Counts normalization and corrections and Poisson noise in and from CI, XL and UV files

Notes on the revised 02/14/07 version: Additional tests have been performed and changes were made in the document. We verified values of dark in XL, UX and CI routines again. Also we run no-light XL and UV scans to verify (a) if Hg or QTH lamps when turned on contribute to signal and (b) measured Dark. The conclusion is that some change in Table I were made (now they are in **bold green font**). For example, we were incorrect about the cy=4 multiplier of DARK in UV procedure. Also **equation 8** was corrected. We also added **Appendix VI** that contains figures form tests performed with no-light with UV and XL scans. Also briefly we mention the issue of spikes that, in our opinion should be further investigated.

Introduction

In this document we are concerned with CI, XL and UV files only. CI files give us ways to monitor instrument stability; XL files lead to instrument responsivity; UV files provide signal proportional to the total horizontal irradiance.

Our goal is to perform all calculations on properly normalized counts using dark correction and dead time correction. Also we want to estimate Poisson noise errors. For these tasks we need to know exactly how photon counts are measured and processed in Brewer routines and how they are reported in the files. The purpose of this document is an attempt to retrieve and reconstruct the information that will be used by us.

Unfortunately there are inconsistencies and idiosyncrasies in Brewer programs and in the files they generate. The manual is not too elaborative and sometimes ambiguous. We appreciate help of Volodya Savastiouk who answered several key questions (see Appendix V. Still more was there to be learned by inspecting Brewer routines and output files but we would be very grateful for further comments and corrections.

The following notation is used in the text:

cts	- counts
cpc	- counts per cycle
cps	- counts per second
pps	- photons (pulses) per second (these are <i>cps</i> corrected for dead time effect)
су	- number of cycles
div	- pre-counter division factor (<i>div=</i> 4)
dt	- dead time
Δt	- integration time ($\Delta t=0.2294s$)
Drk	- refers to dark <i>cts</i>

Tot- refers to total counts (dark cts not removed)Net- refers to Tot - Drk

 $\Delta S/S$ - relative 1-sigma precision

 ΔT - effective integration-averaging time ($\Delta T = cy^* \Delta t$)

<u>Parameters</u>

The results of our investigation are in the Table I. In Appendix I we inspected Brewer routines and performed analysis on some output data to arrive at values in the Table I.

Counts (*cts*) are usually normalized by cy, so cpc are available from files. But in CI files (see Table I) both *cts* and *pps* are listed. There is inconsistency in reporting *cpc* for *Drk* where it is multiplied by cy=4. This makes sense in CI file where because *Tot* is not normalized by cy=4.

Table I These parameters were deduced from CI, XL and UV routines and experiments (see Appendix I)

		CI	XL	UV (ux-scan)
1	DARK	<i>cy=40</i>	<i>cy=40</i>	cy=20
2	WV<3000Å	<i>cy=4</i>	<i>cy=30</i>	<i>cy=4</i>
3	WV>=3000Å	<i>cy=4</i>	<i>cy=20</i>	<i>cy=2</i>
4	Reported in file	<i>cy=4</i>	<i>cy=1</i>	cy=1
5	Reported DARK	cpc*4	срс	срс
6	Reported counts	cts & pps	срс	срс
7	Dark corrected	yes (pps)	no	no
8	DT corrected	yes	no	no

<u>Formulas</u>

The integrated counts (cts) obtained over several cycles are converted to cpc as follows:

$$cpc=cts/cy$$
 (1)

where cy (see Table I) is different for Drk and Tot and may depend on wavelength range.

Then

$$cps = cpc^*(div/\Delta t) \tag{2}$$

where div=4 restores original number of pulses prior to being divided before being counted and Δt normalizes cpc to 1 sec.

Due to the dead time count loss

$$cps = pps^* exp(-dt^* pps) \tag{3}$$

To obtain true number of pulses or photons per second (*pps*) the above formula needs to be inverted. Let's denote the inverse function (Dead Time Correction) as follows:

$$pps=DTC(dt,cps) \tag{4}$$

The inversion can be performed either with the method given in MAIN.ASC program in lines 8350-8370 or with a faster Newton method. Both methods were coded in Igor (see Appendix IV),

Dark is treated the same, i.e., dark counts are converted to cpc and then to cps.

In Brewer dark correction is performed simultaneously (if it is performed like in CI file) with the dead time correction:

$$Net_pps = DTC(dt, Tot_cps - Drk_cps)$$
⁽⁵⁾

However, one may argue that since dark *cts* contribute to dead time count loss, the following formula would be more correct even though, the dark pulses may have no Poisson distribution:

$$Net_pps = DTC(dt, Tot_cps) - DTC(dt, Drk_cps)$$
(6)

There will be no significant difference between results from formulas (5) and (6). When *Tot* is small, dead time correction is negligible and when *Tot* is large, *Drk* is negligible.

While the arrival rate of photons displays Poisson statistics, this is no longer true after the dead time loss. So neither *cts* nor the retrieved (via the *DTS* function) *pps* has Poisson distribution. But without much error we still may use Poisson formula to calculate standard deviation (see Appendix III). It is more accurate to calculate $\Delta S/S$ using *cps* counts than *pps* counts (See Appendix III).

Thus

$$\Delta S/S = sqrt(cps^*\Delta T)/(cps^*\Delta T) = 1/sqrt(cps^*\Delta T)$$
(7)

Since measurement of dark and total are statistically independent, the standard deviation of *Net_cps* is given as follows

$$Net_\Delta S/S = sqrt(Tot_cps/Tot_\Delta T + Drk_cps/Drk_\Delta T)/(Tot_cps - Drk_cps)$$
(8)

Example of calculations

Data from three files (different times and different instruments) were processed using the above formulas and parameters from Table I. The results are in Figure 1. The *Tot_cps* are in lower panel and in the upper panel $\Delta S/S$ and the fraction of dead time correction (pps-cps)/pps are given. Note that for XL and UV scan dead time correction is significantly lower than 1% and that $\Delta S/S$ is large for short wavelengths in UV scan.



Figure 1. Example of calculations of *cps*, (pps-cps)/pps and $\Delta S/S$

Appendix I: Routines evaluation and experimental verification of parameters and formulas

Brewer routines

INIT.RTN 12100 IT=.1147:B1%=1:B2%=2:B3%=64:B4%=128:B5%=4:B6%=16

Observation: Integration time IT=.1147, however real integration time is 0.1147*2=0.2294. To obtain photons per second, counts must be multiplied by 2/0.1147=4/0.2294, where the factor *div*=4 is the divider of pulses prior to the counter.

CI.RTN

14016 IF TYP="mkiv" AND MDD="o3" THEN CX="4": MN=2865: MX=3630: SWV=3250 \leftarrow CI file happens to have only one *cy* value.

14040 PRINT#8,"CI SCAN-INST #"+NO\$+" dt ";T1;CR\$;" cy ";CX\$;CR\$; \leftarrow CI file happens to report true value of *cy*.

31010 O1\$="R,1,1,40;O":GOSUB 9450:DARK=VA/40*VAL(CX\$) 31020 PRINT#8,USING " DARK = ########## ";DARK ← printing not normalized DARK=VA/40*4.

31030 PRINT#4,USING " DARK = #####.### ";DARK

Observation: DARK is scaled by *cy*=4.

31160 VB=VA:CY=VAL(CX\$):F(1)=DARK:GOSUB 8350

8350 REM correct VA for dark/dead time
8355 VA=(VA-F(1))*2/CY/IT:IF VA>1E+07 THEN VA=1E+07 ← Wrong limits.
Should be VA>1/T1/e.
8360 IF VA<2 THEN VA=2 ← This ain't necessary. DTC works fine for 0<=VA<2.
8365 F1=VA:FOR J=0 TO 8:VA=F1*EXP(VA*T1):NEXT
8370 RETURN

Observation: The dead time algorithm fails if VA>1/T1/e.

Observation:VB are counts that were not divided by cy=4. VA=pps is obtained in 8350-8370 section by the equivalent of the following:

A = VB/(cy=4)	B=DARK/(cy=4)
$A = A^*(div = 4)$	$B = B^*(div=4)$
A=A/0.2294	<i>B=B/0.2294</i>

VA = DTC(dt, A-B)

XL.RTN

13060 PRINT#8,"dt ";T1;CR\$;"cy ";CX\$;CR\$;"le ";LE%;CR\$ ← misleading cy value in file (cy=1)

31010 O1\$="R,1,1,40;O":GOSUB 9450:GOSUB 9190: DARK = VA/40 31020 PRINT #8, "DARK": PRINT#8, DARK: PRINT "dark=", DARK: SN\$ = ""

Observation: Unlike in CI file the DARK in XL file is not multiplied by *cy* of Tot scan *cy*.

31120 IF WV<3000 THEN CX\$="30" ELSE CX\$="20"

Observation: Two values of cy (30 or 20) are used. The scanning rates in Table II seem to confirm it.

31160VA = INT(VA*100/VAL(CX\$))/100'normalize output31170PRINT#8,TM;CR\$;WV;CR\$;SN\$;CR\$;VA;CR\$; 'record to file

Observation: *cpc* values are printed to file. To obtain *cps* or rather *pps* the following needs to be done:

 $A = VA^{*}(div=4)/0.2294$ $B = DARK^{*}(div=4)/0.2294$

VA = DTC(dt, A) - DTC(dt, B) (see eq.(6))

UV.RTN

16060 PRINT#8,"cy 1";CR(0); \leftarrow misleading *cy* value in file

32010 IF WV<3000 OR C\$="uf" OR C\$="uv" OR C\$="ua" THEN CX\$="4" ELSE CX\$="2"

33010 O1\$="R,1,1,20:O":GOSUB 9450:GOSUB 9190:DARK=VA/20:IF C\$="uv" THEN DARK=DARK*VAL(CX\$) ← what is the value of CX\$ at this point?

Is DARK=DARK*2 or DARK=DARK*4? Only one DARK value is printed in the file. By adding extra print statement we verified that is not scaled by 4 because C\$="ux"!

33030 PRINT#8,"dark";CR\$;DARK:RETURN

32070 X=VAL(CX\$):IF C\$="uf" OR C\$="uv" THEN X=1 \leftarrow presumably not the case because C\$="ux"? Correct!

32080 VA=INT(VA*100/X)/100 'normalize output ← presumably X=4 or 2 By adding extra print statements we verified that X is 4 WV<3000 and 2 for WV>=3000!

32090 PRINT#8,TM;CR\$;WV;CR\$;STR\$(M1);CR\$;VA 'record to file

Additional verification of results

In Figure 2 we plotted time vs. wavelength and calculated slopes for three regions: 2865Å - 3000Å, 3000Å - 3250Å and 3250Å - 3630Å. The slopes in sec/0.5nm are equal to

$$slope = cy^*(\Delta t/dc) + T_{scn}$$

where dc is measurement duty cycle and T_{scn} is time to move grating 0.5nm. First it should be noted that T_{scn} in 3250Å - 3630Å region is different that for shorter wavelengths. Is it because it is a different diffraction order? But from the data in the first two regions we can verify if cy's from Table I are congruent with the slopes in Table II.



Figure 2. Scan rates from CI, XL and UV files

Table II	Scan rates calculated from Figure 1
----------	-------------------------------------

2865Å - 3000Å	0.0078885 min/Å	0.030575 min/Å	0.0079645 min/Å
	2.36655 sec/0.5nm	9.1725 sec/0.5nm	2.38935 sec/0.5nm
3000Å - 3250Å	0.0078885 min/Å	0.021933 min/Å	0.0063882 min/Å
	2.36655 sec/0.5nm	6.5799 sec/0.5nm	1.91646 sec/0.5nm
3250Å - 3630Å	0.0082992 min/Å	0.022279 min/Å	0.0065307 min/Å
	2.48976 sec/0.5nm	6.6837 sec/0.5nm	1.95921 sec/0.5nm

Using *cy*'s from Table I we obtain that $\Delta t/dc$ is between 0.25926s and 0.23644s and T_{scn} is between 1.3947s and 1.44357s for wavelength less than 3250Å.

Appendix II: Dead time issues

We are not certain how valid the formula

cps=pps*exp(-dt*pps)

is for Brewer systems. The formula applies to the so-called extended (paralyzable) measurement regimen in counting of Poisson process. While there is no question about Poisson nature of photon statistics, there are other regimens of measurements, such as nonextended (nonparalyzable) measurement that yields the following, commonly used, formula (also for Poisson statistics):

cps=pps/(1+dt*pps)

"In the case of extended dead time an event occurring during t belonging to a previous pulse, although it will be lost, still starts a new dead-time period, that is it extends the dead time. In the case of nonextended dead time an event occurring during the dead interval is lost and does not start a new dead period." (see http://www.atomki.hu/atomki/Electr/deadtime.htm)

"Counting systems are usually classified into two categories: non-paralysable (type I) or paralysable (type II). In a non-paralysable system, each recorded photon produces a deadtime of length t; if an arrival is recorded at t, then any arrival from t to t + t will not be recorded. In a paralysable system, each photon arrival, whether recorded or not, produces a deadtime of length t; if there is an arrival at t, then any arrival from t to t + twill not be recorded." (see Yu and Fessler, *Phys. Med. Biol.* 45 (2000) 2043-2056)

In fact the DT(dt, cps) function is linearization function. How well do the shape of this function?

The value of dt is measured in Brewer using two superposition method measurements (three measurements in each range - Manual p. 97). Such measurements do not

necessarily prove that it is or not a paralyzable case. The Brewer method *a fortiori* makes it so. I presume that when this type of correction (linearization) was postulated and applied, it had prior more extensive empirical basis and understanding of Brewer photon counting electronics.

In most cases dt is very small and for low cts the effect of dead time is negligible. This is in particularly true for measurement through the diffuser (UV and XL files) and even for CI files. The dead time correction is more critical in direct sun measurements such as PS or DS routines (not the scope of this document).

Figures 3 and 4 help to identify maximal counts when the dead time correction is necessary or not. For instance if pps < 100,000Hz even dt=6E-08s produces error less than 1% (Fig. 3). And when dt=3E-08s 1% is obtained if pps < 300,000Hz (Fig.4).

Figure 5 demonstrates that for not so low counts (pps < 300,000Hz) the knowledge of dead time is not that critical. Whether dt=70ns or dt=10ns the errors is less than 1% when dt=40ns was assumed in the retrieval process via the DTC function.



Figure 3. Effect of dead time: constant dead time curves.







Figure 5. Effect of dead time uncertainty: *cps* were generated using seven *dt* values and *pps* was retrieved with assumption that *dt*=4.0E-08.

Appendix III: Noise calculations issues

Due to dead time loss cts has no longer Poisson distribution. However when counts are relatively small comparing to 1/dt, the precision can be calculated in approximate way. Below we present four formulas:

A is accurate when counting is ideal, i.e. *dt*=0:

$$A = \frac{\Delta pps}{pps} = \frac{1}{\sqrt{pps \cdot \Delta T}}$$

B is not accurate because *cps* is not Poisson. But this is a formula that we decided to use (see eq. 7):

$$B = \frac{\Delta cps}{cps} = \frac{1}{\sqrt{cps \cdot \Delta T}}$$

C is accurate (see Yu and Fessler, *Phys. Med. Biol.* 45 (2000) 2043-2056) that takes into account no-Poisson nature of *cps*:

$$C = \frac{\Delta cps}{cps} = \sqrt{\frac{1}{\Delta T \cdot cps} - \frac{dt}{\Delta T} \left(2 - \frac{dt}{\Delta T}\right)} = \sqrt{B^2 - \frac{dt}{\Delta T} \left(2 - \frac{dt}{\Delta T}\right)}$$

D is accurate for linearized counts pps=DTS(dt,cps). The formula is derived from formula *C* and the fact that

$$\frac{\Delta DTC(dt,cps)}{DTC(dy,cps)} = \frac{\Delta cps}{cps} \cdot (1 - dt \cdot pps)^{-1}$$
$$D = \frac{\Delta DTC(dt,cps)}{DTC(dy,cps)} = \frac{\Delta cps}{cps} \cdot (1 - dt \cdot pps)^{-1} = \sqrt{B^2 - \frac{dt}{\Delta T} \left(2 - \frac{dt}{\Delta T}\right)} \cdot (1 - dt \cdot pps)^{-1}$$

Note that in Figure 6 *B* formula approximates *D* formula the best.



Figure 6. Comparison of four methods to calculate precision $\Delta S/S$.

Appendix IV: Igor function for dead time correction

//Comments by P.K. 01/10/07 // double prec calculations should be used // sflag value should be kept. If cps>C indicates that something very wrong occurred (spike for example) //Example1 // print/d correct dt(2.8E-08 , 5345678, 9, .00001, 0) // 6393690.87594203 //Example2 // print/d correct dt(2.8E-08 , 5345678, 9, .00001, 1) // 6393691.07036627 function DTC(dt, cps,Niter, Precision, Method) //dt-dead time, cps-counts per sec, Niter - number of iteration for Brewer method // Precision - precision for Newton method //Method =0 - Brewer method variable dt, cps, Niter, Precision, Method variable i, Xn, An, Yn, A, B, C, ee variable/g sflag //global flag: 0 - no solution returns original value, 1 - solution ee= 2.71828182845905 C=1/dt/ee if(cps>=C) //no solution

```
Xn=cps
        sflag=0
else
        sflag=1
        if(Method==0) // Brewer method used with Niter=8 or 9 in Brewer software
                Xn=cps
                i=0
                do
                         Xn=cps*exp(Xn*dt)
                                 //B=abs(Xn*exp(-Xn*dt)/cps-1)
                                 //print i, B
                i=i+1
                while(i<Niter)
        else
                         //Newton method (converges much faster than method used by Brewer)
                Xn=cps
                i=0
                do
                         A = exp(-dt*Xn)
                         Yn=cps-Xn*A
                         An=(dt*Xn-1)*A
                         B=abs(Xn*A/cps-1)
                         Xn=Xn-Yn/An
                                 //print i, B
                i=i+1
                while(B>Precision)
        endif
endif
return Xn
end
```

Appendix V: Previous correspondence

Volodya's comments in red!

```
12/19/06
Hi Volodya,
```

Patrick has forwarded your email to me about counts, etc. (attached below). I hope you won't mind me writing you with more questions as I think I am still in need of further clarification.

I tried to read some Brewer's basic routines (e.g., uv.rtn and main.asc) but I do not have enough confidence in my findings. For this reason we would like to have definitive statements from somebody who knows Brewer's software and firmware through and through.

(1) How to get counts (photons) per second?

You write that counts per second are given by the following formula:

CperS = N*4/2/cy/0.1147

I am not quite sure why there is that factor=2.

For example our UV-files from ux scans read the following:

Integration time is 0.2294 seconds per sample dt 3.6E-08 cy 1

Also, why our integration time is twice as large as that in your example?

The two points above have the same cause: 0.1147 seconds is the integration time for half-cycle (the actual 'unit' of time), so if you use 0.2294 you do not need to dived by 2.

Then the so-called raw counts (N in your notation) are integer when large or have one decimal after the period when small. Where does this decimal come from? Is it because they are divided by that factor 2? (If so, they are not really raw.)

No, they are not dived by *that* factor 2. However, I have to correct myself: the UV files indeed all counts are scaled to one cycle. The actual number of cycles is either 4 or 2, so if you want to know what the actual raw counts were you need to multiply the reported number by 2 (if the wavelength is greater than 3000A) or 4 (for wl<=3000A)

How to convert dark counts into counts per second?

The same way as any other counts.

And just to be sure, the dark counts were not subtracted, were they? NO

(2) <u>How to use dead time?</u>

Say that we arrive at proper CperS formula applicable to our data. Now, if I want to correct for dead time, do I use CperS?

Yes

Does the dead time dt that was derived, was derived for CperS?

Yes

In another words, in the formula

 $CperS = CperSo^*exp(-CperSo^*dt)$

both CperS and CperSo are counts per second? The formula must be solved for CperSo? Yes

In which files dead time was corrected and which it was not? To my knowledge, only QL files have fully corrected values, other files have uncorrected counts (save for a factor of 2 or 4 in UV). There is no loss of information in here.

(3) <u>Standard deviation</u>

When calculating relative standard deviation of signal from Poisson statistics besides CperS we need the effective integration time ΔT . Then,

$\Delta S/S = sqrt(\Delta T^* CperS)/(\Delta T^* CperS)$

Now the question is what is effective integration time? Is it cy*0.2294 or does the factor=2 come into it?

See above

(4) <u>Temperature correction</u>

Your manual mentions temperature corrections as well. When the corrections are applied? For which files they do and for which files the do not?

The temperature correction that is mentioned in the manual is only applicable to ozone measurements and cannot be applied to UV files.

I hope these are not too many questions at once.

No, and I hope my answers were not too short.

Sincerely,

Peter Kiedron

Cheers,

Volodya Savastiouk

Hi Patrick,

the counts in the UV and CI files and SL records in the Bfiles are raw counts.

In these cases the name is used properly: no correction is applied to the numbers received by the computer from the Brewer when they are recorded. As once was very appropriately said: the "raw counts cease to be raw as soon as anyone or anything corrects them".

The only exception is the dark counts: they are not raw counts, they are corrected for the number of cycles to be the same as the rest of the scan. This is why in the UV scans often you can see the dark counts as non-integers.

As to your second question, there is indeed a count divider in the Brewer and the counts that are reported by the Brewer to the computer are predevided by 4.

Just in case: do not forget that the number of cycles is recorded in the files for each scan and the number of actual measurements at each slit-mask position is twice that. A cycle consists of two measurements at each slit-mask position.

In other words, to get the number of photon counts per second (for Poisson statistics) you need to take the raw counts, subtract the dark counts,

multiply that by 4, divide by 2, divide by the number of cycles and divide by the integration time (0.1147 sec):

CperS = N*4/2/cy/0.1147

Hope this helps.

Merry Christmas and a very Happy New Year!

Volodya Savastiouk

Appendix VI: No external light tests

We performed several UX and XL scans with a covered diffuser, so there was no external light. In the case of UX scans we wanted to see what are the values of signal when scans are performed with Hg or QTH lamps turned on. We do that to warm up lamps before HG or CI routines to save time in routine schedules. For example we do bluxhg or b2uxci sequences in the schedules. The results of tests are in Figures 7 and 8.



Figure 7. Several UX scans with and without Hg and QTH lamps on.

We concluded that there is no measurable light leak from Hg or QTH lamps when performing UX scans. Dark signal is not increased. Also it appears that the number of spikes is not larger when Hg or QTH lamps are turned out. However more extensive studies of spikes are necessary to arrive at their statistics.

The same data were used to evaluate what is the dark *cts* when measured from UX scan no-light values. We compared them with the values that are printed in UV files. The

results were close in magnitude (See Figure 9) but they are poorly correlated when averages of scan values with spikes were used. The correlation was less than 0.02. But when the averages were obtained from scans with screened out spikes, then the correlation improved to over 0.5.



Figure 8. The same as Figure 7 but spikes over 1.4 cts were removed.



Figure 9. Dark values form UV files (x-axis) are compared with Dark values obtained as averages of no-light ux scan values (y axis)

We also looked at XL scans and confirmed that printed Dark values in XL file are close to those measured as the average of all scan values for external light condition.

Spikes:

We do not know the origin and nature of the spikes. Their frequency is relatively low (less than 2%), however their magnitude often is large (30cts). This may cause problems when doing short wavelength irradiance measurement.

We need to investigate why the average of 150 samples (each of which is either average of 4 or 2 measurements) produces less correlated DARK with DARK that is an average of 20 from UV file when spike are not removed than when spike are removed. Does it mean that spikes do not occur or are screened out from Brewer's DARK measurement?

Also we should compared spike activity in Brewers that would be operating side by side. Are spikes related to the external source like cosmic rays or rather to internal Brewer's spurious noise.