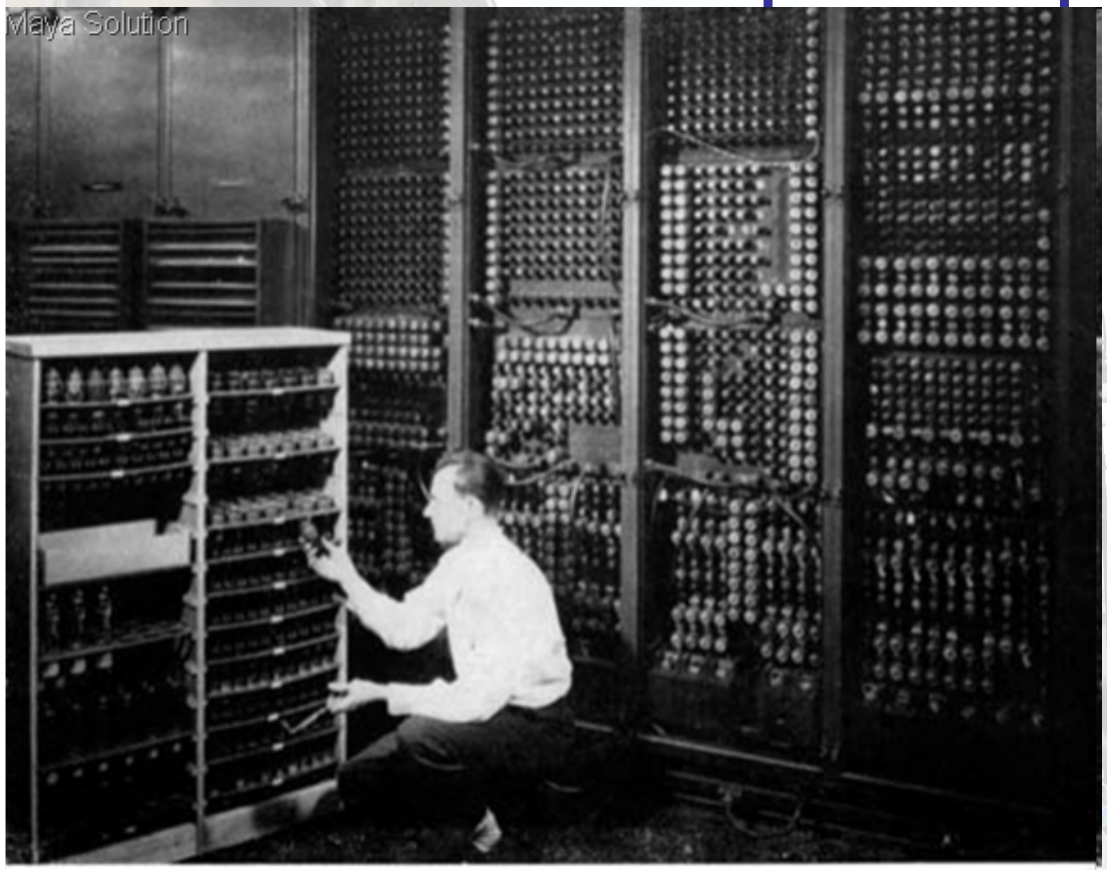
Environment Canada Environnement Canada

A database implementation of data analysis and quality control for the Brewer spectrophotometer



Replacing a bad tube meant checking among ENIAC's 19,000 possibilities.

Volodya Savastiouk

Environment Canada

March 23-26, 2010

WMO, Geneva

The database comment

- **We will face the need to track different changes in data (processing) in the future regardless of the ozone cross-sections**
- **Databases seem like a good solution**

The goals

- Advanced diagnostics in near real-time
- Data quality control
- Data re-analysis in a consistent way

Inspiration: show that this can be done

Web-based applications on your PC

- Open-source software allows running web-based applications on a single computer
- That same computer can be a server and database queries can be send via LAN or the Internet

Web-based applications on your PC

- Open-source software used:
  - Apache web-server
  - PHP interpreter
  - MySQL database engine
- The code is done in PHP
  - Scripting language allows to change the program easily when needed

Implementation

- PHP code fully separated from the data: it is abstract, i.e. the program doesn't know what it is reading. It gets the formatting information from the database.
  - Reads tag-based files (ds 0 ….)
  - The same code reads both existing records (DS, HG, SL,…) and those that have not been yet implemented
  - Database tells what records are known and how to read them
- Web-based interface for accessing data
- Database accumulates useful queries for future use

Brewer diagnostics

- Using the database approach we can prepare reports about the data that have been collected
  - Identify outliers easily
  - Keep track of SL and other tests
  - Use some intelligence in the diagnostics by analyzing several type of tests/observations together
    - What leads to the HG failure?
    - Did we forget to do HP before HG on a MKIII?

Environment Canada Environnement Canada

Warnings about the number of observations and tests

Measurements and tests summary.

_jday	_serial	aod	aode	ds	hg	hp ▲	ls	sl	sl	uv	zs
251	#		99	100	24	75			3	2	30
252	#	106		106	32	63			8	5	8
251	#		51	51	35	60			7	16	7
252	#	96		97	33	59			8	4	3
252	#		98	99	31	58			7	7	3
252	#		86	86	35	56			5	8	8
254	#		80	82	30	52			5	6	4
252	#		75	75	29	52			7	3	6
253	#	77		78	32	51			5	5	3
253	#	100		100	27	48			7	14	6
255	#		52	54	32	47		4	7	16	5

Fully configurable.
All criteria are in the database tables

Helps with scheduling issues for observations.

Environment Canada Environnement Canada

Questionable Run/Stop results. (0 sec)

Serial	Jday	slit 0	dark rs	dark counts	slit 1	slit 2	slit 3	slit 4	slit 5	slits 3+5
	2009256	0.9986	1.1667	112 / 96 (0.9M)	0.9984	0.999	0.9472	0.9983	0.8982	0.9994
	2009256	0.9333	1.2143	17 / 14 (1.4M)	0.8997	0.8999	0.9005	0.8992	0.8995	0.8996
	2009256	1.0001	1.1386	115 / 101 (0.9M)	1.0002	0.9996	0.9998	0.9991	0.9964	1.0005
	2009251	1.0031	0.9787	46 / 47 (0.6M)	0.9998	0.9993	0.9997	0.9999	0.9979	1.0006
	2009253	0.9998	3.4082	167 / 49 (1.6M)	0.9994	0.9998	0.9998	0.9978	0.9979	0.9998
	2009255	0.9958	0.8333	10 / 12 (0.6M)	0.9996	0.9994	0.9999	0.9984	0.998	1.0001
	2009257	0.9988	8.5882	146 / 17 (1.6M)	1.0006	0.9992	0.9988	0.9982	0.9984	1.0009
	2009256	0.9983	0.4	2 / 5 (1.5M)	0.9991	0.9996	0.9996	0.9996	0.9985	1.0012
	2009254	0.9969	1	30 / 30 (0.6M)	1.0004	0.9994	1.0002	0.9995	0.9986	0.999

Measurements and tests summary.

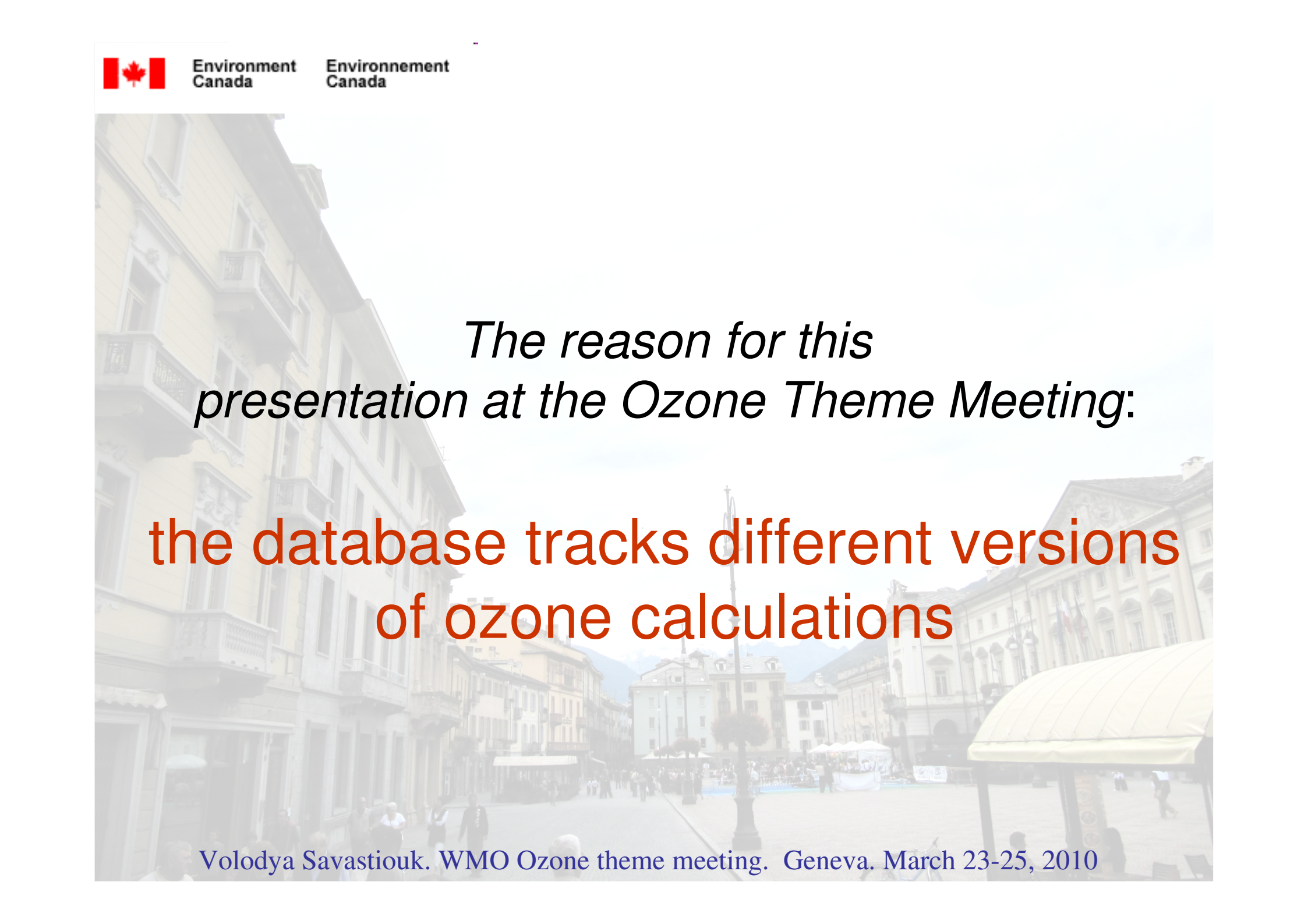
_jday	_serial	aod	aode	ds	hg	hp	ls	sl	sk	uv	zs
254	#			77	28				8	82	3
253	#			43	20				3	71	4
256	#		113	115	29				8	24	6
255	#			116	34				10	24	6
255	#			68	28	30			7	24	3
256	#			68	29	30			8	24	4

Standard lamp R6 (0 sec)

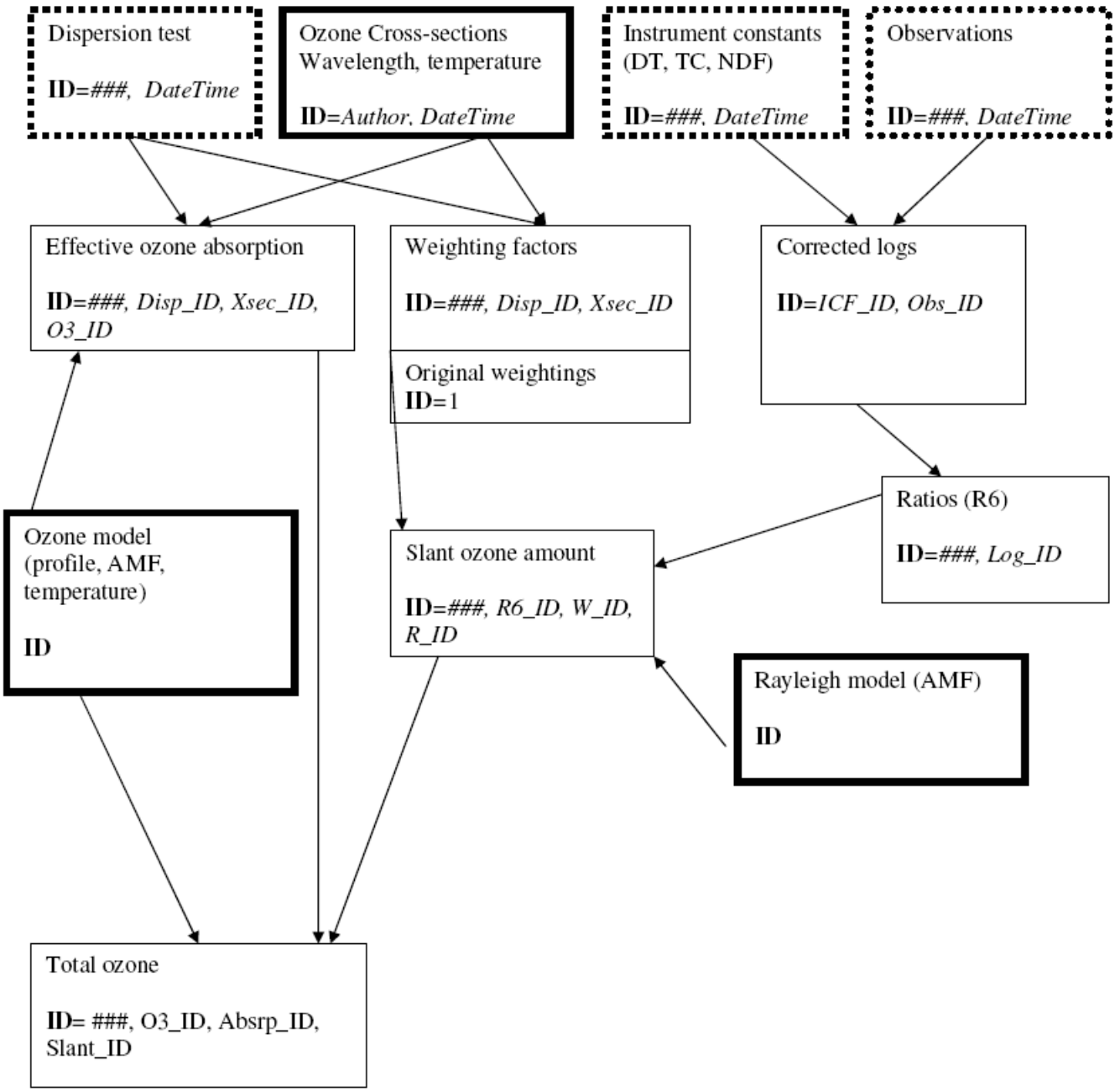
File name	Count	Serial	Expected (+/-20)	Measured	30-day history
b25009.	1	#	2110	2180	
b25709.	1	#	2110	2174	
b25709.	1	#	2230	2333	
b25409.	1	#	2000	1962	

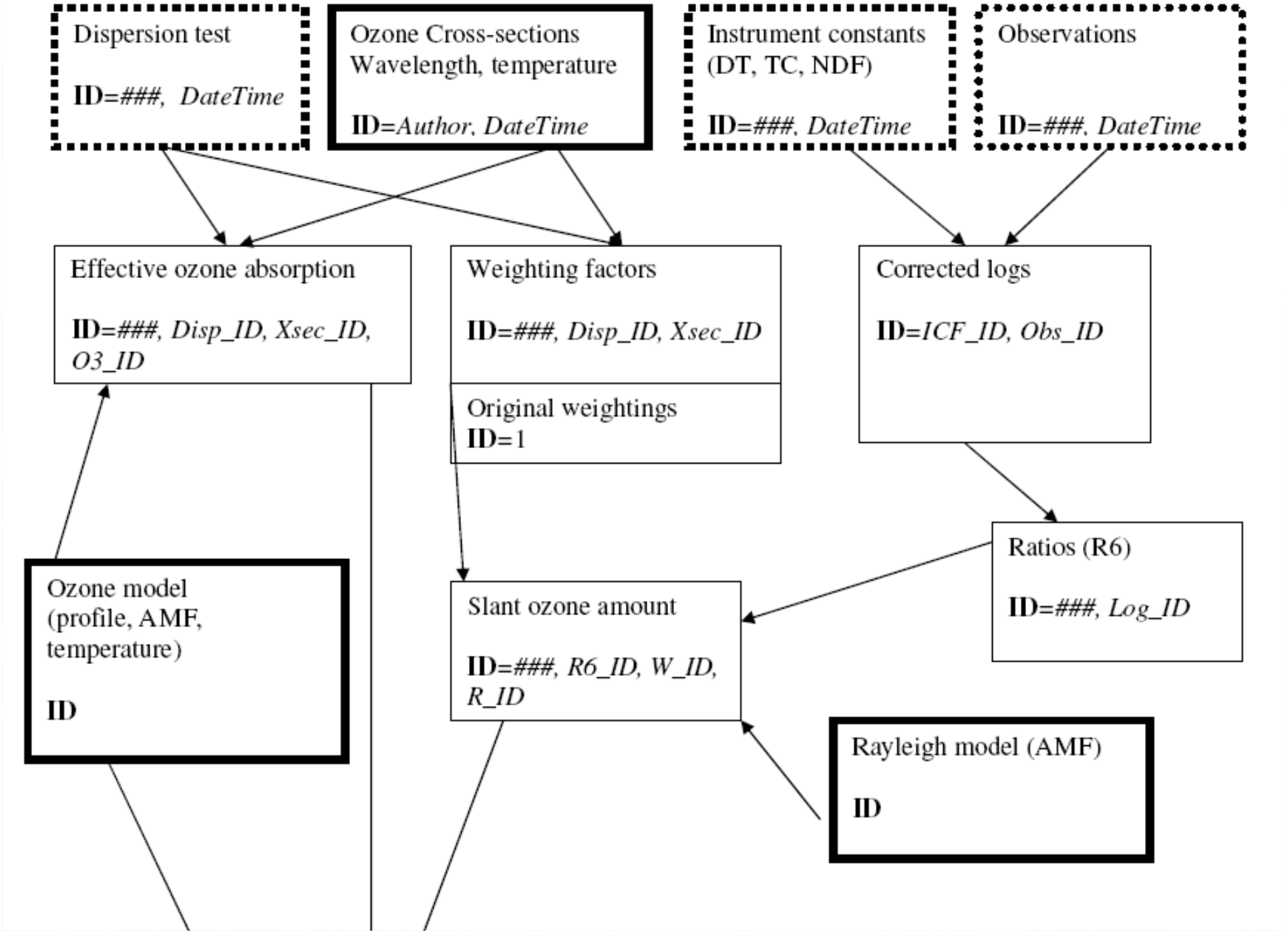
Ozone calculations

- Reprocessing starting as far back as possible
  - Dispersion tests
  - Calibration data (co-located Brewers)
- Tracking the data versions


Environment Canada
Environnement Canada
The reason for this
presentation at the Ozone Theme Meeting:
the database tracks different versions
of ozone calculations
Volodya Savastiouk. WMO Ozone theme meeting. Geneva. March 23-25, 2010

Environment Canada




Dispersion test
ID=###, DateTime
Ozone Cross-sections
Wavelength, temperature
ID=Author, DateTime
Instrument constants
(DT, TC, NDF)
ID=###, DateTime
Observations
ID=###, DateTime
Effective ozone absorption
ID=###, Disp_ID, Xsec_ID, O3_ID
Weighting factors
ID=###, Disp_ID, Xsec_ID
Original weightings
ID=1
Corrected logs
ID=ICF_ID, Obs_ID
Ozone model
(profile, AMF, temperature)
ID
Slant ozone amount
ID=###, R6_ID, W_ID, R_ID
Ratios (R6)
ID=###, Log_ID
Rayleigh model (AMF)
ID

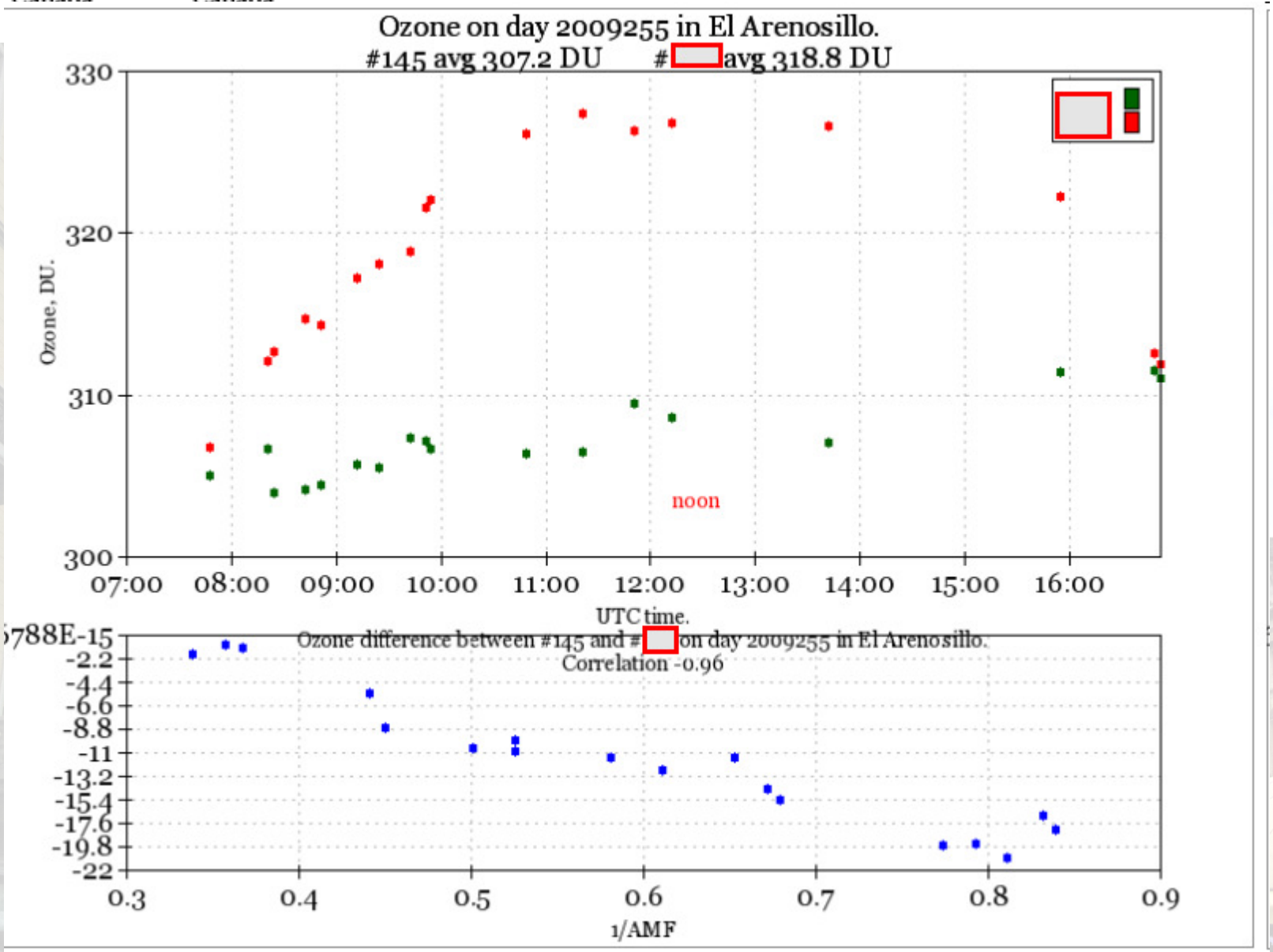
What has been done

- Calculation of ozone from raw counts - done
- Calculation of effective ozone absorption – nearly done

Environment Canada Environnement Canada

Ozone results for good DS obs (std<2.5DU, mu<3) within 3 minutes between 2 Brewers. (2 sec)

Location	Count	Reference ▲	Brewer	Day	Ozone average for B1	Ozone average for B2	% difference	ETC suspect?
El Ar	8	#145	#	2009255	309.9 +/- 1.5 DU	341.4 +/- 10.5 DU	10.2	
El Ar	16	#145	#	2009255	307.2 +/- 2.6 DU	278.7 +/- 4.9 DU	9.3	yes
El Ar	18	#145	#	2009255	307.2 +/- 2.3 DU	318.8 +/- 6.2 DU	3.8	yes
El Ar	16	#145	#	2009255	308.3 +/- 3.2 DU	314.4 +/- 6.3 DU	2.5	
El Ar	25	#145	#	2009255	307.7 +/- 2.8 DU	308.6 +/- 5.2 DU	1.5	yes
El Ar	20	#145	#	2009255	307.2 +/- 2.1 DU	303.7 +/- 2.1 DU	1.1	
El Ar	25	#145	#	2009255	308.0 +/- 2.6 DU	309.1 +/- 4.6 DU	0.9	
El Ar	17	#145	#	2009255	308.1 +/- 3.1 DU	306.0 +/- 3.0 DU	0.9	
El Ar	6	#145	#	2009255	310.7 +/- 2.3 DU	312.9 +/- 0.6 DU	0.8	
El Ar	22	#145	#	2009255	307.5 +/- 2.3 DU	305.5 +/- 1.8 DU	0.8	
El Ar	14	#145	#	2009255	307.6 +/- 2.6 DU	305.6 +/- 2.3 DU	0.7	
El Ar	17	#145	#	2009255	306.8 +/- 2.7 DU	305.6 +/- 3.3 DU	0.7	
El Ar	19	#145	#	2009255	307.7 +/- 2.2 DU	308.4 +/- 2.9 DU	0.6	yes
El Ar	16	#145	#	2009255	307.7 +/- 2.5 DU	308.6 +/- 2.9 DU	0.6	yes
El Ar	16	#145	#	2009255	306.7 +/- 1.9 DU	306.4 +/- 2.9 DU	0.5	
El Ar	9	#145	#	2009255	308.8 +/- 1.6 DU	308.9 +/- 2.7 DU	0.5	
El Ar	*29*	*#145*	*#145*	*2009255*	*307.9 +/- 2.7 DU*	*307.9 +/- 2.7 DU*	*0.0*	


Ozone on day 2009255 in El Arenosillo.
#145 avg 307.2 DU
avg 318.8 DU
330
320
310
300
Ozone, DU.
noon
07:00
08:00
09:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00
UTC time.
Ozone difference between #145 and
on day 2009255 in El Arenosillo.
Correlation -0.96
-2.2
-4.4
-6.6
-8.8
-11
-13.2
-15.4
-17.6
-19.8
-22
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1/AMF

Conclusions

- This database approach has been successfully implemented at the Canadian Brewer Network
- Valle d'Aosta Brewer station uses the database daily to monitor the instrument
- Next step is the implementation of the intercomparison (re)processing and integration with other database tools

I'd like to thank Tom McElroy, David Wardle, Ken Lamb and many others who provided invaluable input to this project.

Recommendations

- Invest into further development of this database system; this will help in the implementation of changes of any kind, not only the cross-sections

- The above only needs to be done once

- The above is applicable to all data, not only to the Brewer data

The database comment

- **We will face the need to track different changes in data (processing) in the future regardless of the ozone cross-sections**
- **Databases seem like a good solution**

Environment Canada Environnement Canada