# Methods of assessing JODI oil data quality


Presented by

**Dr. Hossein Hassani**

**OPEC Secretariat**

# Disclaimer

This Presentation is intended solely for the use of OPEC Member Country officials and may be legally privileged and/or confidential. Any unauthorized use, disclosure or copying of this Presentation or any parts of it or its attachment(s) by any unintended recipient is strictly prohibited. If you have received this Presentation in error, please immediately return or destroy it. The OPEC Secretariat does not warrant or assume any liability or responsibility for the accuracy, completeness, or usefulness of any information contained in this Presentation. Nothing in this Presentation shall be construed as interpreting or modifying any legal obligations under any agreement, treaty, law or other texts; or expressing any legal opinions or having probative legal value in any proceedings.

# Outline

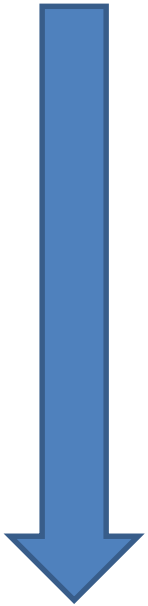
- **Balance Check**
  - Primary side vs. Secondary side
- **Internal consistency checks**
  - Primary side vs. Secondary Side
- **Other checks**
  - Comparison with other information
- **Meta Data**
- **Smiley faces**
- **Cost and burden**

# Balance check

## Table 1 Oil balance

### Primary oil

| | Crude Oil | NGL | … |
|---|---|---|---|
| | (1) | (2) | (3) |
| + Production | | | |
| + From Other sources | | | |
| + Imports | | | |
| - Exports | | | |
| + Products Transferred /Backflows | | | |
| - Direct Use | | | |
| - Stock Change | | | |
| - Statistical Difference | 0 | 0 | 0 |
| = Refinery Intake | | | |
| Closing stocks | | | |



### Secondary oil

| | LPG | Naphtha | … |
|---|---|---|---|
| | (5) | (6) | (7) |
| + Refinery Output | | | |
| + Receipts | | | |
| + Imports | | | |
| - Exports | | | |
| - Products Transferred | | | |
| + Interproduct Transfers | | | |
| - Stock Change | | | |
| - Statistical Difference | 0 | 0 | 0 |
| = Demand | | | |
| Closing stocks | | | |

# Balance check

## Primary oil

| | Crude Oil | NGL | … |
|---|---|---|---|
| | (1) | (2) | (3) |
| + Production | | | |
| + From Other sources | | | |
| + Imports | | | |
| - Exports | | | |
| + Products Transferred /Backflows | | | |
| - Direct Use | | | |
| - Stock Change | | | |
| - Statistical Difference | 0 | 0 | 0 |
| = Refinery Intake | | | |
| Closing stocks | | | |

## Secondary oil

| | LPG | Naphtha | … |
|---|---|---|---|
| | (5) | (6) | (7) |
| + Refinery Output | | | |
| + Receipts | | | |
| + Imports | | | |
| - Exports | | | |
| - Products Transferred | | | |
| + Interproduct Transfers | | | |
| - Stock Change | | | |
| - Statistical Difference | 0 | 0 | 0 |
| = Demand | | | |
| Closing stocks | | | |

# Balance check – primary oil

- **Calculated** refinery intake ≈ **reported** refinery intake
- Calculated refinery intake := production + from other sources + imports – exports + products transferred/backflows – direct use – stock change
- **Calculated** refinery intake - **reported** refinery intake = **Statistical Difference**
- Statistical difference should be small in relative terms (**less than 10% of Reported Intake??**)

# Balance check – primary oil

| | | Crude oil (tb/d) |
|---|---|---|
| + | Production | 3681 |
| + | From other sources | 0 |
| + | Imports | 2 |
| - | Exports | 0 |
| + | Products transferred/backflows | 0 |
| - | Direct use | 200 |
| - | Stock change | -295 |
| - | Reported refinery intake | 3550 |
| = | Statistical difference | 228 |
| % | Percentage statistical difference | 6.4% |

# Balance check – petroleum products

- **Calculated** demand ≈ **reported** demand
- Calculated demand := refinery output + receipts + imports – exports - products transferred + interproduct transfers – stock change
- **Calculated** demand - **reported** demand = **Statistical Difference**
- **Statistical difference** should be small in relative terms (**less than 10% of demand**)

# Balance check – petroleum products

| | | Total products ( tb/d ) |
|---|---|---|
| + | Refinery output | 126 |
| + | Receipts | 0 |
| + | Imports | 59 |
| - | Exports | 13 |
| - | Products transferred | 0 |
| + | Interproduct transfers | 0 |
| - | Stock change | -2 |
| - | Reported demand | 176 |
| = | Statistical difference | -2 |
| % | Percentage statistical difference | -1% |

# Balance check

- Applicable **only** if all data **are complete** and reliable
- **Large deviations** would require **review** and/or **verification/correction**
- **Applicable** to **every column**
- **Range** of **over 10%** is **quite large**

# Internal consistency checks

Quantities of **Refinery Intake** should correspond to quantities in **Gross Refinery output** (processing gains are the source of slight mismatch)

| | Crude Oil | NGL | … |
|---|---|---|---|
| | (1) | (2) | (3) |
| + Production | | | |
| + From Other sources | | | |
| + Imports | | | |
| - Exports | | | |
| + Products Transferred /Backflows | | | |
| - Direct Use | | | |
| - Stock Change | | | |
| - Statistical Difference | 0 | 0 | 0 |
| = Refinery Intake | | | |
| Closing stocks | | | |

| | LPG | Naphtha | … |
|---|---|---|---|
| | (5) | (6) | (7) |
| + Refinery Output | | | |
| + Receipts | | | |
| + Imports | | | |
| - Exports | | | |
| - Products Transferred | | | |
| + Interproduct Transfers | | | |
| - Stock Change | | | |
| - Statistical Difference | 0 | 0 | 0 |
| = Demand | | | |
| Closing stocks | | | |

# Internal consistency checks

- Automatic checks could be incorporated in the questionnaire to point out inconsistencies
- **Fuel checks** – **total oil products** should be equal to the **sum of reported products** (excluding memo items – automotive diesel and jet fuel)
- Statistician should ensure that this property holds in all columns
- Indication of misreporting of data

# Other checks

| | Crude Oil | NGL | … | Total | | Petroleum Products | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | LPG | Naphtha | … | Total Products |
| | (1) | (2) | (3) | (4) | | (5) | (6) | (7) | (13) |
| + Production | 0 | 0 | 0 | 0 | + Refinery Output | 0 | 0 | 0 | 0 |
| + From Other sources | | | 0 | 0 | + Receipts | 0 | 0 | 0 | 0 |
| + Imports | 0 | 0 | 0 | 0 | + Imports | 0 | 0 | 0 | 0 |
| - Exports | 0 | 0 | 0 | 0 | - Exports | 0 | 0 | 0 | 0 |
| + Products Transferred /Backflows | | | 0 | 0 | - Products Transferred | 0 | 0 | 0 | 0 |
| - Direct Use | 0 | 0 | 0 | 0 | + Interproduct Transfers | 0 | 0 | 0 | 0 |
| - Stock Change | 0 | 0 | 0 | 0 | - Stock Change | 0 | 0 | 0 | 0 |
| - Statistical Difference | 0 | 0 | 0 | 0 | - Statistical Difference | 0 | 0 | 0 | 0 |
| = Refinery Intake | 0 | 0 | 0 | 0 | = Demand | 0 | 0 | 0 | 0 |
| Closing stocks | 0 | 0 | 0 | 0 | Closing stocks | 0 | 0 | 0 | 0 |

| |
|---|
| Larger or equal to zero |
| Zero |
| Positive, neagtive or Zero |

# Other Tables

1- For **Crude** – the **Total** should be equal to the **Sum of all Streams** where applicable

2- For **NGLs** – the **Total** should be equal to the **Sum of all NGL** subcategories

- Field (lease) condensates
- Plant condensates
- Gas plant LPG
- Others

3- For **Other Primary oil**– the **Total** should be equal to the **Sum of all subcategories**

- CTL
- GTL
- Others

# Other checks

- Comparison with
  - data from other sources
  - last years' annual data
- Time-series check (seasonality, outliers, etc.)
- Visual checks
- Metadata

# Metadata

- The simplest definition of metadata is that it is **data** about **data**. More specifically **information** (data) about a **particular content** (data)
- **Metadata** describes how and when and by whom a particular set of data was collected; how the data is formatted
- **Metadata** must be updated when there is a change in resource it describes
- It can be useful to keep **metadata** even when the resource no longer exists
- **Metadata** enhances data transparency and is essential for understanding information stored in a database

# Smiley faces

- Timeliness
- Completeness
- Sustainability

# Smiley faces (timeliness)

- The **JODI database** is expected to be updated **regularly**.
- The **timeliness** indicates whether submissions were submitted at the **expected deadline**

"**good**" when **6 submissions** received within **45 days** after the end of the reference month

"**fair**" when **4 or 5 submissions** received

"**less reliable**" when **less** than **4 submissions** received

# Smiley faces (completeness)

**Completeness** measures the number of expected data points out of the **maximum 42** in the **JODI questionnaire** which are filled in

"**good**" when **more** than **90%** of the data are given for production, stock change/closing and demand

"**fair**" when **between 60%** and **90%** of the data are given

"**less reliable**" when **less** than **60%** of the data are given

# Smiley faces (sustainability)

**Sustainability** is the number of the monthly **JODI data** (timely) **submissions** evaluated **2 months** after the end of the **six-month period**

"**good**" if the **6 questionnaires** have been **submitted**

"**fair**" if **4 or 5 questionnaires** have been **submitted**

"**less reliable**" when **less** than **4** questionnaires have been **submitted**

# Cost and burden

- The **quality** of the **data** will be affected by available resources to collect, analyze and store energy statistics
- **Costs**: Specialized equipment, office space, utility bills, staff-hours involved, software, etc.
- **Response burden**: Simplest way to measure is the time spent by the respondent to provide information
- A compromise between **quality** and **cost** and **burden** must be achieved

# Cost and burden

- Functions of cost/burden
  - **Collection** of data
  - **Level** of disaggregation
  - **Time lags**, frequencies of data
  - Applied methodologies

# Thank you.

www.opec.org

# Metadata