



INTERNATIONAL
YEAR OF LIGHT
2015

**COSMIC
LIGHT** IAU

Globe at Night - Sky
Brightness Monitoring
Network

User Workshop Tokyo, Japan - Jan 7-9, 2015

Introduction to data analysis of Globe at Night - Sky Brightness Monitoring Network

Dr SO Chu-wing

The University of Hong Kong



国立天文台
NAOJ
National Astronomical
Observatory of Japan



Introduction to data analysis of Globe at Night Network

- NSB database
 - User login
 - Data structure
- Basic analysis
 - Statistics
 - Temporal variations of NSB
 - Short-term
 - Long-term
 - Geographical variations of NSB
 - Cross-city analysis

NSB database - User login

- All NSB data collected are obtainable from a database namely *iy1* via an user-friendly web-based interface *adminer*
 - IP address: (to be announced in the workshop/ later)
 - Login username: (to be announced in the workshop / later)
 - Password: (to be announced in the workshop / later)
- Database management system: MySQL
- Limited privileges



NSB database - Data structure

- The *iy1* database contains the following tables:
 - ***sqmle_nsb_data_table***: the main table storing NSB data collected from all stations in *real-time*
 - ***non_night_sky_events***: record of non-night-sky events
 - ***non_night_sky_events_category***: list of category of non-night-sky events
- Normal user can only edit the *non_night_sky_events* table

NSB database - User login

- Training dataset:
 - IP address: [http:// \(to be announced in the workshop\) /iyl-adminer.php](http://(to be announced in the workshop) /iyl-adminer.php)
 - Login username: (to be announced in the workshop)
 - Password: (to be announced in the workshop)

NSB database - Data structure

- The training database contains the following tables:
 - *sqmle_nsb_data_table_training*
 - *non_night_sky_events_training*
 - *non_night_sky_events_category_training*
- They have the same data structure as the real one
- The table *sqmle_nsb_data_table_training* stores more than 340,000 entries:
 - 1 urban (HK1) and 1 rural (HK2) stations in Hong Kong
 - Cover multiple lunar cycles and different weather conditions in late-2012 and recent months
 - The recent data collected from Tokyo (AOJ) and Taipei (TAM) are also included

NSB database - Data structure

- Table *sqmle_nsb_data_table* has the following fields for each entry:

Column	Type
id	bigint(20) unsigned <i>Auto Increment</i>
created	datetime
received_utc	datetime
received_adjusted	datetime
sqmle_serial_number	int(11) <i>NULL</i>
nsb	double [9999]
sensor_frequency	double [0]
sensor_period_count	double [0]
sensor_period_second	double [0]
temperature	double [9999]
time_of_resend	int(11) [0]
device_code	varchar(255) <i>NULL</i>
request_time	datetime [0000-00-00 00:00:00]
status	enum('success','failure') [success]

NSB database - Data structure

- Only the following fields are useful for basic analysis:

field	format	description
id	number	Unique ID of each entry
received_utc	YYYY-MM-DD HH:MM:SS	NSB data date & time in <u>UTC</u>
received_adjusted	YYYY-MM-DD HH:MM:SS	NSB data date & time in <u>local time</u>
nsb	number	Raw NSB value
device_code	character	Station code indicating the location
status	“success” or “failure”	Always select “success”

Basic analysis

- Before conducting the analysis, we assumed:
 - Data collected during non-routine lighting events or on-site maintenance / servicing periods (i.e. events with *reading_affected* = yes) were removed
 - Sunlight affected data were avoided if needed
- Will be considered:
 - Excluding unphysical NSB readings
- Not considered yet:
 - Moonlight contribution
 - Scattered city light from clouds
 - Aging of SQM-LE's filter
 - Light attenuation of glass window
 - Other sources of error

Basic analysis - Statistics

- *adminer* provides the following built-in functions to perform basic statistics:

function	target field(s)	description
avg	nsb	Calculate the average
count	id	Count the number of entry (sample size)
count distinct	(any)	Count the number of distinct entry (sample size)
max	nsb	Return the maximum
min	nsb	Return the minimum

Basic analysis - Statistics

- Example 1:
 - Count the sample size collected from AOJ between 2014-12-24 17:00:00 and 2014-12-25 05:00:00 local time

The screenshot shows a database query interface with the following components:

- 1. function:** A dropdown menu set to 'count' and a field containing '(id)'.
- 2. conditions:** A search table with the following entries:

device_code	=	AOJ
received_adjusted	>=	2014-12-24 17:00:00
received_adjusted	<=	2014-12-25 05:00:00
nsb	>	10
nsb	<	23
(anywhere)	=	
- 3. query:** A 'Select' button and a text area containing the SQL query:

```
SELECT COUNT('id') FROM `sqmle_nsb_data_table_training` WHERE `device_code` = 'AOJ' AND `received_adjusted` >= '2014-12-24 17:00:00 ' AND `received_adjusted` <= '2014-12-25 05:00:00 ' AND `nsb` > '10' AND `nsb` < '23' LIMIT 50 (0.186 s) Edit
```
- 4. result:** A box showing the result: **COUNT(id)** with the value 720. Below it, it says '(0 rows) whole result'.

Basic analysis - Statistics

- Example 2:
 - Calculate the **average of NSB** collected from HK1 between 2012-11-29 21:00:00 and 2012-11-30 22:00:00 local time

1. function

Select

avg (nsb)

Search

device_code	=	HK1
received_adjusted	>=	2012-11-29 21:00:00
received_adjusted	<=	2012-11-30 22:00:00
nsb	>	10
nsb	<	23
(anywhere)	=	

Sort

2. conditions

Limit: 50

Text length: 100

Action: Select

3. query

```
SELECT AVG(`nsb`) FROM `sqlme_nsb_data_table_training` WHERE `device_code` = 'HK1' AND `received_adjusted` >= '2012-11-29 21:00:00 ' AND `received_adjusted` <= '2012-11-30 22:00:00 ' AND `nsb` > '10' AND `nsb` < '23' LIMIT 50 (0.395 s) Edit
```

4. result

AVG(ns)

14.88297362110313

(0 rows) whole result

Basic analysis – Short-term temporal variations of NSB

- Aim: plot the NSB light curve of a particular night of a particular station to study the short-term time variation of NSB
- Example 3:
 - Time period (local time): 2014-12-20 15:00:00 - 2014-12-21 08:00:00
 - Location: AOJ

1. fields

2. conditions

3. sorting

4. query

5. results

Select

received_adjusted	
nsb	

Search

device_code	=	AOJ
received_adjusted	>=	2014-12-20 15:00:00
received_adjusted	<=	2014-12-21 08:00:00
nsb	<	23
(anywhere)	=	

Sort

received_adjusted	<input type="checkbox"/> descending
	<input type="checkbox"/> descending

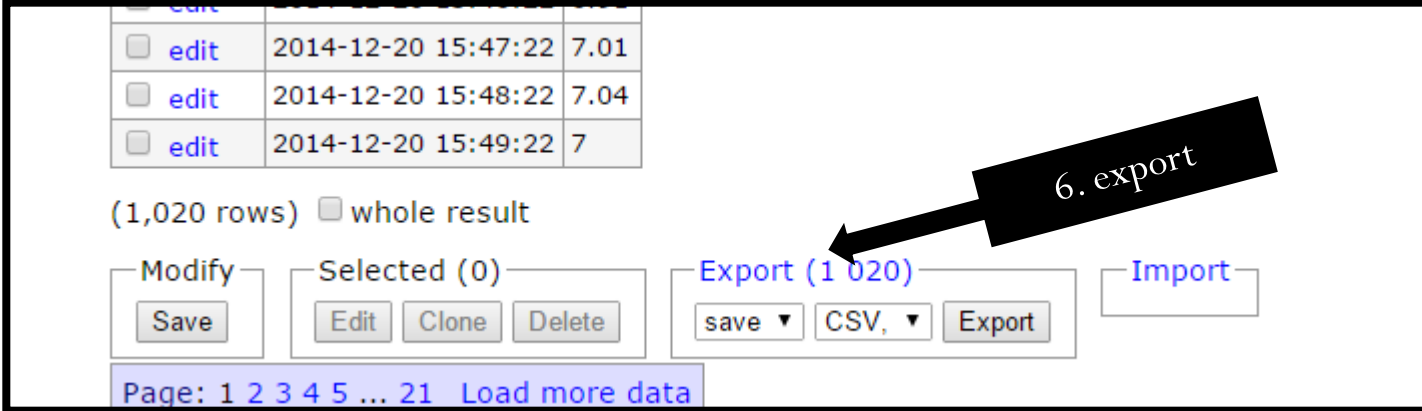
Limit: 50 Text length: 100 Action: Select

```
SELECT `received_adjusted`, `nsb` FROM `sqlme_nsb_data_table_training` WHERE `device_code` = 'AOJ' AND `received_adjusted` >= '2014-12-20 15:00:00 ' AND `received_adjusted` <= '2014-12-21 08:00:00 ' AND `nsb` < '23' ORDER BY `received_adjusted` LIMIT 50 (0.355 s) Edit
```

<input type="checkbox"/> Modify	received_adjusted	nsb
<input type="checkbox"/> edit	2014-12-20 15:00:22	0
<input type="checkbox"/> edit	2014-12-20 15:01:22	0
<input type="checkbox"/> edit	2014-12-20 15:02:22	0

Basic analysis – Short-term temporal variations of NSB

- Export data to csv (or other formats)



The screenshot shows a data table with three rows of data. Below the table are controls for row selection and actions. The 'Export (1 020)' button is highlighted with a black arrow and a label '6. export'.

<input type="checkbox"/>	edit	2014-12-20 15:47:22	7.01
<input type="checkbox"/>	edit	2014-12-20 15:48:22	7.04
<input type="checkbox"/>	edit	2014-12-20 15:49:22	7

(1,020 rows) whole result

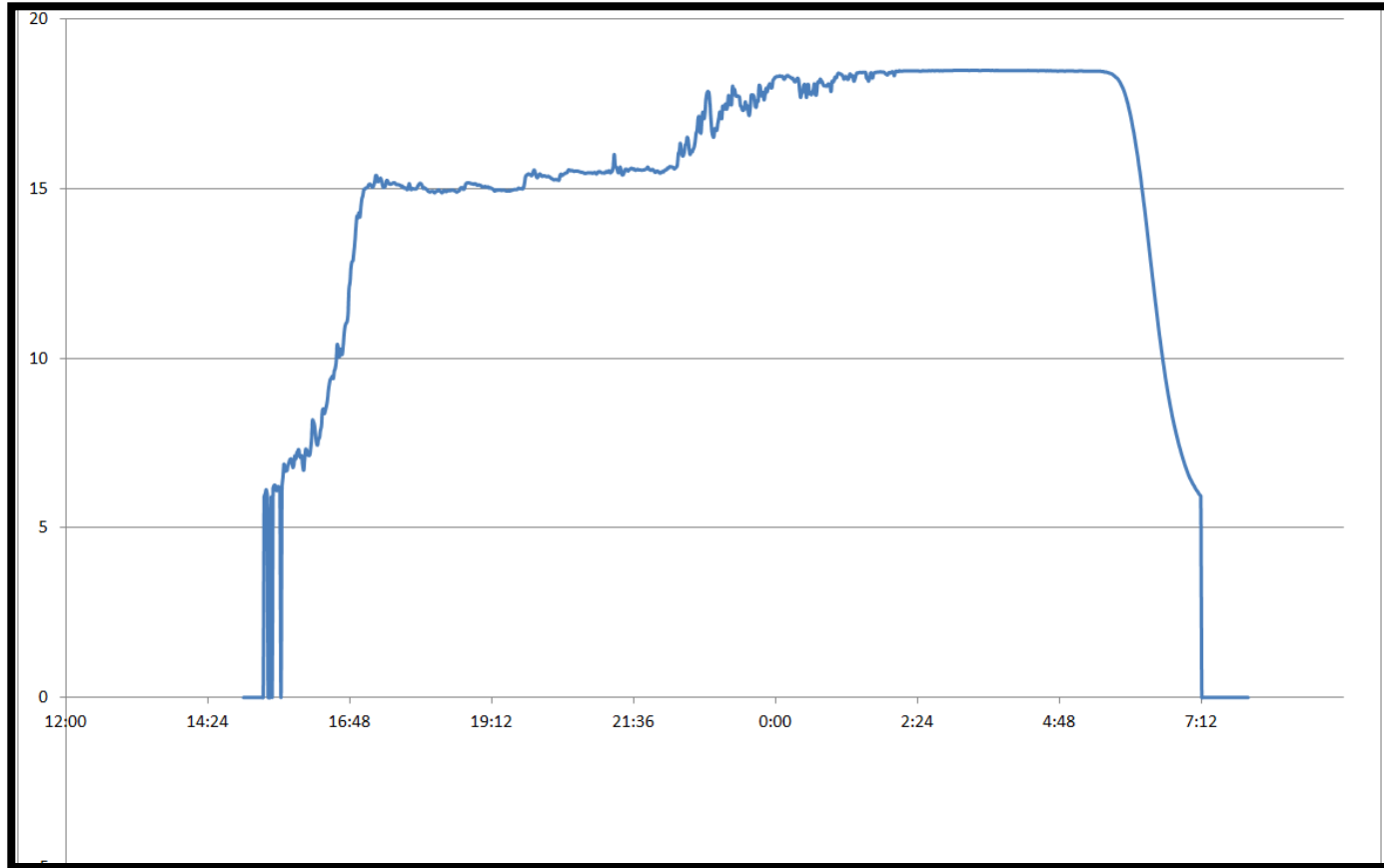
Modify Selected (0) Export (1 020) Import

Save Edit Clone Delete save ▼ CSV, ▼ Export

Page: 1 2 3 4 5 ... 21 Load more data

Basic analysis – Short-term temporal variations of NSB

- Plot in Excel (or other software)



Basic analysis – Long-term temporal variations of NSB

- Aim: plot the nightly average NSB of a particular station to study the long-term time variation of NSB
- Example 4:
 - Date period: from 2012-10-01 to 2012-10-31
 - Time window (local time): from 20:00:00 to 05:00:00
 - Location: HK2
- Calculate the nightly average NSB by **MySQL commands** from the following template (after setting the above parameters, execute the calculation by copying and pasting the codes in the *SQL command* box):

1. click this

DB:

SQL command Import
Dump Create table

```
select non_night_sky_events_cate  
select non_night_sky_events_train  
select sqmle_nsb_data_table_train
```

```
SELECT  
DATE_FORMAT(DATE_SUB(received_adjusted, interval 12 hour),"  
device_code as site, avg(nsb) as average_nsb, std(nsb) as s  
  
FROM sqmle_nsb_data_table  
  
WHERE  
nsb >10 and nsb < 23 AND  
device_code = 'HK2' AND  
DATE_FORMAT(DATE_SUB(received_adju  
(DATE_FORMAT(received_adjusted, "%H:%S") >= "20:00:00" or  
  
GROUP BY DATE_FORMAT(DATE_SUB(received_adjusted, interval 1  
ORDER BY received_adjusted;
```

2. paste here

3. execute

Stop on error Show only errors

Basic analysis – Long-term temporal variations of NSB

```
SELECT
```

```
DATE_FORMAT(DATE_SUB(received_adjusted, interval 12 hour), "%Y-%m-%d")  
as date,  
device_code as site, avg(nsb) as average_nsb, std(nsb) as std_nsb,  
count(id) as smaple_size
```

```
FROM sqmle_nsb_data_table
```

```
WHERE
```

```
nsb >10 and nsb < 23 AND  
device_code = 'HK2' AND
```

```
DATE_FORMAT(DATE_SUB(received_adjusted, interval 12 hour), "%Y-%m-%d")  
between "2012-10-01" and "2012-10-31" AND  
(DATE_FORMAT(received_adjusted, "%H:%i:%S") >= "20:00:00" or  
DATE_FORMAT(received_adjusted, "%H:%i:%S") <= "05:00:00")
```

```
GROUP BY DATE_FORMAT(DATE_SUB(received_adjusted, interval 12  
hour), "%Y-%m-%d")
```

```
ORDER BY received_adjusted;
```

Modify underline
parameters

Basic analysis – Long-term temporal variations of NSB

- Export data to csv (or other formats)

2012-10-25	HK2	18.610386178861788	0.9441538417451565	492
2012-10-26	HK2	17.815496957403628	1.002082216816823	493
2012-10-27	HK2	16.94540489642188	0.6210215567208585	531
2012-10-28	HK2	16.867415254237297	1.0692015936683192	472
2012-10-29	HK2	16.688022727272724	0.7412746696381254	440
2012-10-30	HK2	17.048747433264897	0.3584100640221845	487
2012-10-31	HK2	16.519117043121145	1.5258796367906504	487

4. results

31 rows (0.541 s) Edit, EXPLAIN, Export

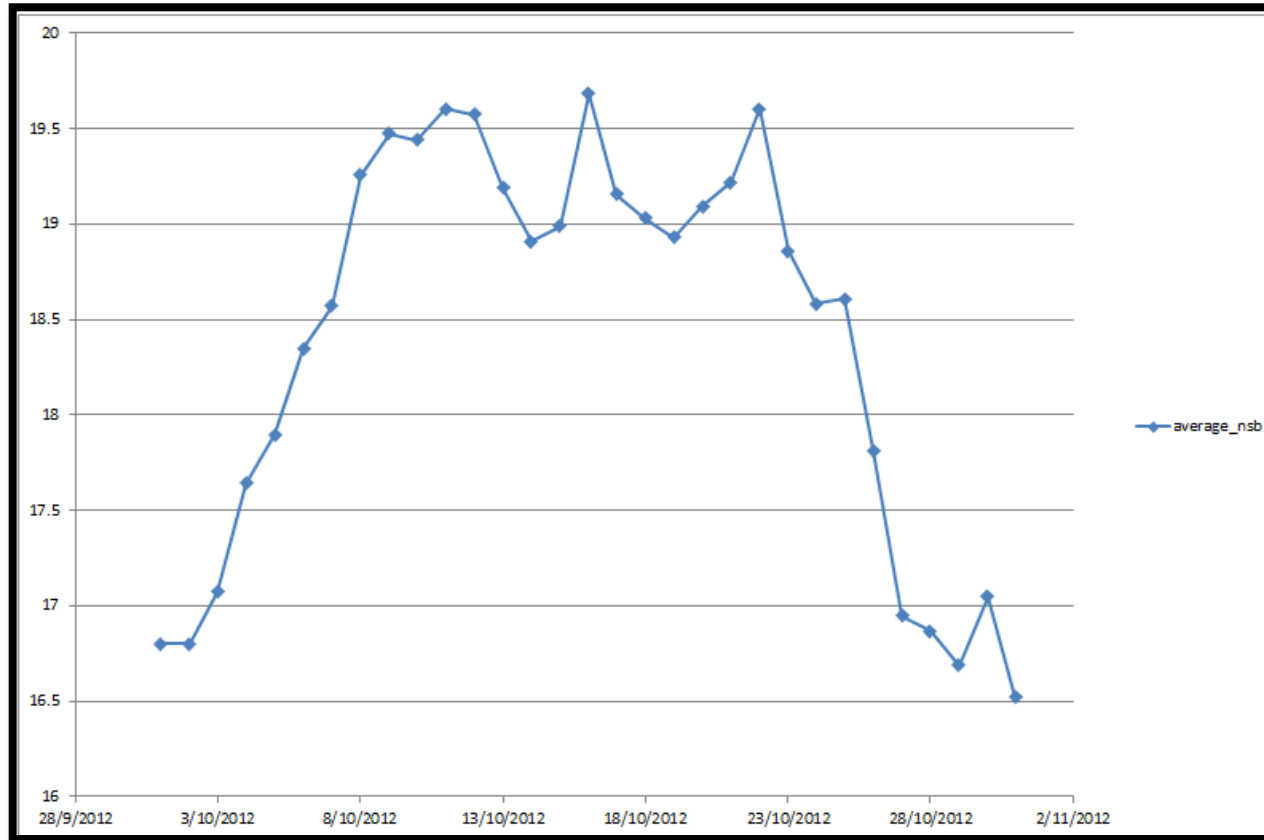
5. export

SELECT

```
DATE_FORMAT(DATE_SUB(received_adjusted, interval 12 hour), "%Y-%m-%d") as date,
device_code as site, avg(nsb) as average_nsb, std(nsb) as std_nsb, count(id) as :
```

Basic analysis – Long-term temporal variations of NSB

- Plot in Excel (or other software)



Basic analysis - Geographical variations of NSB - Cross-city analysis

- NSB would vary significantly among locations due to the difference in light pollution conditions
- Aim: study the geographic variations of NSB across different locations / cities from single-night light curves
 - Method: repeat the steps in example 3 for different stations then compare their light curves
- Cautions!
 - Effects of cloud / Moon would lead to completely different results
 - e.g., Station A (cloudy) vs station B (clear) => station A is brighter
 - => A is more light-polluted?
 - e.g., Station C (full Moon) vs station D (new Moon) => station C is brighter
 - => C is more light-polluted?