

////////////////////////////////////
////////////////////////////////////

Change Initiator

Name: **Craig McLaren**

Requirement Details

Requirement Summary: **initial coherence information to be stored in series knowledge**

Impact: **New functionality**

Due Date: **20/10/2005**

Dependent Events
:

Related Requests:

Background Information:

We need facilities within SEASABS to assist in managing time series which are "conceptually related", to assist in assessing levels of consistency through diagnostics and summary measures, and, to assist in transferring selected priors from one series knowledge to another for a related series pair. This project is broken into three phases:

- Phase 1: creation, storage and management of related series identifiers
- Phase 2: the development of a user-interface to compute, display and store consistency summary measures and diagnostics
- Phase 3: the development of a user-interface to transfer and apply selected prior factors from one series to its related counterpart

Disclaimer: please avoid using words like "coherence", "coherent" or derivatives thereof in any of this work (eg. window titles, labels etc..). Words like "related" or "consistent" (the latter in the right context) should be used.

Requirement Description:

Phase 1: Creation and storage of related series in the series knowledge

Functionality to record group and series name of potential related series in the series knowledge.

a) perhaps record under series options, ability to enter in more than one time series as consistent.

b) group name and link to where group resides in SEASABS, ie. GROUP.SERIES ID

c) it's up to user to enter this information "manually" - preferred option is with a pick list from available time series across groups and owner groups.

d) a warning flag to be issued under a "Research Analysis" (not concurrent) run that displays a message that this series may be related to other time series with a list of those particular time series. This warning should appear at the end of the reanalysis.

e) ability to click/select on a consistent time series and SEASABS opens this series (similar to the aggregation structure which allows clicking on a name aspect).

f) ability to store in the series knowledge information (metadata) about the level of consistency or what makes them consistent. For example, a description of the way in which highlighted series are consistent. Eg. "Series A comprises 90% of series B so movements should be similar, except in April due to seasonal rise in parakeet sales (parakeet sales not recorded in series A, but recorded in series B)".

g) ability to export a list of all related series names to DOS as a text/ascii file. More specifically, all series in a group that have been declared to be related to another series in either same group, other groups (including other owner groups) should be saved. The format of this output list should be for example:

```
<ownergroup>,<group>,<seriesname1a>,<ownergroup>,<group>,<seriesname1b>
```

```
<ownergroup>,<group>,<seriesname2a>,<ownergroup>,<group>,<seriesname2b>
```

```
<ownergroup>,<group>,<seriesname3a>,<ownergroup>,<group>,<seriesname3b>
```

etc..

h) ability to import a text/ascii file containing a list of related series pairs as in the example in (h). This obviates the need for a user to manually select related series from a list within SEASABS. On import, this information should be stored directly into the series knowledge.

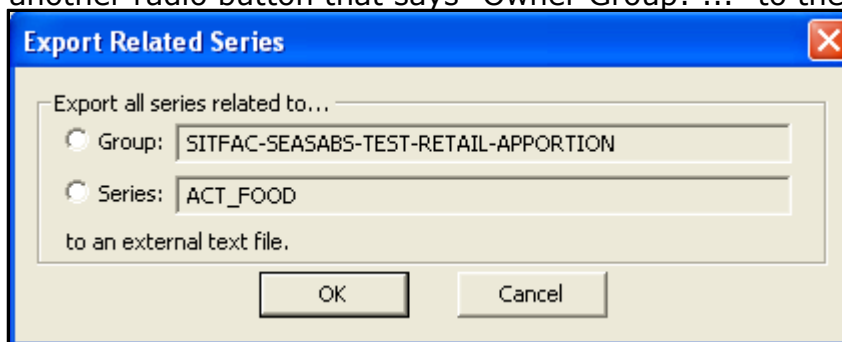
- The relationship between series would be symmetric. For example, if series A is related to series B this means that series B is related to A and so should be reflected in both series knowledge records. This pairing (two-way) association should be implicit in the series knowledge: ie. a user need not explicitly declare A -> B and B -> A separately. When A -> B is selected, then this immediately implies B -> A, ie. "A <-> B" is assumed.
- If A is related to B and B is related to C does that always mean A is related to C?: Possibly - but if we can't reflect this relationship within the series knowledge then we wouldn't be doing this automatically. ie. it would be up to the user to enter this information manually at this stage. The design should be based on explicit user-defined (paired-series) associations.
- We would like to automatically include who declared the association and when.
- If we copy a series or group with the production flag set, then the series knowledge about consistent time series should be retained. The production

group is the "master" version from which deletion of prior declared associations is disallowed without sufficient privileges.

Some fixes to phase 1 of consistency functionality following test version 2.6.13:

These fixes/enhancements to phase-1 can wait until a future test version. These are not a high priority for the version 2.6 production release.

- Ability to import generic text file containing many different related series across any groups and owner groups, not just those series related to the single series loaded in a SEASABS session.
- Ability to export to a text file all series pairs with related information in format: "own1, grp1, sname1, own2, grp2, sname2" across all groups in the owner group that is loaded in a SEASABS session. Right now, this is only being done at the series and group levels. For example, we would like to add another radio button that says "Owner Group: ..." to the following options:



- Some diagnostic windows and messages are still showing "CORS.." in their title. Please replace this with "Related series info..".

Phase 2: Facility/user-interface to compute, display and allow storage of consistency summary measures and diagnostics

We envisage a summary table (visualisation tool) that lists all related series pairs for a group and the "Pass"/"Fail" (P/F) results of a number of diagnostic tests computed (on the fly) on at least three different series components for each pair. There will also be in-built buttons to control processing, set configuration parameters and store finalised consistency measures and quality flags. An example of this table is below. See the notes that follow for further details.

Series Pair	Compute / Review	Originals	ABS-SA	D11-SA	Select to Store
<input type="checkbox"/> pair 1	<input checked="" type="radio"/> <input type="radio"/>	P: 1, 2, 5, 6 F: 3, 4	P: 1, 2, 3 F: 4, 5, 6	P: 1, 2, 3, 6 F: 4, 5	<input checked="" type="radio"/>
<input type="checkbox"/> pair 2	<input type="radio"/> <input checked="" type="radio"/>	P: 2, 4 F: 1, 3, 5, 6	etc..	etc..	<input type="radio"/>
<input type="checkbox"/> pair 3	<input type="radio"/> <input checked="" type="radio"/>	P: 1, 2, 3, 4, 5, 6 F: none	.	.	<input type="radio"/>
<input type="checkbox"/> pair 4	<input type="radio"/> <input checked="" type="radio"/>	P: 1, 2, 3, 6 F: 4, 5	.	.	<input checked="" type="radio"/>
.
.
<input type="button" value="Compute All"/>	<input type="button" value="Configure Tests"/>				<input type="button" value="Select/Deselect All"/>

Clicking on this will populate the Pass and Fail ("P" and "F") test numbers for all three series components and all pairs in this table in one fell swoop.

Depending on "Compute/Review" selection in col. 2, clicking on any one of these will spawn a HTML table containing either:

- Current SEASABS parameters and new computed "on the fly" consistency diagnostics and plots, OR,
- previous SEASABS parameters and consistency diagnostics ("CQMs") from series knowledge for a requested input date.

Same for all Series Components:

- Time spans on which to apply all tests, defaulting to series with shortest available span in each pair
- Critical P-values for tests 2,3,4,5,6 with defaults
- Outlier detection threshold (#sigma)
- Movement difference threshold (test 5)
- Max. identical movement tolerance
- Time spans in temporal-movement plots
- For "Review" mode, requested date of a previous consistency analysis (hence storage) run.
- Other..?

1. All the series pairs listed in column one are all those series in the opened group that are related to any other series from the same group, other groups in the same owner group, or, other groups from different owner groups.

2. If "Compute" mode is selected in column two, then clicking on a series pair in column one will bring up a html formatted table listing (i) SEASABS parameters and settings, and, (ii) consistency test diagnostics, measures and clickable plots for the three series components listed in the next three columns of the above table. These should all be generated "on the fly". See item 3 below if "Review" mode is selected from column two. The SEASABS parameters and settings will be identical for all three series components so should only be shown once in the output. The test diagnostics however will be different for each series component. An example html output for one series component (as generated by our current prototype tool) is:

[S:\data\NAB\Coherence\ExampeOutput Don'tDelete!\example_SA_component.html](#) . The diagnostic measures for the three series components (in the order *Originals* ; *ABS-SA* and *D11-SA*) shall be displayed after the SEASABS parameters summary table in separate sections of the html output. So, for example, for a single series pair ("row" in the above table), the html output format shall be:

Series Pair: <ownergroup.group.series_1>, <ownergroup.group.series_2>

Consistency Analysis Date: <either current date if "Compute" or specific date from series knowledge if "Review" is selected>

SEASABS parameters and settings		
parameter	series 1	series 2

Consistency Diagnostics for Originals
TEMPORAL MOVEMENT PLOTS (hyperlink)...
TEST 1: ...
TEST 2: ...
TEST 3: ...
TEST 4: ...
TEST 5: ...
TEST 6: ...

Consistency Diagnostics for ABS Seasonally Adjusted
TEMPORAL MOVEMENT PLOTS (hyperlink)...
TEST 1: ...
TEST 2: ...
TEST 3: ...
TEST 4: ...
TEST 5: ...
TEST 6: ...

Consistency Diagnostics for D11 (i.e. X11) Seasonally Adjusted
TEMPORAL MOVEMENT PLOTS (hyperlink)...
TEST 1: ...
TEST 2: ...
TEST 3: ...
TEST 4: ...
TEST 5: ...
TEST 6: ...

Example HTML output for a series pair.

3. Column two labelled with "Compute / Review" allows the user to either review Consistency Quality Measures (CQMs) from previous analyses if available (using an input previous consistency analysis date; see item 10), or, re-compute

new measures for the specified analysis span (defaulting to current analysis span) when a series pair in column one is clicked. HTML formatted output tables are expected under each mode. The only difference for the "Review" mode is that (i). previous CQMs are read from the series knowledge history and written to the HTML output (defaulting to N/A if none exist) and (ii). there are no clickable hyperlinks to plots. On startup of the above table/GUI, the "Compute / Review" field should have all selections initialised to "Compute".

4. The summary table of "SEASABS parameters and settings" in the example output:

["S:\data\NAB\Coherence\ExampeOutput_Don'tDelete!\example_SA_component.html"](S:\data\NAB\Coherence\ExampeOutput_Don'tDelete!\example_SA_component.html) is not complete. They were generated by parsing the outputs from TSA-download. We would like to include the following additional parameters and values for each series:

- adjustment type: "concurrent" or "forward factor";
- method: value can be either "direct", "indirect", or "not adjusted". Note: our prototype tool currently only reports "direct" or "indirect" (as available from TSA-download). "not adjusted" cases must be available from the series knowledge, right? If so, we want to be able to report "not adjusted" cases in the method field.
- Star value;
- I/C ratio;
- I/S ratio;
- If static trading day exists, the trading day weights (table C15 SEASABS outputs) for each series listed in the "Trading Day" row with an indication that these are static trading day weights;
- If moving trading day exists, the trading day weights (table C15 SEASABS outputs) for the last year available in each series listed in the "Trading Day" row. There should also be an indication that these are "year XXXX" moving trading day weights;

5. It is requested that all prior corrections (namely the trend breaks, seasonal breaks, large-extremes, additive outliers) be listed one per line and in time order in the "SEASABS parameters and settings" table output.

6. Algorithms for computing all consistency diagnostics and plots (for each of the tests in the example output above) have been implemented in a prototype tool written in "Perl" and "R":


<S:\data\BOP\coherence\Frank'sCode\consistency.pl>

7. In one of the diagnostic plots (namely the movement versus movement plots) that are generically named "mvt_vs_mvt_pair_<?>.png" by TSA's prototype tool, we would like to add another feature. For any outliers in this plot (currently labelled by a "X" symbol; see example in

S:\data\NAB\Coherence\ExampeOutput_Don'tDelete!\example_SA_component.html), we would like the ability to click on an outlier and a pop-up window would appear showing:

- the time point " t " to which the time-matched movements apply
- any prior corrections at that time point for the two series
- simple line plots of series 1 and series 2 (ie. of the series components being examined on input) at times " $t-3$, $t-2$, $t-1$, t , $t+1$, $t+2$, $t+3$ ". These plots should be side-by-side.

8. In all plots and diagnostic outputs where we refer to time points, we would like these written in "long format" (ie. as used in SEASABS plots), eg: Jan 1994 or Qtr1 1994 etc..

9. The "Originals", "ABS - SA" and "D11 - SA" columns in the above table represent respectively the Original, ABS seasonally adjusted and D11 (X11 seasonally adjusted) series components. These are all the ones that matter for now. Under these columns are the "Pass" (P) and "Fail" (F) results of the six diagnostic tests summarised in the html output. These should be computed (or updated with historical values if "Review" is selected) for a specific series pair (for all associated components) "on the fly" each time a user clicks on that series pair in column 1. Note that this operation will also spawn the html table. Alternatively, the user can click on the "Compute All" button which will perform the diagnostic tests internally and populate all the "P" and "F" fields for all series components automatically. On startup of the above table/GUI, the "P" and "F" values should reflect the latest that are stored in the series knowledge (see stored quality flags/measures in item 12 below) if available. The "P" and "F" flags are defined from the outcome of the six diagnostic tests. These are defined in *section 3* of  (Subject: Quantifying and Improving Consistency: version 3; Database: Time Series Analysis WDB; Author: Frank Masci; Created: 16/02/2006; Doc Ref: FMAI-6M33Y3).

10. The button labelled "Configure Tests" allows the user to adjust the input parameters for each of the diagnostic tests (as shown). These should default to appropriate values which are currently hardcoded in our prototype consistency tool: <S:\data\BOP\coherence\Frank'sCode\consistency.pl> . Included in this list is an option to specify the date of the last stored "consistency analysis" (that performed just prior to this date) for use with the "Review" mode in column two. It is advised that the developer contact Frank Masci (x5806) when this stage is reached to ensure the details are well understood.

11. The column labelled "Select to Store" gives the user the option to select those series pairs for which the results of any updated consistency diagnostic results (and quality flags) should be stored in the series knowledge following a Force -> "Update Series Knowledge" at the end of a "Research Analysis" run. For flexibility, the user can also use the "Select/Deselect All" button to select all pairs and store the consistency results/flags for all of them in the series knowledge.


12. As a first cut, the following Consistency Quality Measures (CQMs) and flags should be stored in the series knowledge. These have been generically named CQM-1 to CQM-156..(?) for ease of storage and allowance for changes in definitions if needed in future. We have also defined 'null' measures to serve as placeholders in case more definitions are needed. It is envisaged that these will be stored for a maximum of five years in order to monitor consistency quality over time.

- CQM-1: Date string "dd/mm/yyyy" pertaining to when consistency diagnostics were last stored for the series pair. Note that the "Review" option in column two will use this string to retrieve historical CQMs.
- CQM-2: String of test numbers with P-flag for Orig, ABS-SA, D11-SA components: e.g. pair 1 in table above will have: "1,2,5,6-1,2,3-1,2,3,6".
- CQM-3: String of test numbers with F-flag for Orig, ABS-SA, D11-SA components: e.g. pair 1 in table above will have: "3,4-4,5,6-4,5".




- CQM 4: 'null' --- (placeholder)
- CQM 5: 'null' --- (placeholder)
- CQMs-6 to ~156: Diagnostic measures represented by all those strings starting with a "*" in the HTML output (see prototype example output: S:\data\NAB\Coherence\ExampeOutput_Don'tDelete!\example_SA_component.html). This does not include strings which are hyperlinks to plots. This CQM range is expected to list, in consecutively order, diagnostics for each of the components: Orig, ABS-SA and D11-SA. Each component should be given a CQM range of length 50.

Phase 3: Facility to transfer and apply selected prior factors from one series to its related counterpart

In all cases, a warning should be issued displaying a message that initial prior(s) in series "A" for example will be overridden with those from series "B" in a SEASABS research run. The new priors should be forced (overriding all previous priors with new proposed corrections) to the series knowledge if confirmed by the user at the end of the reanalysis.

This phase is still under design and some suggestions are outlined in sections 8 and 9 of  (Subject: 887/34: comparison of series knowledge across groups; Database: TA Service Request System; Author: Craig McLaren; Created: 22/06/2003).

Summary of doclinks/files relevant to the above:

- Example html output from TSA's prototype tool: S:\data\NAB\Coherence\ExampeOutput_Don'tDelete!\example_SA_component.html
- TSA's prototype tool for computing consistency diagnostics: <S:\data\BOP\coherence\Frank'sCode\consistency.pl>
- Definitions of "Pass" and "Failure" modes for diagnostic tests:  (Subject: Quantifying and Improving Consistency: version 3; Database: Time Series Analysis WDB; Author: Frank Masci; Created: 16/02/2006; Doc Ref: FMAI-6M33Y3)
- Phase 3 suggested design:  (Subject: 887/34: comparison of series knowledge across groups; Database: TA Service Request System; Author: Craig McLaren; Created: 22/06/2003)
- Subsidiary SR to enable consistency checks between indirect and component series (mostly TSUpdate):  (Subject: 887/222: Functionality to enable consistency checks between indirect and component series; Database: TA Service Request System; Author: Frank Masci; Created: 21/03/2006)

Acceptance Criteria: Works as specified

////////////////////////////////////
 //////////////////////////////////

Change Initiator

Name: **Frank Masci**

Requirement Details

Requirement Summary: **Copy of initial coherence information to be stored in series knowledge**

Impact: **New functionality**

Due Date: **15/12/2006**

Dependent Events
:

Related Requests: **887/180**

Background Information:

We need facilities within SEASABS to assist in managing time series which are "conceptually related" and to assist in assessing levels of consistency through diagnostics and summary measures. This project is broken into two phases:

- Phase 1: creation, storage and management of related series identifiers
- Phase 2: the development of a user-interface to compute and display consistency summary measures

Disclaimer: please avoid using words like "coherence", "coherent" or derivatives thereof in any of this work (eg. window titles, labels etc..). Words like "related" or "consistent" (the latter in the right context) should be used.

Requirement Description:

Phase 1: Creation and storage of related series in the series knowledge

a) perhaps record under series options, ability to enter in more than one time series as consistent.

b) group name and link to where group resides in SEASABS, ie. GROUP.SERIES ID

c) it's up to user to enter this information "manually" - preferred option is with a pick list from available time series across groups and owner groups.

d) a warning flag to be issued under a "Research Analysis" (not concurrent) run that displays a message that this series may be related to other time series with a list of those particular time series. This warning should appear

at the end of the reanalysis.

e) ability to click/select on a consistent time series and SEASABS opens this series (similar to the aggregation structure which allows clicking on a name aspect).

f) ability to store in the series knowledge information (metadata) about the level of consistency or what makes them consistent. For example, a description of the way in which highlighted series are consistent. Eg. "Series A comprises 90% of series B so movements should be similar, except in April due to seasonal rise in parakeet sales (parakeet sales not recorded in series A, but recorded in series B)".

g) ability to export a list of all related series names to DOS as a text/ascii file. More specifically, all series in a group that have been declared to be related to another series in either same group, other groups (including other owner groups) should be saved. The format of this output list should be for example:

```
<ownergroup>,<group>,<seriesname1a>,<ownergroup>,<group>,<seriesname1b>
```

```
<ownergroup>,<group>,<seriesname2a>,<ownergroup>,<group>,<seriesname2b>
```

```
<ownergroup>,<group>,<seriesname3a>,<ownergroup>,<group>,<seriesname3b>
```

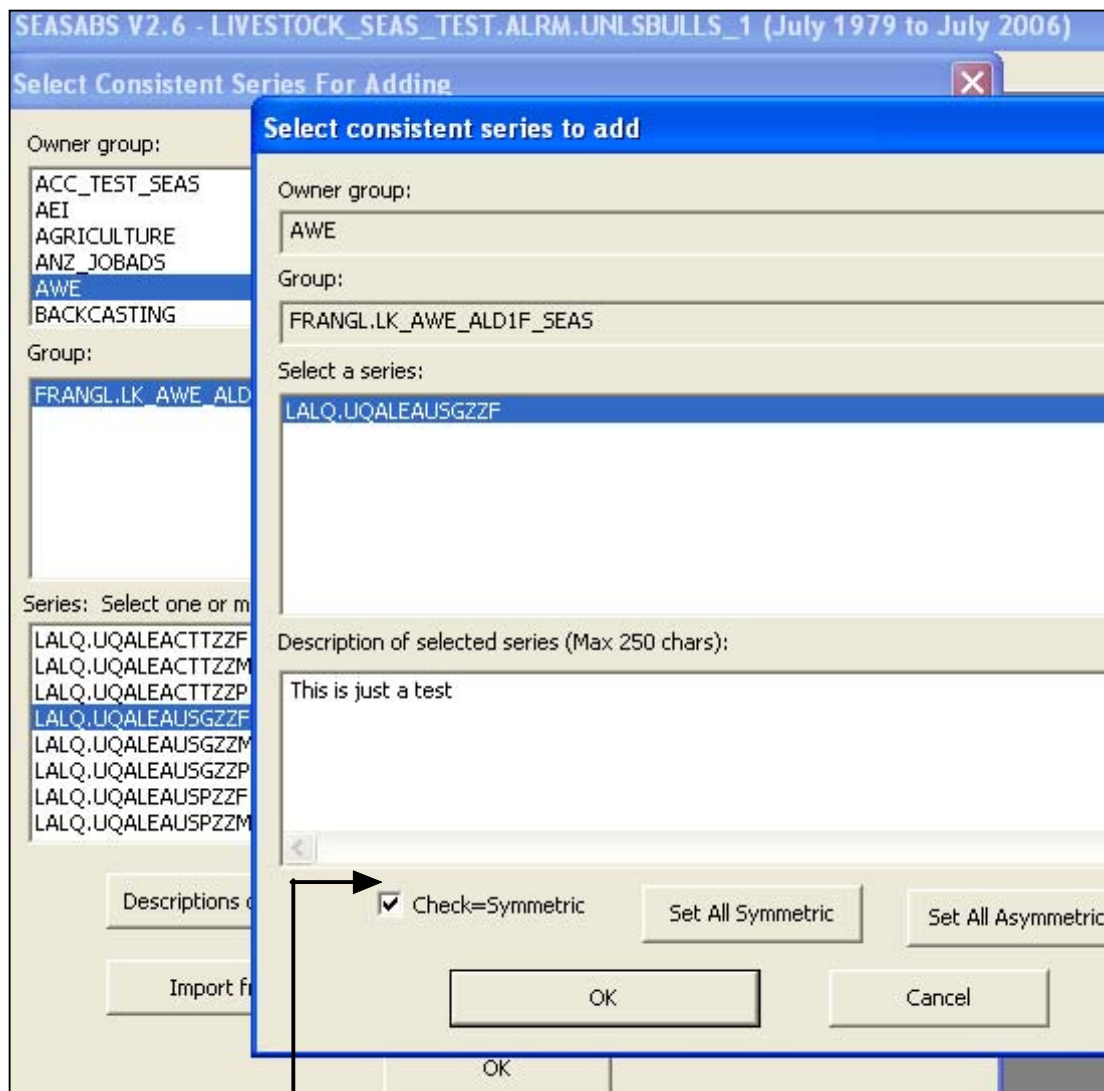
etc..

h) ability to import a text/ascii file containing a list of related series pairs as in the example in (h). This obviates the need for a user to manually select related series from a list within SEASABS. On import, this information should be stored directly into the series knowledge.

- The relationship between series would be symmetric. For example, if series A is related to series B this means that series B is related to A and so should be reflected in both series knowledge records. This pairing (two-way) association should be implicit in the series knowledge: ie. a user need not explicitly declare A -> B and B -> A separately. When A -> B is selected, then this immediately implies B -> A, ie. "A <-> B" is assumed.
- If A is related to B and B is related to C does that always mean A is related to C?: Possibly - but if we can't reflect this relationship within the series knowledge then we wouldn't be doing this automatically. ie. it would be up to the user to enter this information manually at this stage. The design should be based on explicit user-defined (paired-series) associations.
- We would like to automatically include who declared the association and when.
- If we copy a series or group with the production flag set, then the series knowledge about consistent time series should be retained. The production group is the "master" version from which deletion of prior declared associations is disallowed without sufficient privileges.

Some fixes/enhancements to PHASE 1 functionality following production release v2.6:

- Ability to import generic text file containing many different related series across any groups and/or owner groups, not just those series related to the single series loaded in a SEASABS session. See point (g) above for input format.
- Ability to assign attributes (or "dimension" information) to each series when a user declares a related series pair. This will enable a user to display all related series according to pre-selected attributes. This will later assist when generating consistency-diagnostic information by allowing users to focus on those related series that have a common theme or attribute. The specification for this enhancement can be built on the existing GUI. Here are the details:



Place new dialogue window here

Set series attributes (check all that apply for this pair):

Series 1:

LIVESTOCK_SEAS_TEST.ALRM.UNLSBULLS_1

Series 2:

FRANGL.LK_AWE_ALD1F_SEAS.LALQ.UQALEAUSGZZF

<u>Attribute</u>	<u>Series 1</u>	<u>Series 2</u>
Grand total	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Total industries/one state	<input type="checkbox"/>	<input type="checkbox"/>
Total states/one industry	<input type="checkbox"/>	<input type="checkbox"/>
General lower level	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Chain Volume Measure (CVM)	<input type="checkbox"/>	<input type="checkbox"/>
Current Price Value (CPV)	<input type="checkbox"/>	<input type="checkbox"/>
Monthly	<input type="checkbox"/>	<input type="checkbox"/>
Quarterly	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Conceptually related	<input type="checkbox"/>	<input type="checkbox"/>
Source series	<input type="checkbox"/>	<input type="checkbox"/>
Derived series	<input type="checkbox"/>	<input type="checkbox"/>

Current Price Value (CPV)	<input type="checkbox"/>	<input type="checkbox"/>
Monthly	<input type="checkbox"/>	<input type="checkbox"/>
Quarterly	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Conceptually related	<input type="checkbox"/>	<input type="checkbox"/>
Source series	<input type="checkbox"/>	<input type="checkbox"/>
Derived series	<input type="checkbox"/>	<input type="checkbox"/>

Notes:

- The above is purely illustrative, showing the boxes that a user may check when declaring a related series pair. By default, we would like no boxes checked on start-up. An error should be issued if a user did not check any boxes, i.e. if they inadvertently clicked OK in the above panel without checking any boxes. This error should be issued as a message on the screen, e.g: "no attributes were checked; please go back and tick all applicable attributes".
- The user is free to check more than one box for each series of a related pair. This information should obviously also be stored in the consistency metadata for later reference (see below).
- "**Series 1**" in the above panel should always refer to the series that is loaded in the current SEASABS session (in above example="LIVESTOCK_SEAS_TEST.ALRM.UNLSBULLS_1"). "**Series 2**" is then the series that is manually selected by the user in the "Select Consistent Series For Adding" panel (in above example="FRANGL.LK_AWE_ALD1F_SEAS.LALQ.UQALEAUSGZZF").
- For your information, here are the specific definitions for the above attributes. We envisage this will make its way into the SEASABS help.
 - **Grand total:** E.g. series representing national aggregate over all states and/or all industries within.
 - **Total industries/one state:** E.g. in the case of a 2D aggregation structure, this series could be the "marginal" sum over all industries for a particular state. Note that here "industry" and "state" are arbitrary attributes. These can stand in for other generic classification breakdowns, e.g. "number of juniors employed" versus "employment type".
 - **Total states/one industry:** E.g. in the case of a 2D aggregation structure, this series could be the "marginal" sum over all states for a particular industry or category. Note that here "industry" and "state" are arbitrary attributes. These can stand in for other generic classification breakdowns, e.g. "number of juniors employed" versus "employment type".
 - **General lower level:** Generic low level component (child) series that may be part of a larger aggregation structure. E.g. a series defined for a particular single industry and single state.
 - **Chain Volume Measure (CVM):** a series representing the "chained volume" equivalent of the CPV.
 - **Current Price Value (CPV):** a series representing the "current price" equivalent of the CVM.
 - **Monthly:** a series whose periodicity is monthly.
 - **Quarterly:** a series whose periodicity is quarterly.
 - **Conceptually related:** series that has an explicit conceptual relationship to another series (often declared by client).
 - **Source series:** series representing the "raw input" or progenitor of a collection that feeds into a series used/published by another subject matter area (e.g. input into National Accounts). See also "Derived series".
 - **Derived series:** series that is derived from an input series originating from

another collection (from the "Source series"). Most National Accounts series are "derived".

Phase 2: Facility/user-interface to compute and display consistency summary measures and diagnostics

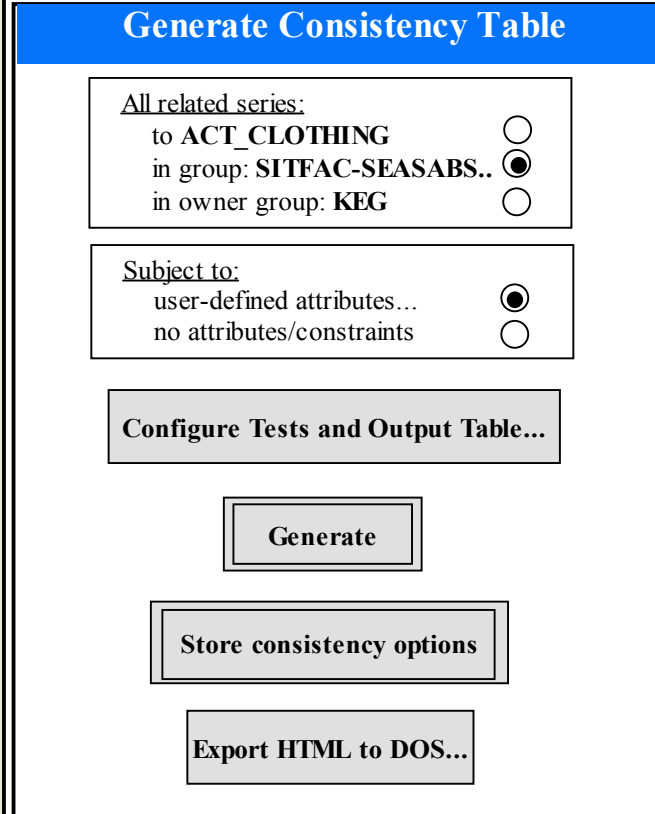
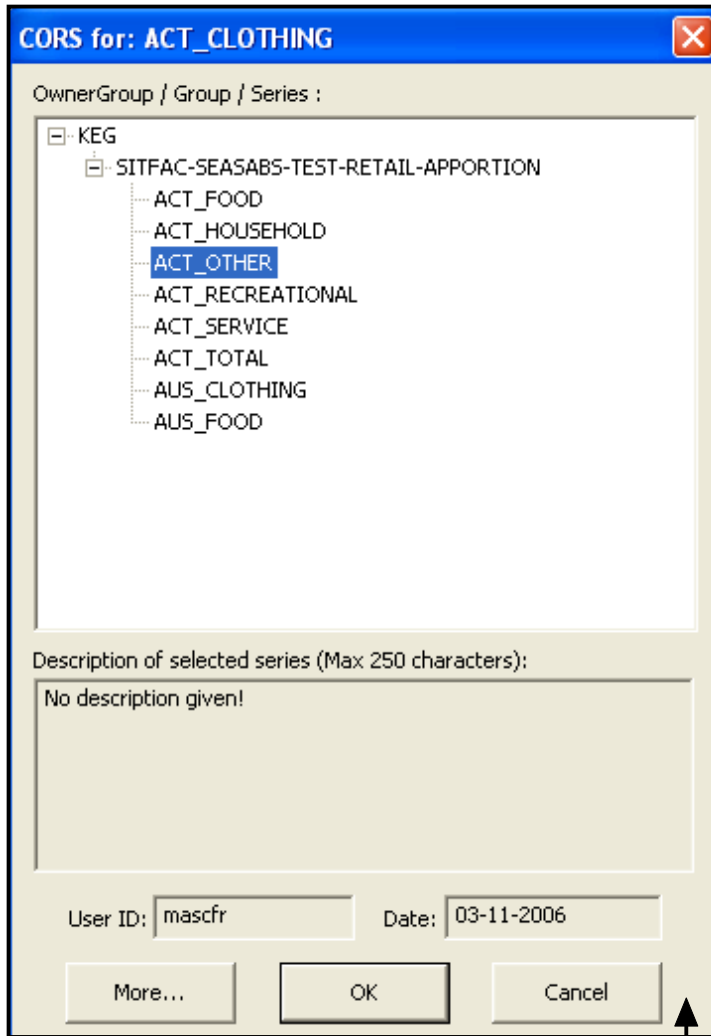
(1). We would like to generate consistency diagnostics for a list of related series pairs:

(a) either related to the single series open in a SEASABS session;

(b) for all series in the open group that are related to any other series across all other SEASABS groups/owner groups;

(c) for all series in the open owner group that are related to any other series across all other SEASABS groups/owner groups.

We envisage the following GUI for generating consistency diagnostics. This should be accessed from the existing "Related Series..." window in SEASABS version 2.6, e.g:



Consistency Table...

We'll leave it to the developer's imagination and creativity to come up with an esthetic interface.

(2). Having picked the desired "**All related series:**" radio button option (first panel), the user can now fine tune the list of series pairs by only keeping those "**Subject to:**" either "**user-defined attributes**" or "**no attributes/constraints**". If "**user-defined attributes**" is selected, a window of all attributes with checkboxes should pop up (e.g. same attributes that were used to declare all related series from the phase 1 implementation):

Select desired series attributes:	
Grand total	<input checked="" type="checkbox"/>
Total industries/one state	<input type="checkbox"/>
Total states/one industry	<input type="checkbox"/>
General lower level	<input type="checkbox"/>
Chain Volume Measure (CVM)	<input type="checkbox"/>
Current Price Value (CPV)	<input type="checkbox"/>
Monthly	<input type="checkbox"/>
Quarterly	<input type="checkbox"/>
Conceptually related	<input type="checkbox"/>
Source series	<input type="checkbox"/>
Derived series	<input type="checkbox"/>


The user is free to select as many "attributes" as they wish. The initial list of series pairs will be filtered against these attributes. If a single series in the list (not a related pair) has any of the selected attributes set (i.e. as set when first declaring related series maps), then it, together with its related counterpart should be included in the 'final list'. This 'final list' of series pairs will be used for generating the consistency table. So in the above example, the checked box would mean that all series pairs where one of its members is a "Grand total" would be selected. If for instance "Grand total" and "Quarterly" were both selected above, then the final list will contain pairs where either member is a "Grand total" OR "Quarterly" series. Each series pair in the final list should be unique, i.e. no duplicates should exist prior to generating the consistency table.

If the user had selected "**no attributes/constraints**" in the second panel of the "Generate Consistency Table" window, then no filtering of the initial list of series pairs is needed.

(3). If the "**Configure Tests and Output Table...**" button is pressed, an interface of dialogue boxes should pop-up that allows a user to specify the following configuration parameters and output options:

- Compute and report consistency diagnostics in table? Yes / No; default = Yes. "Diagnostics" are defined in (4) below. Otherwise, only report SEASABS parameters and settings in the output HTML table (see 3 below).
- Report prior corrections in HTML table (namely TBs, SBs, LGEXs and AOs)? Yes / No; default = Yes.
- Time spans on which to apply all tests. Default = minimum common span of a pair.
- Time spans for temporal-movement plots; default = minimum common span of a pair.
- Critical P-values for tests 2,3,4,5,6 with global default = 5%.
- Outlier detection threshold (#sigma); default = 3 sigma.
- Movement difference threshold (test 5); default = 2 sigma.

- Max. identical movement tolerance; default = 98%

On this interface, there should also be an option to retrieve the options/settings that were set in a previous (historical) run of this tool (previously stored by pressing the "Store consistency options button"; see 11 below). We don't want to store the consistency measures themselves. We want to be able to re-compute them on the fly using any stored options (if they exist), the stored historical series knowledge and original series. We are currently not sure how to reference/label the historical runs. A version number of the consistency run and pointers to the original data and historical series knowledge (by end analysis-date) may be enough. We'll leave this to the programmer. This design is strongly recommended by Duncan McCaskill:  (Subject: SEASABS Consistency relationships, phase 2 - issues; Database: TA-Statistical IT Facilities WDB; Author: Duncan McCaskill; Created: 05/01/2007; Doc Ref: DMCL-6X69Z3).

(4). If the "**Generate**" button is pressed, a HTML table of SEASABS parameters, settings and diagnostics should be generated and spawned to the screen. Here's an example from our prototype tool:



results.html

The first series of a pair is in a yellow row and the corresponding second (related) series is in a subsequent blue row. "First" and "second" here is arbitrary. Note that if a series of a pair is repeated in any of the subsequent pairs of a list (e.g. if comparing an aggregate series against state components), then this series should only be shown once, i.e. in a yellow row in the output table. The series that are related (paired) to this single series should then be listed in subsequent blue rows. For example:



results_aggstruct.html

We request some modifications to this table which were not possible to program in our prototype tool. These are outlined in 5 and 6 below.

(5). The links labelled with "[Diagnostics tests failed: ...](#)" in the above HTML tables give a summary of the failed diagnostic tests. Clicking on these shows the diagnostics computed for the "ABS seasonally adjusted" component series only. We would like to additionally compute diagnostics for the "original" and "D11" seasonally adjusted components aswell. So in the end, there will be three separate links to the diagnostics for each series pair in this table. For example:
[Diagnostics tests failed for Originals: 1, 2, 4](#)
[Diagnostics tests failed for ABS-SA: 1, 3, 5, 6](#)
[Diagnostics tests failed for D11-SA: 1, 2](#)

(6). The summary table of "SEASABS parameters and settings" in the example HTML tables above is not complete. We would like to include the following additional parameters and values:

- adjustment type: "concurrent" or "forward factor";
- method: value can be either "direct", "indirect", or "not adjusted". Note: our prototype tool currently only reports "direct" or "indirect" (as available from TSA-download). "not adjusted" cases must be available from the series knowledge, right? If so, we want to be able to report "not adjusted" cases in

the method field too.


- Star value;
- I/C ratio;
- I/S ratio;
- If static trading day exists, the trading day weights (table C15 SEASABS outputs) for each series listed in the "Trading Day" row with an indication that these are static trading day weights;
- If moving trading day exists, the trading day weights (table C15 SEASABS outputs) for the last year available in each series listed in the "Trading Day" row. There should also be an indication that these are "year XXXX" moving trading day weights;

(7). It is requested that all prior corrections in the HTML table (namely trend breaks, seasonal breaks, large-extremes and additive outliers) be listed one per line and in time order (latest to earliest).

(8). Algorithms for computing all consistency diagnostics with plots have been implemented in a prototype tool written in "Perl" and "R":
<S:\consistency\consistency.pl>. Complete specifications for each algorithm (for each diagnostic test) will be provided later. These may need to be re-programmed - a difficult undertaking since most algorithms are coded as complicated "R" function libraries. Other suggestions include: making calls to the R-libraries directly, or, invoking "R" as an executable from SEASABS. The latter will be slow and inefficient. This aspect is still on the drawing board and should not prevent implementation of items 1 - 7 above. The priority is to get the HTML table of parameters/settings coded up. The ancillary diagnostics that appear under the links mentioned in 5 can wait.

(9). In one of the diagnostic output plots (namely the movement versus movement plots) that are generically named "mvt_vs_mvt_pair_#.png" by TSA's prototype tool (generated under TEST 4), we would like to add another feature. For any outliers in this plot (currently labelled by a "X" symbol; see examples linked in HTML table above), we would like the ability to click on an outlier and a pop-up window would appear showing the time point to which the time-matched movements apply.

(10). In all plots and diagnostic outputs where we refer to time points, we would like these written in "long format" as used in SEASABS plots, e.g: Jan 1994, Qtr1 1994 etc..

(11). The button labelled "**Store consistency options**" allows the user to store the options and settings defined in item 3 above. These, together with the historical series knowledge and original series can be used to recreate a HTML table on the fly. This obviates storing any consistency measures. This design is strongly recommended by Duncan McCaskill:  (Subject: SEASABS Consistency relationships, phase 2 - issues; Database: TA-Statistical IT Facilities WDB; Author: Duncan McCaskill; Created: 05/01/2007; Doc Ref: DMCL-6X69Z3).

(12). The "**Export HTML to DOS...**" button allows a user to save the HTML table of SEASABS parameters and settings to a filename and directory of one's choice. The plots buried within the diagnostic links outlined in 5 above should not be saved (since they will take up a lot of room and be rather cumbersome to

manage). Only the textual contents of these diagnostic links should be saved. The user of course should be able to copy and paste any diagnostic plot from the SEASABS interface (like now) on a case-by-case basis for insertion into a Notes document.

Acceptance
Criteria:

Works as specified subject to programming and infrastructure limitations