

StocksAndFisheriesKB - Bug #13301

Missing catches data in RAM time series

Feb 07, 2019 04:17 PM - Aureliano Gentile

Status:	Closed	Start date:	Feb 07, 2019
Priority:	Normal	Due date:	
Assignee:	Aureliano Gentile	% Done:	90%
Category:		Estimated time:	0.00 hour
Target version:	GRSF		
Description			
In example in this record Hake - Bay of Biscay and Iberian Basin http://data.d4science.org/ctlq/GRSF_Admin/1bbbe48c-3410-3e19-bcae-fe8b144e852c GRSF has 2 records (2012 and 2013) and RAM has 32. See also attached screenshot other examples are provided in this excel file provided by RAM https://support.d4science.org/attachments/download/3085/GRSF-public-site-time-series-checks-2019-02-06.xlsx			
Related issues:			
Related to StocksAndFisheriesKB - Task #16375: Bulk approval of RAM records -...			Closed Apr 11, 2064

History

#1 - Feb 11, 2019 03:30 PM - Yannis Marketakis

@aureliano.gentile@fao.org I do not understand the issue here.

In fact, I see that the GRSF record has 97 values for catches and the same number of values does the RAM record has.
Can you please elaborate on this?

#2 - Feb 15, 2019 04:41 PM - Aureliano Gentile

Thanks Yannis.

As you can see in the excel file attached in this ticket (i.e. cell D21): GRSF has 97 and RAM has 170 and GRSF has 4 out of 5 of the assessid's.

Specifically, regarding the error mentioned in D23 "GRSF has 2 records (2012 and 2013) and RAM has 32" of the following stock:

Merluccius merluccius Portuguese Waters - East (Division 27.9.a) Bay of Biscay - South (Division 27.8.c)
Short Name: Hake - Bay of Biscay and Iberian Basin
GRSF Semantic identifier: asfis:HKE+fao:27.8.C;fao:27.9.A
Record URL: http://data.d4science.org/ctlq/GRSF_Admin/1bbbe48c-3410-3e19-bcae-fe8b144e852c

for the assessment_id **WGBIE-HAKESOTH-1982-2014-ICESIMP2016** the GRSF contains only two values (2013, 2012) while the source DB RAM (not the legacy record but the source database) contains 32 values (from 1982 to 2013). I just double checked and I confirm, you can check on your own also by checking the source RAM CSV file at <https://goo.gl/DMXkzE>, go to the worksheet "timeseries", filter by WGBIE-HAKESOTH-1982-2014-ICESIMP2016 (assessid column) and by TCBEST-MT (tsid column) and you get 32 rows.

In conclusion it seems a data harvest problem where the legacy record is already missing some data.

#3 - Apr 03, 2019 02:50 PM - Yannis Marketakis

- Status changed from New to In Progress

- % Done changed from 0 to 20

I investigated the issue and noticed that the missing timeseries exist in the GRSF KB but not in the catalog.
This means that it is not a harvesting issue.
The issue is related to updates we have done for including the assessment ID ([#12261](#)).

Although the assessment ID has been added in the corresponding timeseries, it was not omitted in some places (during the publishing/update of records).

I modified the services and it now works as expected. In the sequel, I will update all the records (from RAM) and obviously the corresponding ones in GRSF (Admin and Public)

#4 - Apr 03, 2019 02:52 PM - Yannis Marketakis

- Related to Task #16375: Bulk approval of RAM records - type assessment unit - without similar records added

#5 - Apr 03, 2019 03:48 PM - Aureliano Gentile

great, thanks a lot, this is one of the more 'disturbing' bugs...

#6 - Apr 04, 2019 03:08 PM - Yannis Marketakis

- Status changed from In Progress to Feedback
- Assignee changed from Yannis Marketakis to Aureliano Gentile
- % Done changed from 20 to 100

All the records from RAM have been updated.

The issue seems to be fixed for the corresponding record (http://data.d4science.org/ctlq/GRSF_Admin/ecea9850-e0d1-3191-bfef-eec9b0de4f8f). Eventually, any other issues related to this should have been resolved.

@aureliano.gentile@fao.org you are more than welcome to check this out.

In the sequel, we'll proceed with the bulk approval of GRSF records coming from RAM ([#16375](#)) and the update of GRSF records ([#16441](#))

#7 - Apr 05, 2019 11:47 AM - Aureliano Gentile

Thanks Yannis,

I am not sure if the fix has been already propagated to GRSF (RAM) records or only to the legacy records.

I just checked my original example (quoted below) and still there are only 2 entries for "WGBIE-HAKESOTH-1982-2014-ICESIMP2016" (assess_id) while there should be 32.

(it is a catch time series for the record Short Name: Hake - Bay of Biscay and Iberian Basin Record URL:

http://data.d4science.org/ctlq/GRSF_Admin/1bbbe48c-3410-3e19-bcae-fe8b144e852c)

Aureliano Gentile wrote:

Thanks Yannis.

As you can see in the excel file attached in this ticket (i.e. cell D21): GRSF has 97 and RAM has 170 and GRSF has 4 out of 5 of the assessid's.

Specifically, regarding the error mentioned in D23 "GRSF has 2 records (2012 and 2013) and RAM has 32" of the following stock:

Merluccius merluccius Portuguese Waters - East (Division 27.9.a) Bay of Biscay - South (Division 27.8.c)

Short Name: Hake - Bay of Biscay and Iberian Basin

GRSF Semantic identifier: asfis:HKE+fao:27.8.C;fao:27.9.A

Record URL: http://data.d4science.org/ctlq/GRSF_Admin/1bbbe48c-3410-3e19-bcae-fe8b144e852c

for the assessment_id **WGBIE-HAKESOTH-1982-2014-ICESIMP2016** the GRSF contains only two values (2013, 2012) while the source DB RAM (not the legacy record but the source database) contains 32 values (from 1982 to 2013). I just double checked and I confirm, you can check on your own also by checking the source RAM CSV file at <https://goo.gl/DMXkzE>, go to the worksheet "timeseries", filter by WGBIE-HAKESOTH-1982-2014-ICESIMP2016 (assessid column) and by TCBEST-MT (tsid column) and you get 32 rows.

#8 - Apr 05, 2019 02:26 PM - Yannis Marketakis

That's true. Only the legacy records have been updated.

The GRSF ones will be updated after proceeding with the bulk approval of RAM records ([#16375](#))

#9 - Jun 02, 2022 02:50 PM - Aureliano Gentile

- Status changed from Feedback to Closed

Files

grsf_missing-catches-data_07feb2019.jpg	72.6 KB	Feb 07, 2019	Aureliano Gentile
---	---------	--------------	-------------------