

StocksAndFisheriesKB - Bug #12261

Multiple time series in GRSF from RAM data

Jul 27, 2018 04:46 PM - Aureliano Gentile

Status:	Closed	Start date:	Jul 27, 2018
Priority:	High	Due date:	
Assignee:	Yannis Marketakis	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:	GRSF		
Description			
We are finding errors/unexplained behavior in several RAM records, in example here: http://data.d4science.org/ctlg/GRSF_Admin/151d2126-c13f-3599-8259-b46388bff388 http://data.d4science.org/ctlg/GRSF_Admin/b165b8a4-d259-3796-9922-813ad8ba5290			
Where there are multiple BdivBmsypref / UdivUmsypref / Tbest values for the same year, while apparently in the RAM db there is only one value per year.			
Where are such additional data coming from?			
Please see also screenshots.			
Related issues:			
Blocks StocksAndFisheriesKB - Bug #12906: Update "problematic" records in the...			Closed Nov 20, 2018

History

#1 - Jul 31, 2018 12:09 PM - Aureliano Gentile

See also the GRSF API http://62.217.127.124:8080/grsf-api/resources/stock_data?uuid=151d2126-c13f-3599-8259-b46388bff388 it has the same issue with multiple values for the same year.

#2 - Jul 31, 2018 04:05 PM - Yannis Marketakis

- File *Capture.jpg* added

I inspected this a little bit and found out that these values come from RAM database.

More specifically, if I filter for the time-series with the following information

- stockID = HERRNORSS
- tsid = BdivBmgttouse-dimensionless
- tsyear = 2014 and 2013

I see that there are several values for these reference years (from different assessments I guess)

assessid	stockid	stocklong	tsid	tsyear	tsvalue
WGWISE-HERRNORSS-1950-2014-ICESIMP2016	HERRNORSS	Norwegian Spri	BdivBmgttouse-dimensionless	2013	0.9452
WGWISE-HERRNORSS-1988-2013-ICESIMP2016	HERRNORSS	Norwegian Spri	BdivBmgttouse-dimensionless	2013	1.0012
WGWISE-HERRNORSS-1988-2015-ICESIMP2016	HERRNORSS	Norwegian Spri	BdivBmgttouse-dimensionless	2013	1
WGWISE-HERRNORSS-1950-2014-ICESIMP2016	HERRNORSS	Norwegian Spri	BdivBmgttouse-dimensionless	2014	0.8132
WGWISE-HERRNORSS-1988-2015-ICESIMP2016	HERRNORSS	Norwegian Spri	BdivBmgttouse-dimensionless	2014	0.891

#3 - Jul 31, 2018 04:18 PM - Aureliano Gentile

Thanks Yannis, I am referring to BdivBmsypref and UdivUmsypref (Biomass and Mortality) which even if they come from a different assessment, then we should have the assessment ID (assessid) which is not the case in the above reported error/unexpected behavior of the GRSF catalogue. If the harvest of data is correct, then I guess you concur thinking that we need to distinguish such time series from different assessments.

#4 - Sep 12, 2018 05:54 PM - Aureliano Gentile

From: Gentile, Aureliano (FIAS)

Sent: Friday, 07 September, 2018 9:23 PM

To: Nikos Minadakis minadakn@ics.forth.gr

Cc: Yannis Marketakis marketak@ics.forth.gr; Pasquale Pagano pasquale.pagano@isti.cnr.it; Francesco Mangiacrapa francesco.mangiacrapa@isti.cnr.it; Luca Frosini luca.frosini@isti.cnr.it; Νίκος Μηνυαδάκης minadakn@ics.forth.gr; Taconet, Marc (FIAS) Marc.Taconet@fao.org; Ellenbroek, Anton (FIAS) Anton.Ellenbroek@fao.org; Anton, Paula (FIAS) Paula.Anton@fao.org

Subject: Re: Issues with GRSF

Yes exactly, thanks for that Nikos, **the issue would be to display the ram assessment id next to each value so to distinguish the different time series.**

Thanks and wish you a nice weekend.

Aureliano

Da: Nikos Minadakis

Inviato: venerdì 7 settembre, 18:11

Oggetto: RE: Issues with GRSF

A: Gentile, Aureliano (FIAS)

Cc: Yannis Marketakis, Pasquale Pagano, Francesco Mangiacrapa, Luca Frosini, Νικός Μηνάδακης, Taconet, Marc (FIAS), Ellenbroek, Anton (FIAS), Anton, Paula (FIAS)

Dear Aureliano, I investigated on the timeseries. There is nothing wrong. There are many values for the same year because they are coming from different assessments as noted by the assessment id in ram. The raw data of them are like: WGWISE-HERRNORSS-1988-2013-ICESIMP2016 HERRNORSS Norwegian Spring Spawning Herring BdivBmgttouse 2013.0 1.0012 WGWISE-HERRNORSS-1988-2015-ICESIMP2016 HERRNORSS Norwegian Spring Spawning Herring BdivBmgttouse 2013.0 1.0 WGWISE-HERRNORSS-1950-2014-ICESIMP2016 HERRNORSS Norwegian Spring Spawning Herring BdivBmgttouse 2014.0 0.8132 WGWISE-HERRNORSS-1988-2015-ICESIMP2016 HERRNORSS Norwegian Spring Spawning Herring BdivBmgttouse 2014.0 0.891

This is an example according to the issue that you included in the ticket.

As you can see for the same stock and the same year there are many values because of a different assessment id.

This is why you see multiple values for the same stock data page.

Now if the issue is why the assessment id is not visible then this is something that we should check together with our cnr colleagues.

We can discuss further in monday. Best, Nikos

#5 - Sep 24, 2018 11:36 AM - Pasquale Pagano

- Assignee changed from Nikos Minadakis to Francesco Mangiacrapa

Please @francesco.mangiacrapa@isti.cnr.it clarify the issue about the visibility of the assessment id.

#6 - Sep 26, 2018 03:34 PM - Francesco Mangiacrapa

Sorry for my late reply.

Reading the above comments I understood that assessment ID should be contained in the timeseries, if this is correct... the field 'assessid' is not visible because it is missing in the resource attached to record (i.e. <https://goo.gl/ddwV7E>), so this is not a catalogue displaying problem.

I suppose that the source json of a record has also the field 'assessid'... so I think there was a problem publishing the resources by the publishing-service and @luca.frosini@isti.cnr.it should check on this. Otherwise, @minadakis@ics.forth.gr or @marketak@ics.forth.gr can clarify, how are 'assessid' published in the catalogue?

Anyway, it would be really useful to attach to the current ticket at least one source json related to above records.

#7 - Sep 28, 2018 12:42 PM - Yannis Marketakis

The assessment IDs are published for the legacy records. So, for the record

http://data.d4science.org/ctlg/GRSF_Admin/151d2126-c13f-3599-8259-b46388bff388, you should check it's source record (in particular http://data.d4science.org/ctlg/GRSF_Admin/9046a39b-7da8-39f5-a8ea-ef32c522f32e) and then check the timeseries (e.g. Abundance Level), where you'll see the assessment IDs.

#8 - Sep 29, 2018 05:06 PM - Pasquale Pagano

@marketak@ics.forth.gr, i am not sure to understand your explanation.

Are you saying that we should combine the information contained in two or more records and perform this operation at publication time?

If the answer is yes, i believe this is not doable and this is not in line with what we did until now. The content is prepared by the KB that has to remain the authoritative source of information.

If instead, I have not understood what you were proposing, please be so kind to add additional details.

#9 - Sep 30, 2018 12:03 PM - Aureliano Gentile

- File screenshot-bluebridge.d4science.org-2018.09.30-11-58-27.png added

I think Yannis wants to say that the AssessmentID(s) (and the reporting years) are available in the source legacy record

see https://bluebridge.d4science.org/group/grsf_admin/data-catalogue?path=/dataset/9046a39b-7da8-39f5-a8ea-ef32c522f32e and attached screenshot

while in the GRSF record this information is lost and only the reporting year is provided thus making different data for the same REFERENCE YEAR indistinguishable.

Hope this helps. Thanks.

Grid

Graph

Map

108 records

«

1

–

100

»

Filters

Q

Search data ...

Go »

reference_year	value	unit	data_owner	reporting_year_or_assessment_id
2014	0.88061...	UdivUmsypref	WGWIDE	WGWIDE-HERRNORSS-1988-2015-ICESIMP2016
2013	1.176	UdivUmsypref	WGWIDE	WGWIDE-HERRNORSS-1950-2014-ICESIMP2016
2013	1.10473...	UdivUmsypref	WGWIDE	WGWIDE-HERRNORSS-1988-2015-ICESIMP2016
2012	1.208	UdivUmsypref	WGWIDE	WGWIDE-HERRNORSS-1950-2014-ICESIMP2016
2012	1.166965	UdivUmsypref	WGWIDE	WGWIDE-HERRNORSS-1988-2015-ICESIMP2016
2012	1.152	UdivUmsypref	WGWIDE	WGWIDE-HERRNORSS-1988-2013-ICESIMP2016

#10 - Oct 01, 2018 09:14 AM - Yannis Marketakis

Thanks @aureliano.gentile@fao.org

This is exactly what I described. Assessment IDs are valuable for the legacy record (in this case for the record coming from RAM). When publishing in GRSF they are omitted. If we want to keep this type of information for GRSF records, I see two options:

1. We store in the same field with the reporting year (as we do with legacy records)
2. We store them in a new field dedicated for this purpose.

The first option does not require any changes from the catalog and the publisher, while the second one does need the creation of a new field.

#11 - Oct 01, 2018 11:30 AM - Aureliano Gentile

Either the two options are fine, indeed as separate field might be better but for now also in the same field is fine, also considering the original requirement, in fact the field is called "reporting year or assessment id". Whoever will use such data will easily understand that, for example, assessment id "WGWIDE-HERRNORSS-1950-2014-ICESIMP2016" refers to data for the reporting year = 2016.

In conclusion, and also to speed up this bug fix, in case of RAM, when assessment ID is available, this needs to be maintained in the GRSF record and not substituted with an unclear reporting year value. (Actually this rule was already working properly, it broke recently...)

Thanks a lot, please, if things are not clear enough do not hesitate to launch a quick call and we can try to better explain and understand each other through Skype.

#12 - Oct 04, 2018 02:52 PM - Francesco Mangiacrapa

Hi @marketak@ics.forth.gr,

In order to make available Assessment IDs also for GRSF records, I didn't understand if they are omitted in GRSF source json or there was a problem during the publication stage in the GRSF publishing service.

In the first case, you should add Assessment IDs and then republish the GRSF records, while the second one for us it would be useful to have a GRSF source json (e.g. the source json of http://data.d4science.org/ctlg/GRSF_Admin/151d2126-c13f-3599-8259-b46388bff388) to investigate and then fix the issue.

#13 - Oct 05, 2018 01:43 PM - Yannis Marketakis

@francesco.mangiacrapa@isti.cnr.it they were omitted during the publication of GRSF records (as they should), so you don't have to investigate something.

#14 - Oct 05, 2018 02:06 PM - Francesco Mangiacrapa

- Assignee changed from Francesco Mangiacrapa to Yannis Marketakis

Yannis Marketakis wrote:

@francesco.mangiacrapa@isti.cnr.it they were omitted during the publication of GRSF records (as they should), so you don't have to investigate something.

ok, thanks.

#15 - Nov 16, 2018 01:21 PM - Yannis Marketakis

- Status changed from New to In Progress

Today we have discussed with FAO, about how to proceed with this issue.

We ended up that the best way to resolve this is to include the assessment ID which can be found in the timeseries of legacy records in the timeseries of the GRSF record.

Practically, this means that in GRSF for all timeseries coming from RAM we will use their assessment ID, instead of the reporting year.

#16 - Nov 19, 2018 03:58 PM - Yannis Marketakis

- Status changed from In Progress to Feedback
- % Done changed from 0 to 90

The publisher client has been updated.
Whenever there are time-series coming from RAM data source, then instead of their reporting year we use their assessment ID (as it happens with the legacy records).

Once the catalogue and the services are up, I will update some indicative records (and then of course I will update all of them), just to make sure that everything is OK.

#17 - Nov 20, 2018 09:33 AM - Yannis Marketakis

- Blocks Bug #12906: Update "problematic" records in the GRSF admin catalog added

#18 - Nov 20, 2018 09:44 AM - Yannis Marketakis

The issue has been resolved.

As a proof of concept, I ve updated one GRSF record (http://data.d4science.org/ctlg/GRSF_Admin/151d2126-c13f-3599-8259-b46388bff388).

After resolving other issues, I will update all the GRSF records ([#12906](#))

#19 - Nov 20, 2018 09:44 AM - Yannis Marketakis

- Status changed from Feedback to Closed
- % Done changed from 90 to 100

Files			
GRSF_151d2126-c13f-3599-8259-b46388bff388.JPG	58.4 KB	Jul 27, 2018	Aureliano Gentile
RAM_151d2126-c13f-3599-8259-b46388bff388.JPG	136 KB	Jul 27, 2018	Aureliano Gentile
Capture.jpg	52.2 KB	Jul 31, 2018	Yannis Marketakis
screenshot-bluebridge.d4science.org-2018.09.30-11-58-27.png	83 KB	Sep 30, 2018	Aureliano Gentile