JODI Gas Data Quality Assessment

Alex Blackburn
United Nations Statistics Division

Presentation Outline

1. The revised JODI gas questionnaire
2. Smiley faces
3. Preliminary data checks
4. Data quality assessment
5. Summary

The Revised JODI Gas Questionnaire

Revised JODI Gas Format

JOINT ORGANISATIONS DATA INITIATIVE GAS QUESTIONNAIRE

Country
Month
Year

	Natural Gas million m^3 (at 15°C, 760 mm hg)	Natural Gas Terajoules	*Natural Gas 1000 tonnes*
	A	*B*	*C*
Production			
Receipts from Other Sources			
Imports			
LNG			
Pipeline			
Exports			
LNG			
Pipeline			
Stock Change			
Gross Inland Deliveries (Calculated)	0	0	
Statistical Difference (Calculated)	*0*	*0*	
Gross Inland Deliveries (Observed)			
of which: Electricity and Heat Generation			
Closing stocks			

Mass to volume conversion factor of LNG (if you have a specific figure)

m^3/tonne	LNG
Conversion factor	

Smiley Faces

Smiley faces: Motivation

- A concrete method to objectively assess JODI participation
- A way to encourage countries to participate and/or improve their data
- Has been successful in improving the JODI oil data system

Smiley Faces

a) Sustainability

Rates number of submissions within six-month period. JODI organisations proposed a fixed date to determine this rating.

Example :
For January-June 2013, ratings are determined on 20th September 2013.

If data are submitted after this date, while not very "sustainable" it is possible to revise the smiley face afterwards, to encourage countries to submit historical data and improve the database.

Smiley Faces

a) Sustainability ratings

☺ - if there are 6 monthly submissions before the cut-off date

😐 - 4-5 submissions

☹ - less than 4 submissions

Smiley Faces

b) Timeliness

A rating that assesses if a country submits data on time. Data are expected by 25th of the following month (M-1). But in recognition of organisations' different data collection systems, M-2 data can be considered timely. *[IEA and EUROSTAT use M-1 in rating timeliness of their member's submissions.]*

Ratings:

☺ - if there are 6 M-1 or M-2 submissions

😐 - if there are 4-5 M-1 or M-2 submissions

☹ - if there are less than 4 M-1 or M-2 submissions

Smiley Faces

c) Completeness

A rating based on the number of **relevant** data submitted within the rating period. However, data in both volume and energy units will be included. All participants are encouraged to submit data in both units, otherwise completeness ratings will be low. **Currently, the JODI organisations have not yet agreed on which flows should be included in the assessment.**

Ratings:

☺ - if more than 90% of relevant data points are submitted

😐 - if between 60% to 90% of relevant data points are submitted

☹ - if less than 60% of data points are submitted

Note: Countries/economies should put 0 to flows that are really zero and put N.A. for data that are not available

Preliminary Data Checks

Preliminary Data Checks

a) Balance Check –calculated supply should not be too different from observed demand

This is the same as checking the size of the statistical difference.

	Natural Gas million m^3 (at 15°C, 760 mm hg)
	A
Production	6840
Receipts from other sources	
Imports	40
LNG	
Pipeline	40
Exports	
LNG	
Pipeline	
Stock Change	200
Gross Inland Deliveries (Calculated)	7080
Statistical Difference (Calculated)	1120
Gross Inland Deliveries (Observed)	5960
Percentage statistical difference	*15.8%*

Statistical Difference= Production + Other sources + Imports – Exports - Stock change- Gross Inland Deliveries (Observed)

Statistical discrepancy should be as small as possible.
(5%? 15%? It depends…)

Preliminary Data Checks

b) Stocks check

Reported Data	Country A	Country B
Closing level of stocks M-1 (March)	11550	8440
Closing level of stocks M-2 (February)	12562	7121
Stock change in M-1 (March)	-1007	1122
Validation checks		
Calculated stock change	-1012	**1319**
Difference	-5	**197**
Percentage difference (Difference/Reported stock change)	0.5%	**17.5%**

The reported stock **change** should be exactly equal to the difference between the current month's and the previous month's stock **levels**. While a small calculation/reporting error is tolerated, a large difference should be queried.

Preliminary Data Checks

c) *Calorific Value Checks*

	Natural Gas million m^3 (at 15°C, 760 mm hg)	Natural Gas Terajoules	GCV Check
	A	*B*	
Production	45	1700	38
Receipts from Other Sources	5	120	24
Imports	20	800	40
LNG	*0*		
Pipeline	*0*		
Exports	10	491	49
LNG			
Pipeline			
Stock Change			
Gross Inland Deliveries (Calculated)	55	2009	37
Statistical Difference (Calculated)	*7*	*691*	*99*
Gross Inland Deliveries (Observed)	**62**	**2700**	**44**
of which: Electricity and Heat Generation			
Closing stocks			

Calorific values are calculated from reported data in million cubic meters and TJ. If values are not within the 35 to 48 MJ/m^3 range, these values should be checked.

Preliminary Data Checks

d) Conversion Factor Check

	Natural Gas million m^3 (at 15°C, 760 mm hg)	Natural Gas Terajoules	*Natural Gas 1000 tonnes*
	A	*B*	*C*
Production			
Receipts from Other Sources			
Imports			
LNG			
Pipeline			
Exports			
LNG			
Pipeline			
Stock Change			
Gross Inland Deliveries (Calculated)	0	0	
Statistical Difference (Calculated)	*0*	*0*	
Gross Inland Deliveries (Observed)			
of which: Electricity and Heat Generation			
Closing stocks			

Mass to volume conversion factor of LNG (if you have a specific figure)

m^3/tonne	LNG
Conversion factor	

Based on reported total imports/exports, the conversion factor is calculated to check how many cubic meters of NG are there in 1 ton of LNG, if values are very different from 1360 m^3/ton, please check the data again

Preliminary Data Checks

b) Fuel Balance – Total = LNG (m^3) + Pipeline (m^3)

	Natural Gas million m^3 (at 15°C, 760 mm hg)	Natural Gas Terajoules	*Natural Gas 1000 tonnes*
	A	*B*	*C*
Production	5	160	
Receipts from Other Sources			
Imports	20	800	
LNG	*35*		*35*
Pipeline	*0*		
Exports			
LNG			
Pipeline			
Stock Change			
Gross Inland Deliveries (Calculated)	25	960	
Statistical Difference (Calculated)	*37*	*1740*	
Gross Inland Deliveries (Observed)	**62**	**2700**	
of which: Electricity and Heat Generation			
Closing stocks			

LNG + Pipeline = Imports

Sometimes caused by different units

Data Quality Assessment

Data Quality Assessment

a) *Various Quality Checking Procedures*

- Internal checks (similar to preliminary data checks)
- Checks with recent months
- Checking against established monthly data
- Checks with quarterly or annual data
- Checks against other external sources

Each method has its (dis)advantages!

Data Quality Assessment

a) Comparison with Quarterly or Annual Data

- Quarterly and annual data are believed to be more accurate than monthly data.
- This can be a good check if these data are available timely and are collected on the same basis as JODI gas data.
- However, the time lag normally associated with annual data makes this check too late. Seasonality in gas data (increased consumption during the winter) can often make this only valid as a rough "order of magnitude" check too

Data Quality Assessment

a) Comparison with Other Monthly Data

- For partners with parallel monthly data collection procedures, a comparison with these data is often the most logical solution. For example comparing JODI M-1 with the following month's M-2.
- Example: January JODI data submitted at end of February are compared with January M-2 data submitted at end of March.

A useful check, IF the M-2 data are good quality.

Data Quality Assessment

a) Comparison with Recent Months

- A time series check compares the latest submitted month against recent months
- Calculating growth rates month-to-month allows erratic data to be highlighted
- This is a good quality control check on production and demand data; less useful for trade and stock changes
- But even erratic production and demand data can be correct, due to production outages or unusual weather

Data Quality Assessment

a) Comparison with External Sources

- A comparison with external sources can sometimes be useful, BUT should be done with caution:
  - JODI data are supposed to be official, why would an external source be more accurate?
  - What is the methodology of the external source? Transparent?
  - Does the source have a particular motive to provide accurate data (e.g. a company with business interests)?
- This check can be useful to double check a suspicious number or check the order of magnitude, but should not be relied upon.

Data Quality Assessment

UNSD Checks

- Balance checks
- Time Series checks
- Order of magnitude checks with annual data
- External sources checked when available

These are the checks employed by UNSD.

(In the first instance countries are always asked about any data we have questions or doubts on.)

Summary

Summary

- To ensure comparability of data, all countries should use the latest JODI Gas questionnaire and definitions from the JODI Gas manual
- The ratings for sustainability, timeliness and completeness are agreed upon and will be implemented similarly across organisations
- For sustainability, all data for a six-month period should be submitted by the cut-off date
- All participants are requested to carry out preliminary data checks before submission of the data. UNSD also carries out such checks upon receipt of JODI data, and will use it as a basis for data quality assessment
- All countries/economies are requested to fill-out the real 0s and Not Available (N.A.)'s in the JODI formats

Thank you

BlackburnA@un.org Energy_Stat@un.org

For more information at

www.jodidata.org