Next>>" and then "Search". This will bring up a search window:



Select the "5 Keyword" and "2 Dataset Id" buttons and enter your original dataset ID from above, e.g:

Dataset Title Keyword :

Dataset Title Keyword 2 :

Dataset Id Keyword :

Return Datasets containing

ANY Keyword

ALL Keywords

After clicking OK, go back to the "Dataset Wizard", the dataset title will now appear, e.g:

Select the dataset to be copied

Search

**Select Items to Copy**

All

Parts

Title  Consultants Notes

Title Footnote  Population

Collection  Items

Publication Links  Item / Column Links


**Select Tables**

Select the following option if you want to copy tables of the dataset

None  Use existing tables  Copy existing tables

On the above window, ensure you select the "Copy existing tables" button. Then click "Next >". This will bring up the following dialogue window:

**Question**

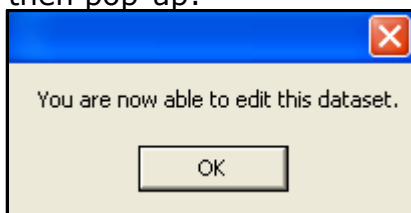
 If the new dataset is to be structurally independent of the existing dataset answer Yes. If you answer No, then any structural changes to new dataset will impact on the existing dataset and vice versa. See help for more details.

Click "Yes" here and then go back to the "Dataset Wizard" and click "Finish". The following window will then pop-up, summarising the new dataset tables created.

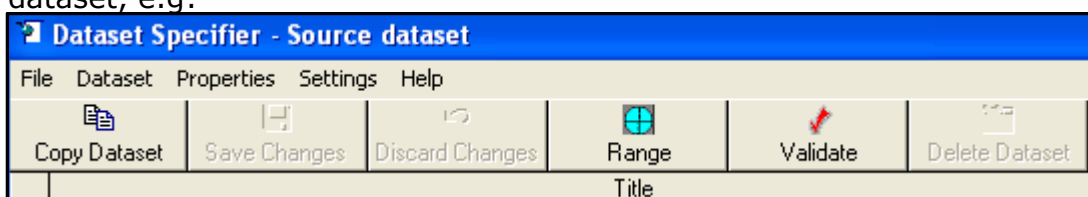
**Congratulations !**  
**You have created a dataset with the following details:**

Dataset Id	894843
Title	Private New Capital - Original, Seasonally Adjusted and Trend Data
Population	Business Management Units
Collection	New Capital Expenditure (Survey of)
Owner Group	Time Series Analysis
Created By	MASCFR
Date Specified	04-OCT-06
Status	Unlocked
Tables	MASCFR.AN_3046436 MASCFR.MD_3046435

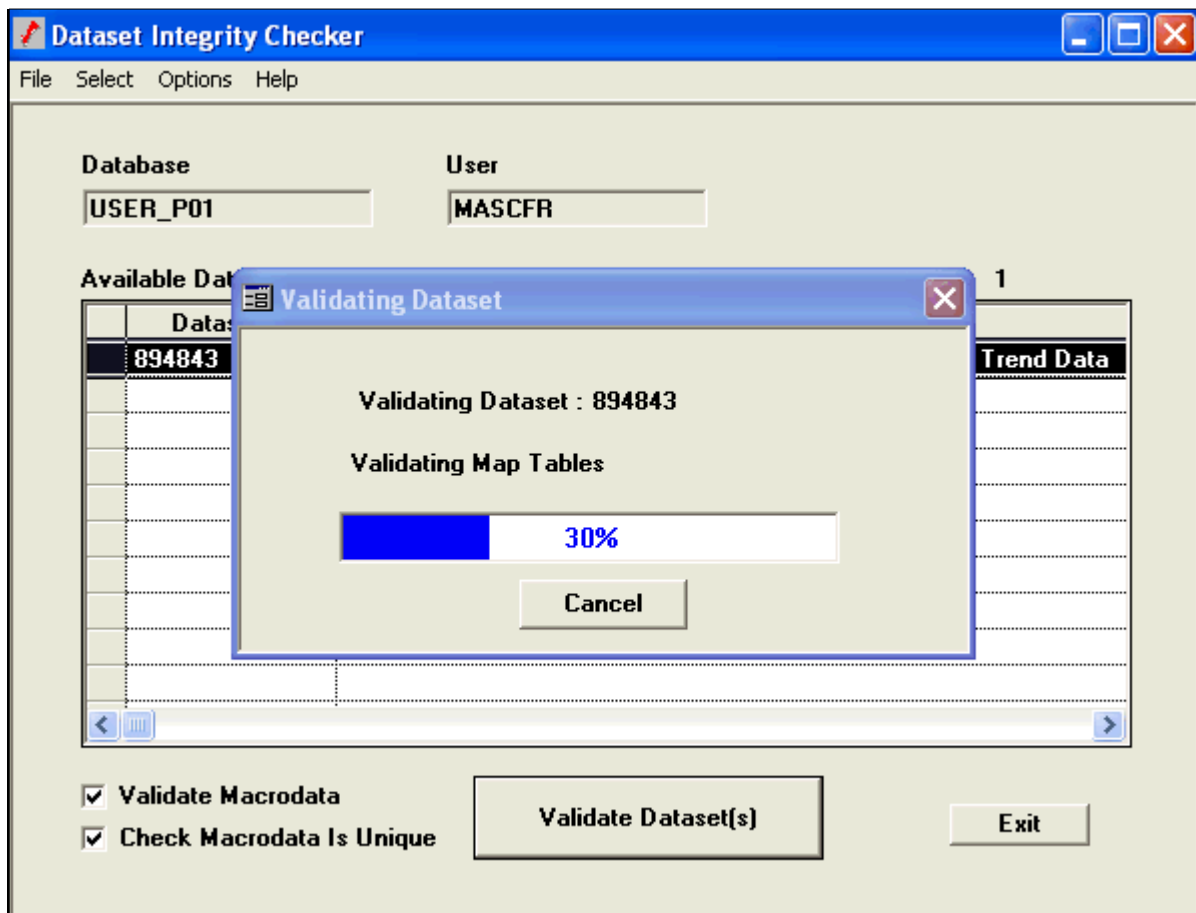
The new dataset ID is: 894843. Make a note of this for use below. Click on "Ok" at the bottom of the above window (not shown here). A validation window will then pop-up:



Click "OK" and then another window will pop-up prompting you to validate the dataset, e.g:

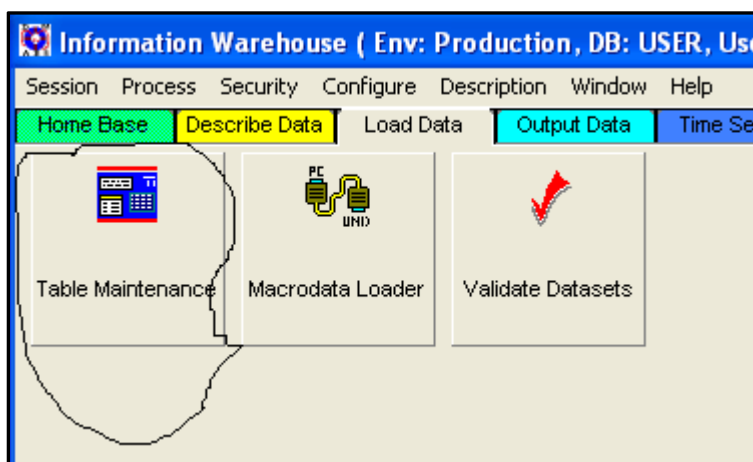


Click on the "Validate" button and then on another validation button in the next window to complete the validation, e.g:

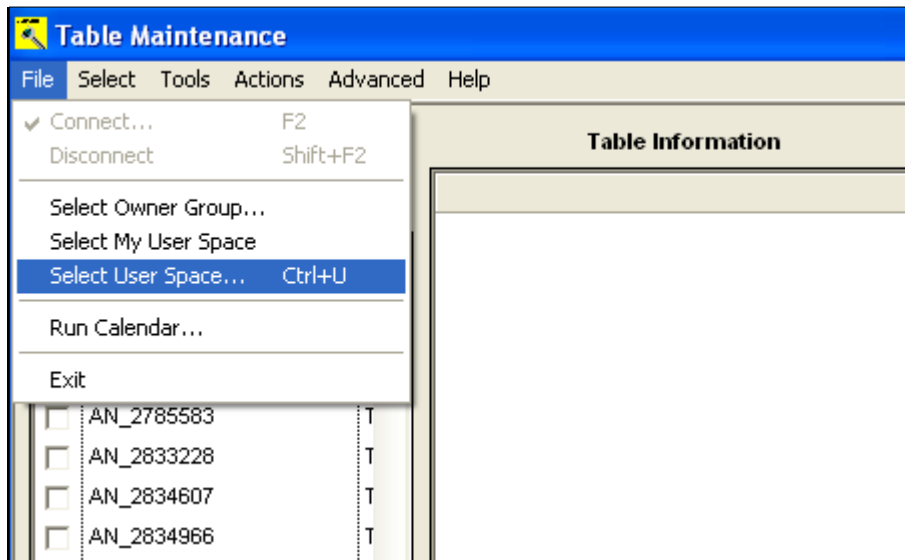


After validation, you can close all the validation and "dataset specifier tool" windows.

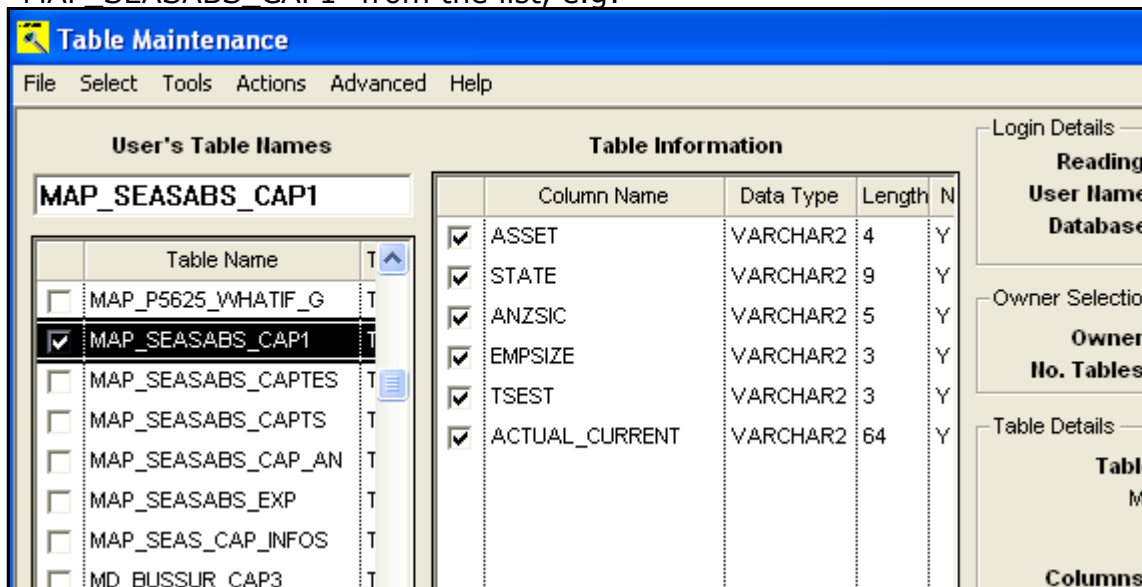
**(4).** Go back to the ABSIW and bring up the "Table Maintenance" tool, e.g:



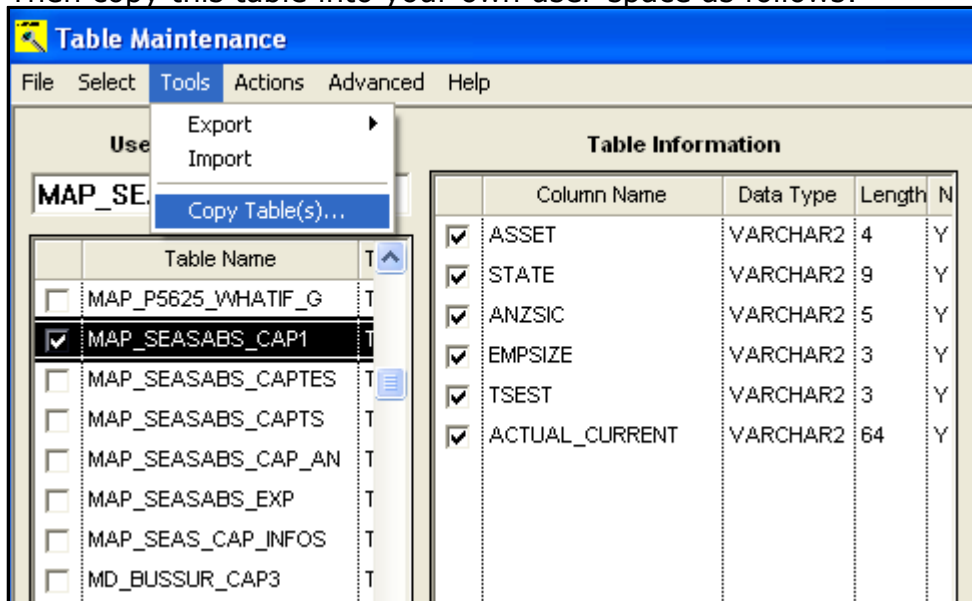
You are going to copy the original map table identified from step 2 above from the original user-space to your own user space. On the window that pops-up, select the user-space that you identified from step 2, e.g:



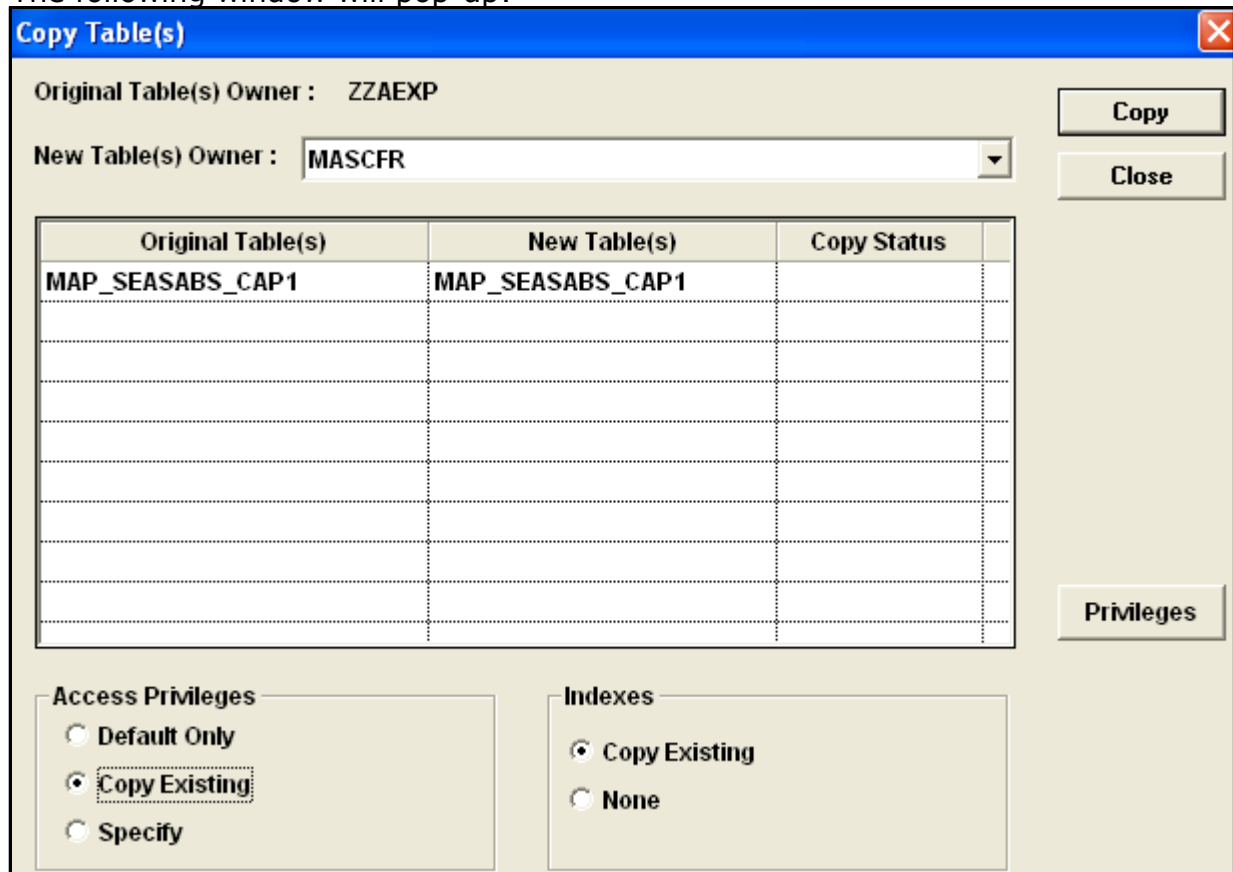
and then enter "ZZAEXP". You can then select the original map name "MAP\_SEASABS\_CAP1" from the list, e.g:



Then copy this table into your own user-space as follows:

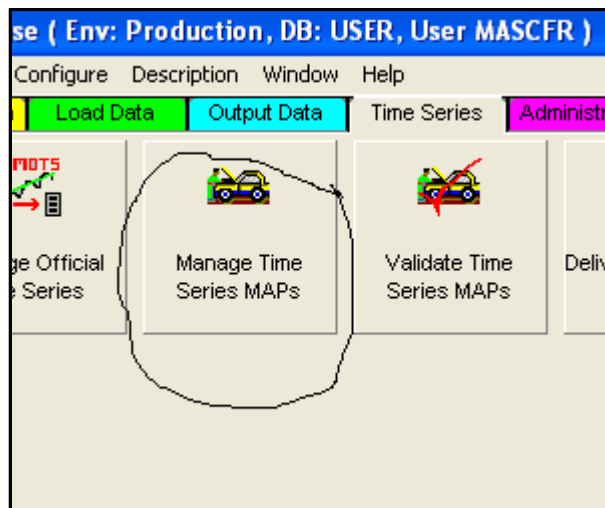


The following window will pop-up:



Ensure you select the "Copy Existing" button and then click "Copy". This will copy the table into your user space (here it's MASCFR). You can now close all windows associated with the "Table Maintenance" tool.

**(5).** Go back to the ABSIW and select the "Manage Time Series MAPs" tool, e.g:



We are now going to create a new TSLink (=new SEASABS group name) that links the new dataset copy to the map table (see Figure 1). Click OK in the following window to bring up your own datasets and TSLinks (never mind if you don't have any):

**External Link Search Criteria**

**Select By**

- All External Links
- Created by myself
- Created by ...
- Owned by ...

**With a Status of**

- All (Except Deleted)
- Locked
- Unlocked
- Deleted

**Order By**

- Dataset Id
- Dataset Title
- Link Name

**System**

ALL

Ok Cancel

Click on the "New Link" button in the new window that appears:

Copy Link

New Link

Select the new dataset copy that you created in step 3 by searching for its dataset ID (894843) in the following window (i.e. click on "Select Dataset"):

**ELM - New Link**

Dataset ID:

Dataset Title:

**Link**

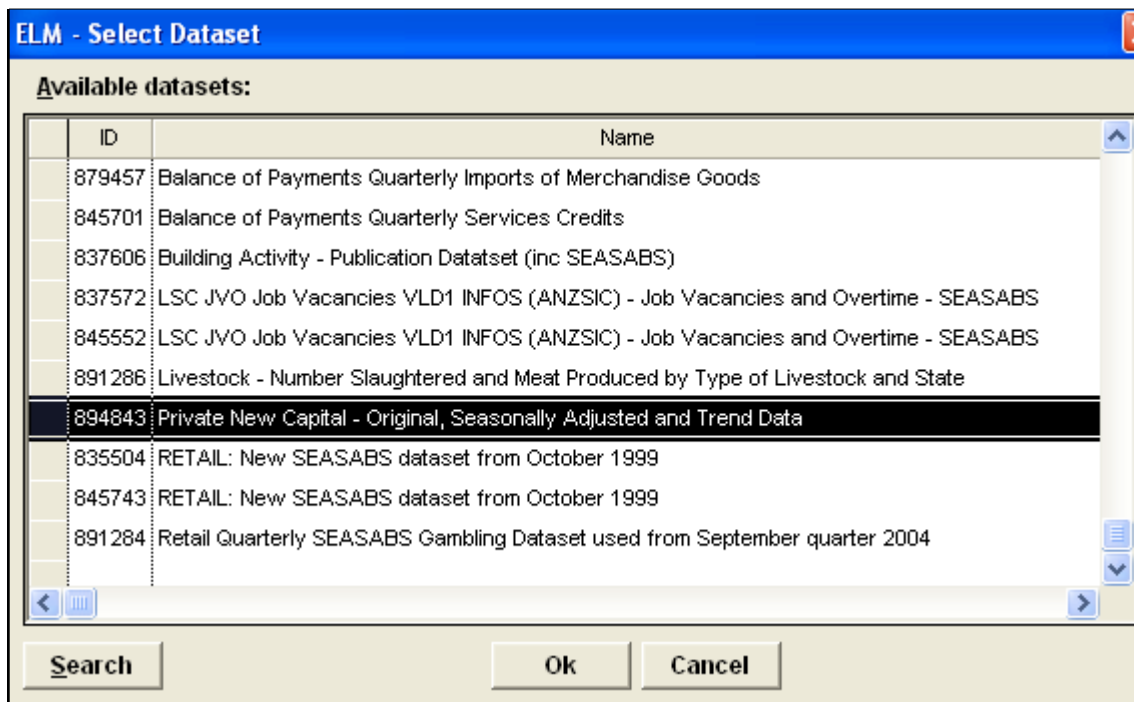
Name:

Comments:

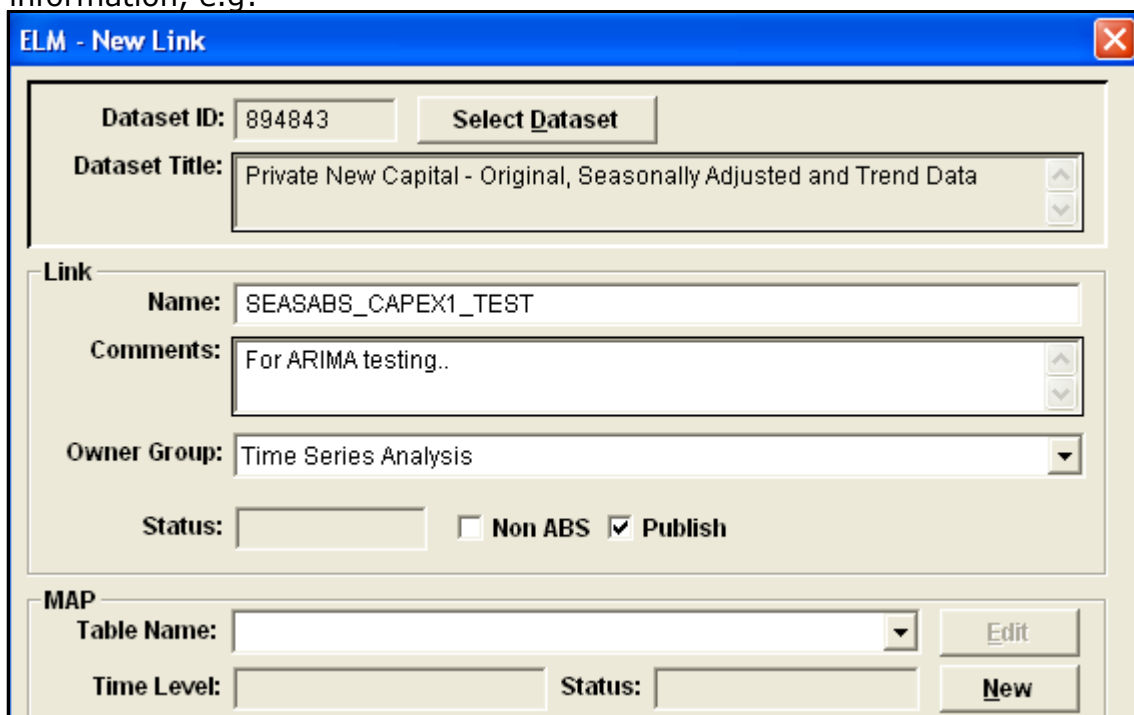
Owner Group:

Status:   Non ABS  Publish

The new dataset ID should be listed in a new window that pops up, e.g:



After clicking OK, a window will pop up where you can enter the new TSLink information, e.g:



After entering the new "Link" details, click on "New" in the "MAP" panel section in the above window. A new window will appear:

**ELM - New MAP Table**

Dataset ID: 894843

Dataset Title: Private New Capital - Original, Seasonally Adjusted and Trend Data

Mode:  Create New Map Table  Select Existing MAP Table

MAP

Table Owner: MASCFR Table Name:

Owner Group: Time Series Analysis

Time Level: QUARTER

Systems:

Valid	System
<input type="checkbox"/>	FAME
<input type="checkbox"/>	GENERIC

Dataset Par Columns:

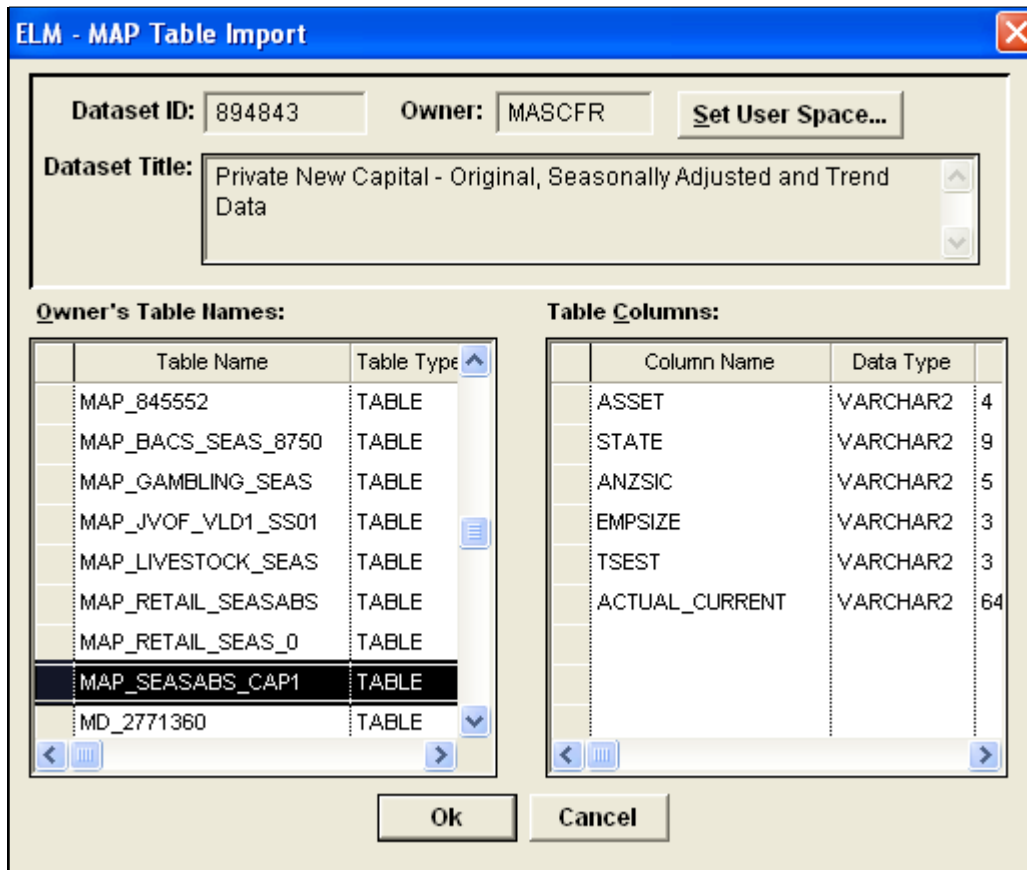
ID	Col Name
78450	ACTUAL_CURRENT

Dataset CIs Columns:

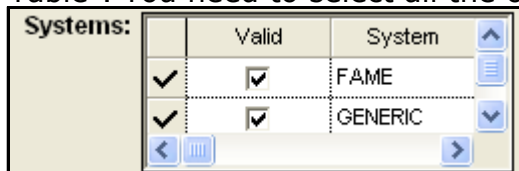
ID	Col Name	Item Name	Include	Theme	Item Type
-2159	QTR	Quarter	<input checked="" type="checkbox"/>	TIME	Column
101512	STATE	State	<input checked="" type="checkbox"/>	N/A	Column
101516	ANZSIC	Industry	<input checked="" type="checkbox"/>	N/A	Column

More Info Select Ok Cancel

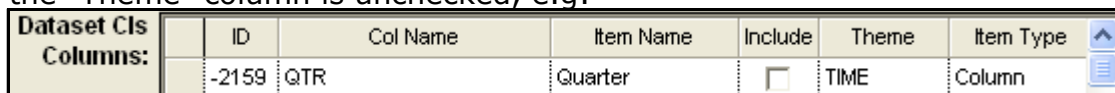
Ensure you select the button: "Select Existing MAP table" and then click on "Select" at the bottom of the above window. The following window will pop-up where you can select the appropriate map table to link to. In this example it's "MAP\_SEASABS\_CAP1":



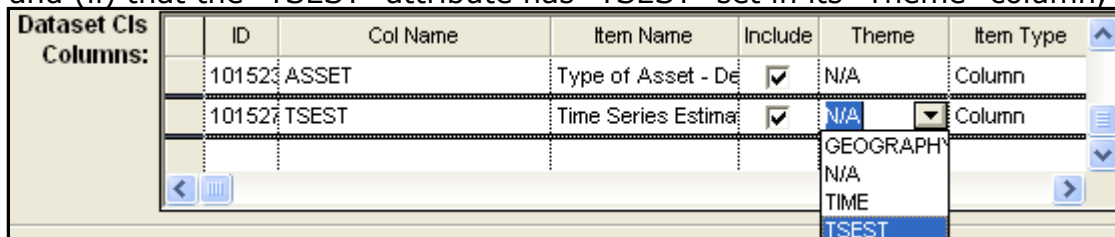
Click OK in this window and then go back to the window titled: "ELM - New MAP Table". You need to select all the options in the "Systems:" box, e.g:



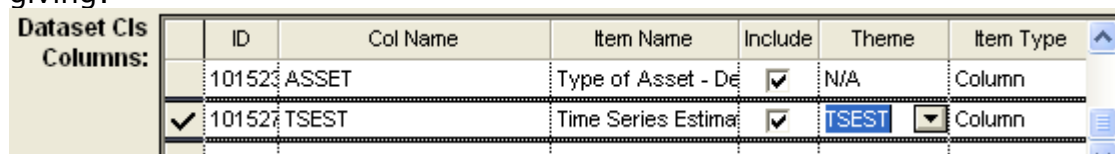
and then in the "Dataset CIs Columns" box, ensure that (i) the "TIME" option in the "Theme" column is unchecked, e.g:



and (ii) that the "TSEST" attribute has "TSEST" set in its "Theme" column, e.g:



giving:



Click OK in this window and also the window titled: "ELM - New Link". The new TSLink that links to the new dataset will now be visible in the "External Link Manager" window, e.g:

**External Link Manager**

File Edit Link Select Options Help

**Available External Links**

Sort by

- Dataset ID
- Dataset Title
- Link Name

Dataset Id	Link Name	Dataset Title
879457	TEST_SEASABS_QGOODSDR	Balance of Payments Quarterly Imports of Me
879517	TEST_SEASABS_QGOODSCR	Balance of Payments Quarterly Exports of M
891284	RETAIL_SEASABS_GAMBLING	Retail Quarterly SEASABS Gambling Dataset
891286	LIVESTOCK_SEAS_TEST	Livestock - Number Slaughtered and Meat Pr
894843	SEASABS_CAPEX1_TEST	Private New Capital - Original, Seasonally Ad

Link Name: SEASABS\_CAPEX1\_TEST      System: FAME, GENERIC, SEAS      **Edit Link**

MAP Table Name: MASCFR.MAP\_SEASABS\_CAP1      Status: Unlocked      **Copy Link**

Type: ABS Internal      **New Link**

Owner Group: Time Series Analysis

Comments: For ARIMA testing..      **Validate**

Created by: MASCFR on 04-Oct-2006      Modified by: MASCFR on 04-Oct-2006      **Exit**

You are now ready to validate this new TSLink. Click on the "Validate" button in the above window. Another window will pop-up confirming the TSLink you wish to validate, e.g:

**External Link Manager - Validation Tool**

File Edit Select Integrity Log Help

**Sort by**

- Dataset ID
- Dataset Title
- Link Name

**Available External Links**

	Dataset Id	Link Name	Dataset Title
<input checked="" type="checkbox"/>	894843	SEASABS_CAPEX1_TEST	Private New Capital - Original, Seasonally Ad
<input type="checkbox"/>			
<input type="checkbox"/>			
<input type="checkbox"/>			
<input type="checkbox"/>			
<input type="checkbox"/>			

**Link Name:** SEASABS\_CAPEX1\_TEST      **System:** FAME, GENERIC, SEASABS

**MAP Table Name:** MASCFR.MAP\_SEASABS\_CAP1      **Status:** Unlocked

**Owner Group:** Time Series Analysis

**Type:** ABS Internal

**Comments:** For ARIMA testing..

**Created by:** MASCFR on 04-Oct-2006      **Modified by:** MASCFR on 04-Oct-2006

**Validate**      **Exit**

Click on the "Validate" button to complete the validation. You will be prompted for an output log filename after which, if all is successful, you will get the message:

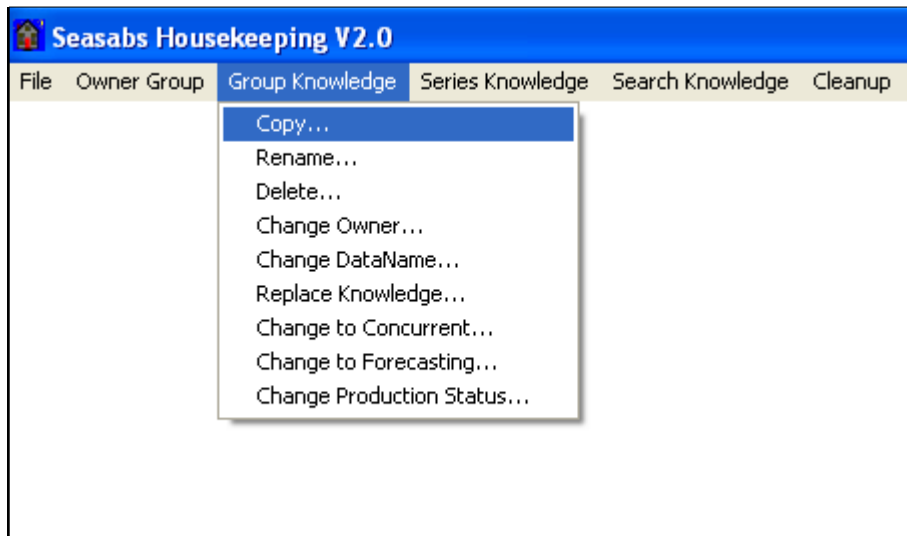
**Validation Results**

A total of 1 links were checked.

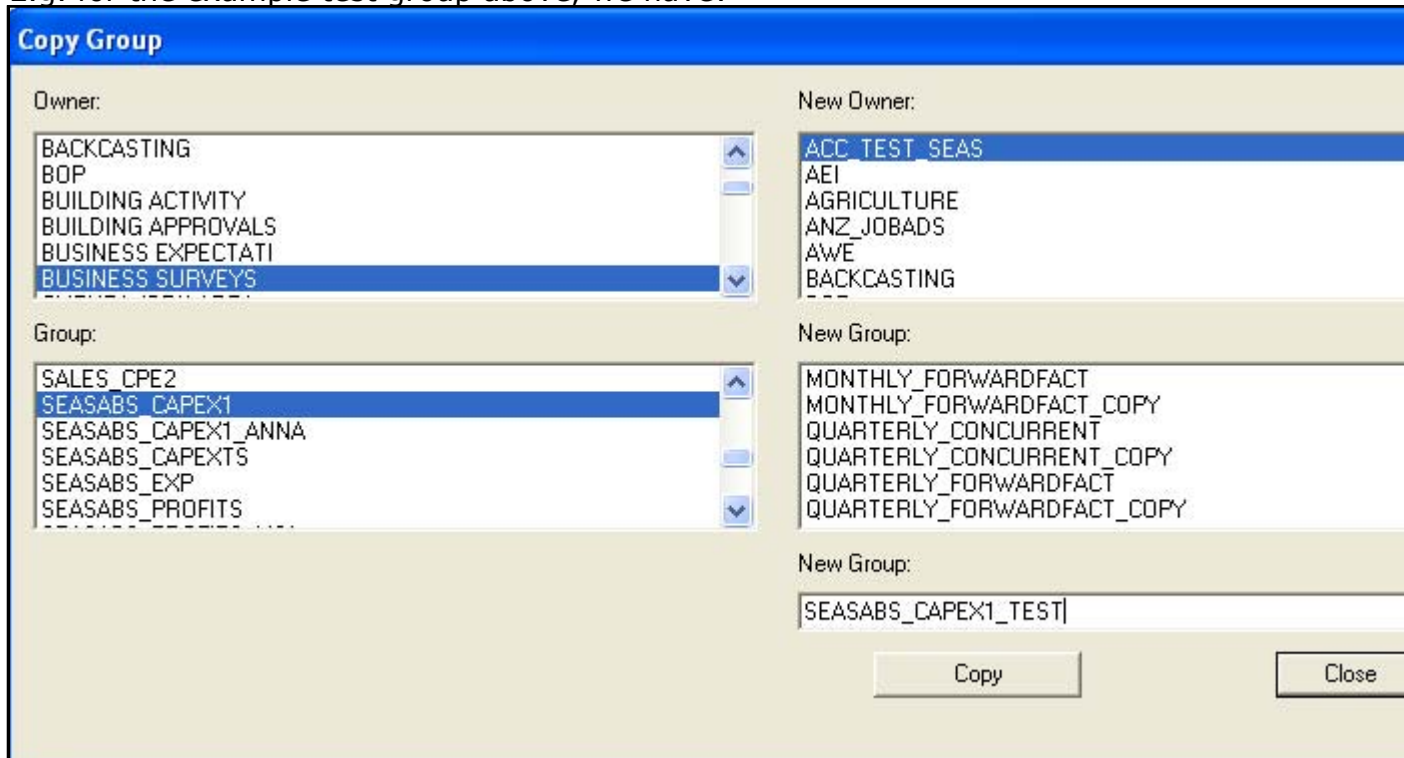
Link	Passed
894843: SEASABS_CAPEX1_TEST	Y

Make a note of the new TSLink you created. This will be used as the group name in SEASABS. You are now done with the ABSIW. Feel free to close all ABSIW windows.

**(6).** Bring up the SEASABS "Housekeeping" facility and copy over the "Group Knowledge" from the original group into your new group. The name of this new group **must be exactly the same** as the TSLink name created in step 5 above (i.e. 'SEASABS\_CAPEX1\_TEST').



E.g. for the example test group above, we have:



**(7).** Now comes the crucial checking step. We want to ensure that modifying the series knowledge of any series in our new test group does not update (hence corrupt) products from the original group that's being used for production.

Here is the checking procedure:

- (i)** bring up the expert version of SEASABS;
- (ii)** open up a series from your new test group;
- (iii)** modify a large extreme or any other prior that will lead to a noticeable difference in say the trend value;
- (iv)** perform a 'research' on this to see the effects of the modified prior. Note down some value (e.g. a trend value);

**(v)** force this 'temporary' modification to the series knowledge. We will undo this later;

**(vi)** open up the same series from the original (old) group (i.e. that's used for production);

**(vii)** perform a "View -> Virtual Update" and examine the same value from a product under the "Publication" menu that was modified in step (iv) (e.g. the 'ABS Trend' is a good one);

**(viii)** if the value in this production group series is different from that forced to be modified in step (iv) above, you have succeeded, otherwise, you're in trouble. The latter means that the new test series knowledge is being used to create/update publication products. This is bad and you must retrace your steps.


**(ix)** assuming you succeeded, go back to the test group series whose prior was modified in step (iii) and restore to its original value.

**(8).** As mentioned in section 4, if you intend to use your new test group for 'intensive' continuous testing that may place a strain on the USER DB, you will need to perform the above procedure again in the user-acceptance test database UAC4. Note that the contents of USER are automatically copied to UAC4 ~every 6 as part of the replication cycle. So, you may get away with repeating the above depending on when the next replication date will be. It is a good idea to check with data management (e.g. Phil Carruthers).

Congratulations, you are now done.

## **6. GUAT versus UUAT**

The following document describes the "Global User Acceptance Testing" (GUAT) infrastructure:

 (Subject: SEASABS/TSUpdate acceptance test procedure.; Database: Time Series Analysis WDB; Author: Frank Masci; Created: 03/04/2006; Doc Ref: FMAI-6NH4R7).

GUAT is normally conducted by a single member of TSA as a new SEASABS version release date approaches. For the version 2.6 release (Dec 2006), this will be performed by Frank Masci. This is opposed to individual Service Request (SR) "User Unit Acceptance Testing" (UUAT) which is normally performed by the initiator of the SR.