Profiling the Long-Exposure Performance of a Canon DSLR *by Craig Stark*

Note: Originally published on Cloudy Nights, 7/2/12. Minor edits in place.

Introduction

I've had something stuck in the back of my head for several years now and a brief interchange with CN's own Uncle Rod motivated me to finally get it un-stuck and to try to come to a resolution. The thing that was stuck concerns Canon DSLRs. We see the term "RAW" and we'd like to think that this is the pure, raw, unadulterated data. We at times know it's not. I've not followed the Nikon DSLRs for some time, but I do know that in the past at least, a ~~sharpening~~ blurring filter was passed over their images before being dumped into their raw NEF format (unless you did a trick with the power to the camera).

Canons didn't have this, but I've noticed something else when I've had them on the bench. Over the years, for various venues, I've solidly tested a 350XT, 40D, XSi, 50D, and a 5D Mark II. They all had similar odd issues in their dark frames. If you try to measure the dark current, you end up at your wits end as increasing the exposure duration would by and large either do nothing to the average image intensity or it would *decrease* the average image intensity. Now, longer exposure times decreasing the image intensity sure is an odd thing and sure points to internal processing of the data before it hits the CR2 file. In one test, I did demonstrate a change in contrast between 1s and 30s images, but I've never been satisfied that I've understood and really demonstrated what is going on.

One goal of this article is to explore this in much more detail. I should note at the outset that despite what one might think, I'm not an "anti-DSLR" guy. Sure, I don't typically shoot with them, but I'm in an odd situation of having a good dozen CCDs around here. DSLRs are amazing devices that do double-duty for us and give us a great way (and pioneered the way) to get big, wide fields. But, they're not perfect and they do have their limitations. Understanding these limitations, though, lets us understand how to get the most out of the tools we have. That's why I've been pushed to characterize the performance here (and to take over 11G of flats and darks!). This brings me to the second goal – to answer the question of what one gains and loses by changing the gain (ISO) setting of the camera.

I should also say at the outset here that I am testing my own Rebel XSi camera. It's not a current generation and it wasn't top-of-the-line when new. But, it's what I (and many others) have and I believe it's typical of what's going on. It's certainly possible that other Canons don't have these same issues. I invite you to use this article as a guide and test your own gear to see if the same holds true. It has for the others I've tested and there are actually good reasons why Canon would process the data when we remember that astrophotography is not their target market (and no, I don't think the "Da" cameras differ here – I think they've just changed the IR filter on these). The processing makes sense for normal use – just not for astrophotography.

Basic Transfer Functions (System Gain)

The System Gain represents a basic property of the camera and describes how many electrons it takes to make a unit change in the image. So, how many electrons does it take to go from an intensity value of say 1028 in your image to 1029. We can calculate the system gain by taking a series of pairs of flats and plotting the mean intensity as a function of the variance of the difference between them (divided by two - this difference bit helps accommodate any unevenness in the flat as it will be the same for both members of the pair, leaving just the noise component). Here, I used exposure durations of 1/5 – 1/4000 s running at f/4 (ISO 100) through f/16 (ISO 1600) to equate the intensity across ISOs.

The figure below shows the system gain for my XSi at each of its ISO values:
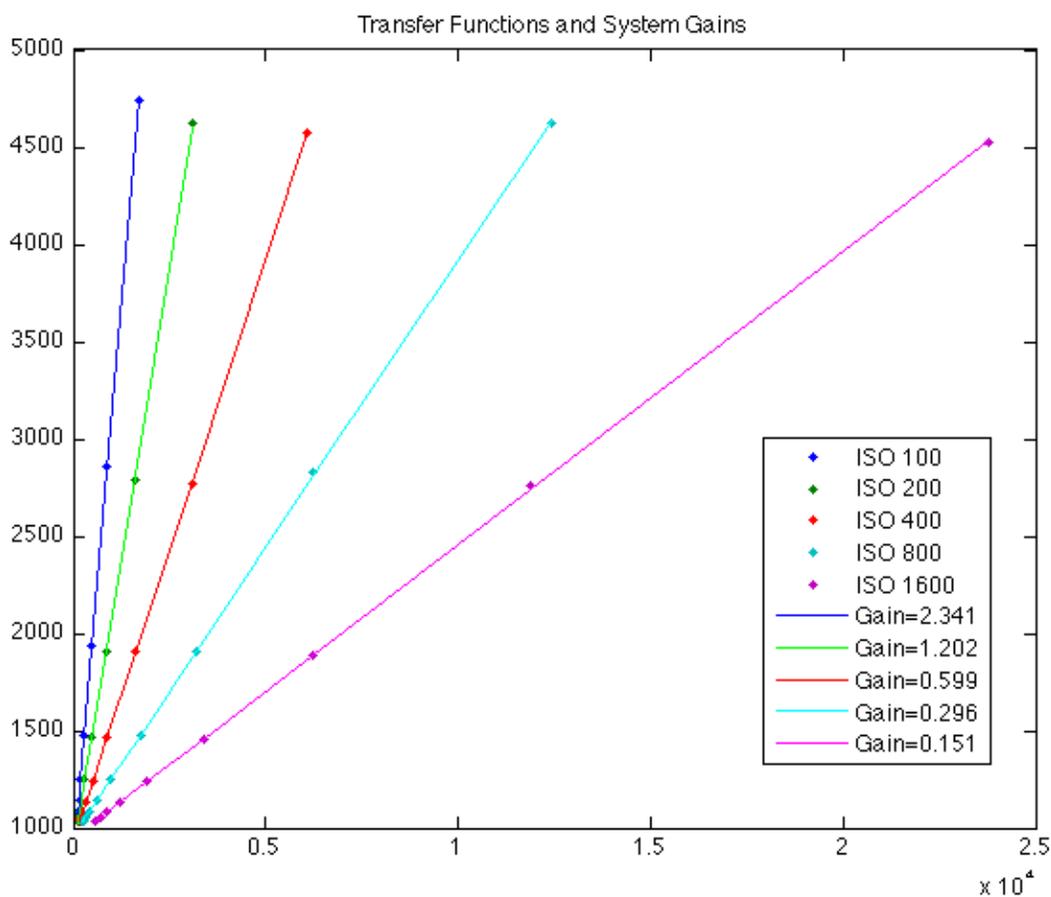

We can see a few things here. First, notice that each of the points nicely falls along a line. The sensor is behaving in a nice, linear fashion and everything looks clean in my data. We can see that at ISO 100, there are 2.3 e-/ADU (here, this is in terms of the 14-bit ADU and Nebulosity's 16-bit scaling is not enabled). As we increase the ISO, the system gain goes up in lock step just as it should with 1.202, 0.599, 0.296, and 0.151 e-/ADU for 200, 400, 800, and 1600 ISO respectively. Each gain is almost exactly the 2x difference we'd predict from the ISO (as ISO is just another term here for system gain).

One thing we can take from this is that there is little to be gained (at least from this perspective) by running at ISOs higher than 400 on this sensor. Photons and electrons are discrete (recording half a photon is like being half pregnant). At ISO 400, you've got about 0.6 e- for each ADU step. Or, put another way, you've got 1.67 ADU steps for every electron. Even running at ISO 200 where you've got 1.2 e-/ADU isn't losing much in the way of intensity resolution.

Dark Current and Distribution

As I mentioned at the outset, my prior tests of the Canon DSLRs have shown some very odd behavior in the dark current. Longer exposures could lead to a decrease or no change in the mean dark signal, even though the variance in the dark signal went up. To me, this clearly showed something was going on in the processing of our long exposures. The question was just, what?
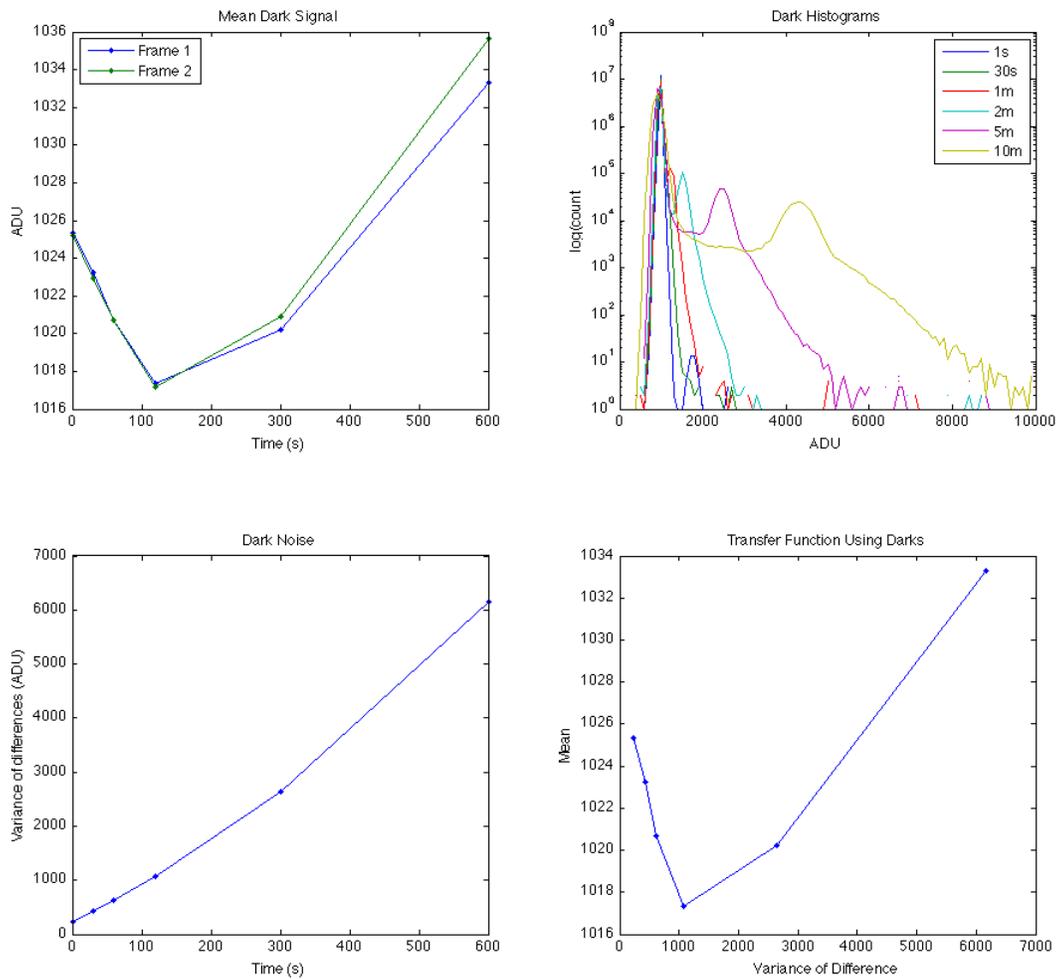
Here, I'm revisiting this issue in a bit more detail. To begin with, I took pairs of darks at 1s, 30s, 1m, 2m, 5m, and 10m at ISO 800 after the camera had warmed up (by taking 30 1m exposures). The upper-left plot shows the mean signal in each of the darks taken at each of the six durations. You can see this has a rather curious shape to it. This should be a straight line that increases the

mean intensity (ADU) with time. But, it's clearly not. Up until about two minutes, the mean intensity in the dark *decreases* with time (more dark current photons yielding less output signal), at which point it performs an about-face and increases with time. Clearly, something is happening to the signal.

To make it clear that strange things are indeed afoot and that it's not something of a feature rather than a bug, take a look at the upper-right image. This shows histograms of the darks for each of the six durations. Clearly, longer durations are leading to broader histograms. You can see a big warm-pixel bump move progressively to the right. Both of these are what we would expect and what any camera should do. However, you should be able to make out, at least for the 10 minute histogram, a progression not only of the right side of the histogram moving more right, but also of the left side moving more left. That should not happen. A nice background pixel that is very slowly building dark current would still move to the right (brighter) with time, rather than move to the left (dimmer) with time, were no processing going on. These histograms show there is more noise with time as we'd expect.

We can picture that perfectly by directly measuring the amount of dark noise. We do this just as we would measure read noise, by looking at the variance in the difference between frames (divided by two). The lower-left shows this and shows exactly what we would expect. The noise increases linearly with time (as these are Poisson processes, the variance should be proportional to the mean and the mean should also go up linearly). So, the noise behaves as it should (or at least close to it as the line is not quite linear, but it's close). The mean, however, isn't.

Finally, we can do one more thing here and that's to plot the transfer function or system gain using the darks. If all is operating in a predictable, linear fashion, the sensor shouldn't care at all whether the photons used to make the "flat field image" come from short exposures with light going through a lens or whether they come from longer exposures with the current coming from heat. Hot pixels would distort things a bit perhaps, but they should be roughly the same. The lower-right plot shows the transfer function using dark current. This looks nothing like the first figure. Yes, Bill, strange things are afoot at the Circle-K. There is some form of processing happening in long-exposure images, likely owing to the thermal signal.

Dark Stability

One of the clear things to grapple with in any off-the-shelf DSLR is the stability of the dark signal. These cameras aren't cooled and so they will have an appreciable dark signal and they warm up during use so that the dark signal will increase with time. How long does this continue and how much of a concern is it?

To investigate this, I ran the camera through a series of 60 2-minute darks indoors so that the ambient temperature would be relatively constant. Below, I'm showing what the dark signal looks like over two hours (ISO 800) both in terms of the mean signal and in terms of the noise (variance of the difference between adjacent darks, divided by two).
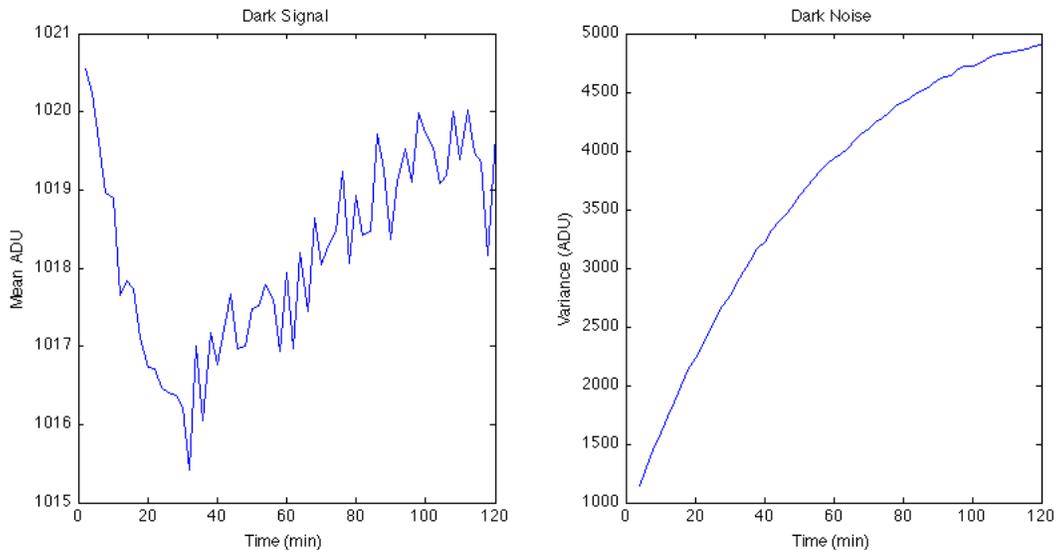
On the left, we see the same check-mark looking plot we got before when looking at mean dark signal vs. exposure duration. This again supports the notion not only that there is scaling going on, but that it is a function of the thermal signal in the image (and not the raw exposure duration). Here, the exposure duration is the same and the only thing that is happening is the camera is slowly warming up (from 21C to 34C according to the EXIF data in the CR2 file). Not to beat a dead horse, but as the camera warms up, if nothing is going on, the signal should never go down.

On the right, we have the dark noise over the interval. You can see that this increases substantially from just over 1000 ADU to almost 5000 ADU over the two hours. The chip has gone up 13C in this time and a general rule of thumb is that the dark current doubles roughly every 6C. As this variance should be proportional to the mean, a rise of 12C would have put the variance term up at about 4600. We're a touch warmer and up around 4900 – not bad for a rule of thumb. The noise has gone up then about what we would have expected given the temperature rise.

Looking at the noise plot, it seems to be the case that by two hours in, while not quite at equilibrium, we're close to it as the curve is getting flatter. If we extrapolate here, it may be about 3 hours before it's constant. One thing worth noting though is that you can see that about a half hour in, we begin a reasonably linear rise in the dark current. If we calculate the slope of that rise, we see that it's only 0.0016 – far off of the system gain. So, while the camera is behaving in a more predictable manner after this warm-up time, it's not that any image processing has been disabled. But, perhaps at this point, whatever the algorithm is, it's behaving a bit more predictably.

System Gain Versus Exposure Duration

In the first section, the system gain was computed using very short exposure durations that minimize any thermal effects. Looking at the darks in the last two sections showed some clearly odd behavior indicating processing of the images going on. What if we increase the exposure duration to allow thermal effects to take hold and re-measure the gain? Here, the same rig was used (decreasing the amount of light from the flat panel) to let me run system gain transfer

functions at 1/30s, 10s, and 2m.
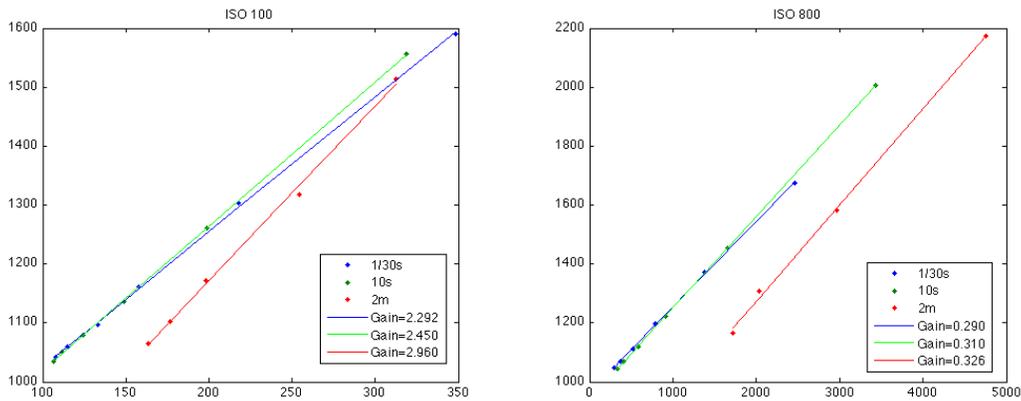
Interesting things can be seen in these plots. For ISO 100, the system gain went from 2.292 e-/ADU to 2.45 and then 2.96 e-/ADU as the exposure duration increased. That's almost a 30% change in the gain function. At ISO 800, we went from 0.29 up to 0.31, and then 0.326 e-/ADU – a 12% increase in gain. Gain here translates into contrast in your image. *It takes more electrons (and more photons) to change the intensity in the image as the exposure duration goes up (and as the thermal component kicks in more), lowering the contrast in your image. So, as the temperature changes, the contrast in your image changes as well.*

There's another component we can see here as well. Notice where each curve hits the y-axis? In the previous figure and for the short exposures here, it's somewhere around 800 plus or minus a bit. For ISO 100, this goes from 796 to 774 to 557. For ISO 800 this goes from 964 to 940 to 620. So, the gain (slope of the line) is increasing and the offset (y-intercept) is decreasing in both cases as we increase the exposure. This would imply that for something very dim (dimmer than the flats I recorded here as the first points you see on the line), the output image intensity would read darker if recorded with a 2 minute exposure than if recorded with a short exposure. *Not only is it artificially lowering the contrast, it's darkening the image as well.*

Together, these data indicate that Canon is both scaling the intensity of the image and shifting the intensity of the image before it hits the CR2 file.

## Estimating the Actual Dark Current?
We can use all of the above to try to get some handle on what the dark current actually is. In the two-minute exposures at ISO 800, we had a system gain of 0.326 and the offset (y-intercept) was 620. These parameters let us pass in a variance as the x term in y=mx+b and estimate the mean signal we should expect. If we pass in the 4900 variance we were getting after 2 hours of use, we get 2561 for the expected mean. As the offset is our estimate of a zero exposure (by not using a bias here we get around shifts in this offset actually), we get an estimated 798 ADU/minute or 260 e-/minute. If we pass in the 1200 or so variance early in the run, we get 64 e-/minute. Given the 13C temperature rise, the dark current should have gone up by just more than a factor of 4 and that lines up here with the 64 e-/min to 260 e-/min rate. While that aspect lines up with expectations, the numbers are far higher than I would have expected. We may just not know enough about the processing to accurately estimate the true dark current. The dark noise, though, does behave predictably and does paint a clear picture of an increase in noise with exposure duration and with camera warm-up, pointing to a significant dark current.

Dynamic Range vs. ISO

What ISO to shoot at is often the topic of discussion in Internet forums. From the basic transfer functions, I suggested that running beyond 400 ISO certainly isn't helping you in terms of picking out finer gradations amongst the wisps of galaxy arms and the like. Each intensity unit (ADU) is less than one electron at and above 400 ISO. But, there's the firm belief among some that it's pulling out fainter details to run at higher ISOs. The image is, after all, brighter. The brightness here is deceiving, however, as it's just an artificial boost according to the transfer function. (Note, for normal terrestrial shooting of JPEGs, the higher ISO does have a real use as it gets the image bright enough so that, when truncated to 8-bits, you've still got something.)

So, is it getting you more low-end detail? Have you lost anything in the process? From what I see, the answers are "no" and "yes", leading to the solid recommendation that you do not shoot with the higher ISOs (the actual optimal will vary from model to model).

Here, we have a series of flats taken at ISO 100, 400, and 1600. The intensity level was varied in a consistent manner by adjusting the shutter speed and the f-stop. The first panel here shows a "bright" setting to see where the various ISOs saturate relative to each other.
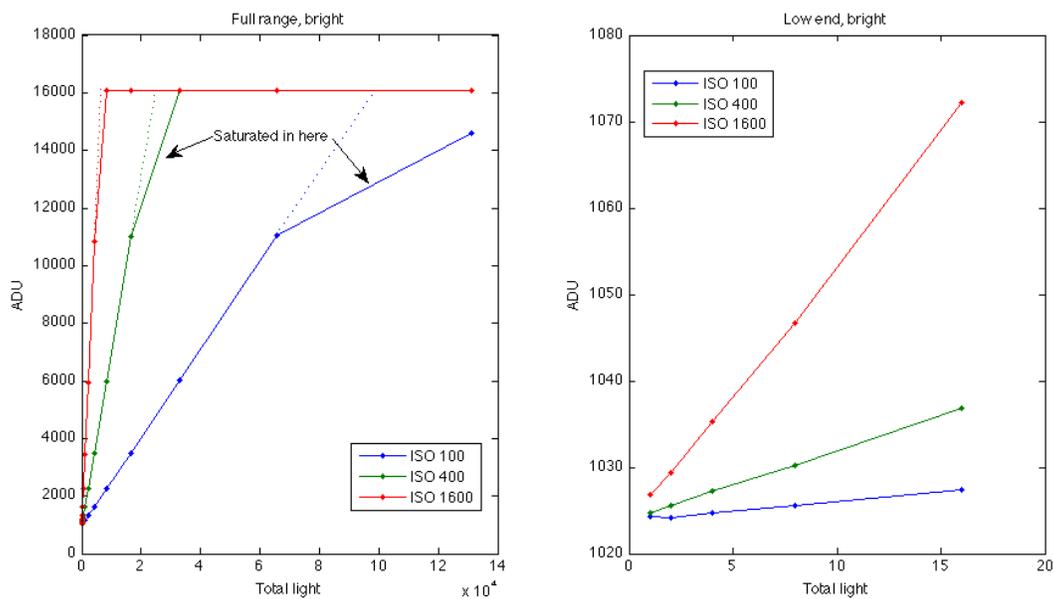

On the left, I'm showing the full range of the data with the x-axis being the total amount of light (some unknown scaling constant times a number of photons) and the y-axis being the raw ADU. There are actually 18 intensity levels in these data. The ISO 1600 (red) line has 5 saturated points with the first of these saturated ones not being too far off of from that red dotted line. That dotted line shows where that solid line segment should have gone if things were linear (there's no reason to expect we'd saturate exactly at one of my points). So, point #13 is good (the one at about 11000 ADU). If we look at the panel on the right, we have zoomed in on just the left-most 5 points to get a view of the dim end. ISO 1600 is going great here. So, to a close guess, we'd say that we've got 13.5 good points. The key for now, is that we've lost about 5.5

points (f-stops, bits) on the top end.

Turning to the ISO 400 (green) line we see we've lost about 3.5 points (f-stops, bits) here. The blue has 1.5 lost. This makes sense as we've got two f-stops worth of shift between each ISO (100, skip 200, 400, skip 800, 1600) and two get lost to saturation at the top end here (1.5 vs. 3.5 vs. 5.5). Thus, by increasing the ISO we've given up the ability to accurately record the bright bits as well as they'll hit saturation sooner. Has this been a fair trade?

To see if it's a fair trade, we need to look at the low-end of the response more than this "bright" series lets us as it's not clear from the right-hand panel above exactly what's going on. So, I ran another series with less light hitting the sensor (filter in place). That is shown below:
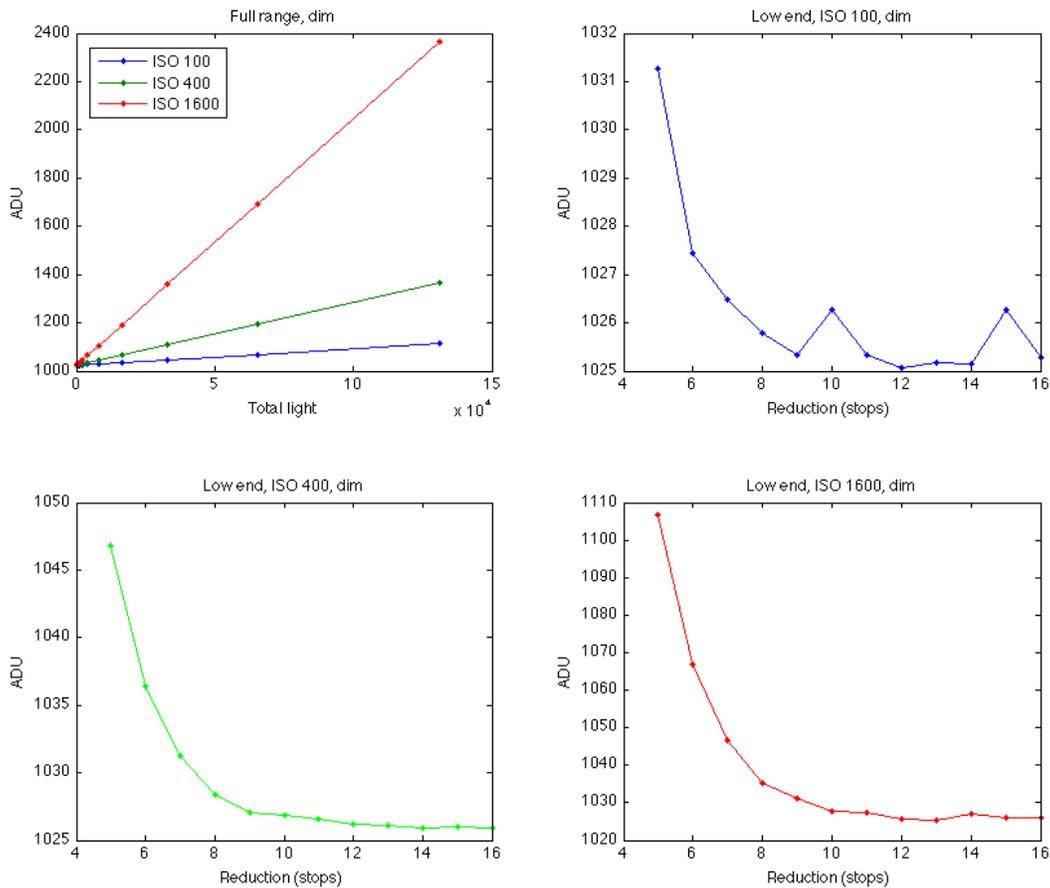
The upper-left panel is much as before (except we're not hitting saturation with any of the exposure settings). But, it shows the data are behaving cleanly. The other three plots now break out the response for each ISO setting. Note, I'm also plotting them differently so that we can get a much cleaner view of the low-end of the scale (where we spend so much of our time as DSO photographers). Now, the x-axis is in powers of two (f-stops) and dimmer is to the right as this is reduction in light).

For the ISO 100 case, we have 3 or perhaps 4 points that are clean (points #5-7 or 5-8) before we

hit the floor (the odd bump at 10 and at 15, noise here, prevents me from clearly giving it 4, but remember at ISO 100 we are undersampling our intensity scale being at 2.3 e-/ADU). At ISO 400, those first four are clearly good and really we may be good out to point #12. I'd be hard-pressed to say that ISO 1600 (red) is doing better than 400 as it certainly appears to have flat-lined before ISO 400 even did. We can confirm this by zooming in on points 9-16 for just these two ISOs (ISO 100 had clearly lost it).
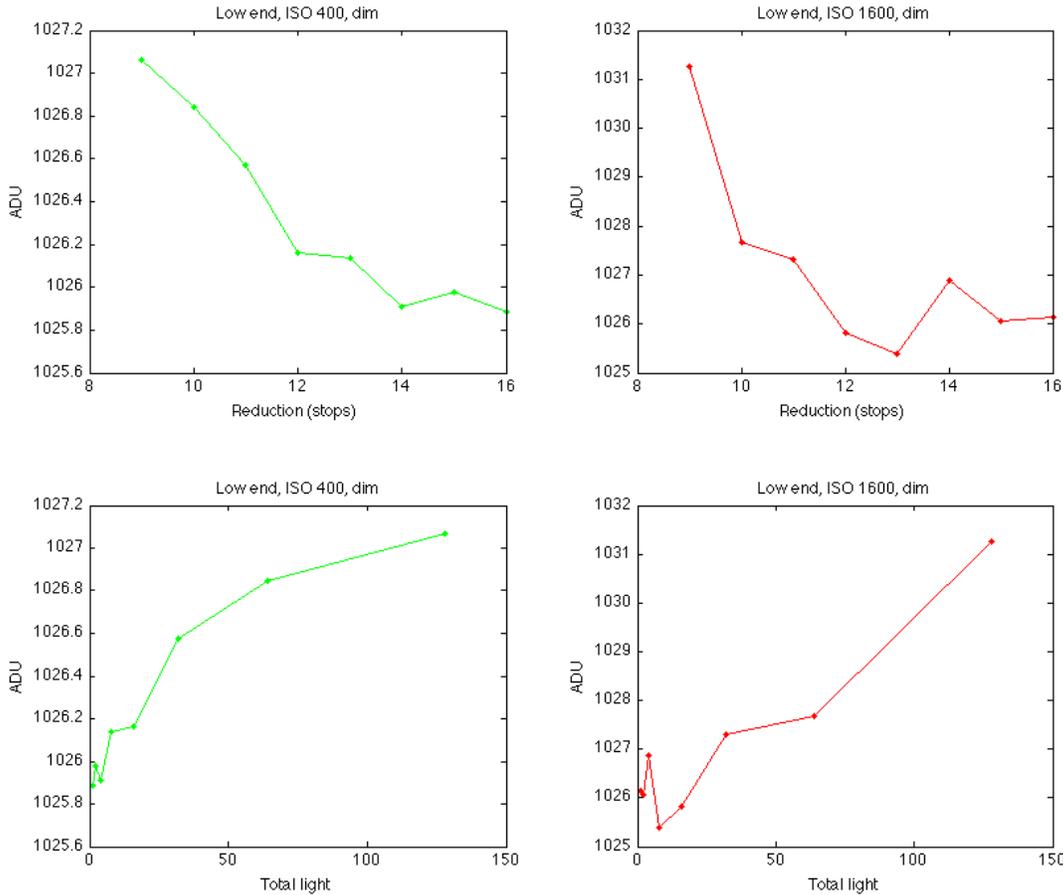
Here, looking at the top row, ISO 400 is going strong for perhaps 5 points (9-13). ISO 1600 on the other hand is good for 3 (9-11) and the rest are all the same (12-16) being at the level of the noise. If we go back to the other style of plot, with the x-axis being proportional to the total number of photons (bottom row), the response should be linear as it was before. Remember here, the axes are flipped so the left 5 points were were looking at in the top row of ISO 400 (aka points 9-13) are now the right 5 points (light levels 8-128). It's clear we're not perfectly linear here (I can't say if this is measurement error or true non-linearity), but I'd be happy with those 5 intensity levels. Again on the ISO 1600, we've got 3 good points, and the rest are bouncing around as noise.

So, let's bring this all together here. Relative to ISO 400, ISO 100 gained 2 points (f-stops, bits) on the bright end but lost 5 on the low end (being good to point 8 vs. 13). So, the net is a loss of 3 points on our dynamic range scale. Relative to ISO 400, ISO 1600 lost 2 on the bright end. It

also lost 2 points on the dim end, meaning its dynamic range is down 4 points on this metric. Even if you claim the low end is identical and don't buy the idea that ISO 400 is reaching dimmer levels more cleanly, nothing in this data suggests ISO 1600 is pulling more out of the faint bits. If you do this, you're still shy 2 points on this scale when using ISO 1600 versus 400.

You may have noticed when I refer to points here, I put "f-stops, bits" in parentheses. Each one of my points doubles or halves the amount of light hitting the sensor, just like moving one f-stop. These points, are a power-of-two scale. Two "points" is two f-stops, or four times the amount of light. The dynamic range in ISO 400 is therefore 4 times higher than the other two ISOs. Or, to put it another way, it's as if you went from 14-bits to 12-bits. Sure, running at ISO 1600 is brighter. But, you've not got any more information and, in fact, you've got less.

What Can We Do (or) Why Don't I See This?

If you try to run through and test your camera, there's a good chance you won't see the things I'm pointing out here. You may do things like load your CR2 files into PhotoShop or even convert them with dcraw and see that your darks do, in fact, increase their mean intensity with exposure duration. At this point, though, resist the temptation to say all is right with the world. The problem again is that we're working with tools designed to make great photos and not tools designed to preserve the accuracy required for things like dark subtraction to work.

Here's one way you can check this. David Coffin's "dcraw" is available for pretty much anything under the sun and it is the *de facto* standard for RAW file conversion. Many programs out there use it as a back-end for their RAW file decoding. A nice feature of dcraw is that it lets you control how much processing is happening. So, for example, you might type "dcraw –D –r 1 1 1 1 -4 –T IMG_1000.CR2" (substituting your actual CR2 file there at the end. What this does is to tell dcraw to write a TIFF file (-T), using no white balance correction (-r 1 1 1 1), with 16-bit linear output (-4) and to just pull out the raw sensor data (no color interpolation) with no scaling (-D). You should now see the issues. If you substitute "-d" for "-D" here, you engage a color scaling routine (this also gets engaged if you run the debayers). What this routine does (it's the "scale_color" routine in dcraw.c) is to determine the range of the data (a white point minus a black point) and, from that, a scaling constant (you can see the constant with –v). It uses this range and the black point to shift (it subtracts the black point) and then scale (multiplies by that scaling constant) the data. Run this on your images and you'll see that the darks do, in fact, increase with exposure duration.

So, we just should do that, right? Well, no, not necessarily. What that relies on is the black point. It gets that black point by looking in the optical black area for the mean signal there and not from something Canon recorded. This optical black area has been scaled just the same as the rest (I've verified this). So, what we're doing here then is to use an estimate of the signal level where no light hits and to set this to be zero. That's great for daytime use but any thermal current (which we'd want to estimate) for non-hot pixels gets set to zero. It's fixing the appearance of the issue but not actually restoring the accurate data. In addition, I should add that while the mean now shows a nice linear increase, the median – or middle-most number – is now zero. In "fixing" the problem, it's zeroed out over half the pixels. Indeed anything that wasn't a warm pixel got set to zero here in the darks and that's not something we'd like to do to our data. Since the optical black is a dark frame and since it's using this to estimate the black level, it

should come as little surprise then that the whole dark frame got wiped out as a result of this processing

I re-ran the ISO 100 1/30s vs. 2m transfer function using Canon's Digital Photo Professional and using dcraw to see if their treatment of the flats (which won't get zeroed out) would do anything to fix the contrast alteration. Try as I might, I could not get any reasonable transfer functions out of the image pairs from Canon's software. Despite trying to make it as linear as possible and to disable everything I could, the points would never fall along a straight line. It's just doing too much processing to the data to get meaningful results. Using dcraw, things behaved well and again a difference was noted between the short-exposure and the long-exposure gain. Running with "-D" to apply no scaling (and then extracting the red channel), the numbers were identical to that in Nebulosity. Allowing it to run some scaling with "-d" and the gain values themselves change (as it scales the intensity by a factor of ~4.8, the gain values scaled by this as well). But, the pattern was exactly the same. The contrast alteration is still in the data even if the mean dark current versus exposure plot is no longer U-shaped.

I'm sometimes asked why frames look different in Nebulosity than they do in Canon's software, in Photoshop, or other daytime programs. This is why. I don't apply any processing apart from the decoding to the image so that we have as good a chance as possible to let pre-processing work. These other packages do a much better job at making normal, daytime shots look great as that's what they're designed for. But, they are processing the data and adding to what's happening inside the camera. This will make bias, flats, and darks even that much more problematic. Note that you can use dcraw (or any front-end to dcraw) in these modes as well.

Conclusions

So where does this leave us? There are several take-home messages from these results:

1. Canon is re-scaling your data before it hits the CR2 file. Based on the thermal signal (likely based on stats from the optical black portion), it is both shifting your histogram left (i.e., subtracting a constant from the whole image) and scaling the intensity (changing the contrast or gain).
2. This makes the camera appear to have very low dark current as the background never gets appreciably brighter. But, the noise increase shows that the dark current is there and it is really the noise from the dark current that's the trouble – far more than the constant component of the increase.
3. The camera warms up for a substantial period of time.
4. The above makes dark subtraction a real challenge. Not only is the current changing with time as the camera warms up (which we might easily account for), but what scaling is being applied to the data is changing as well.
5. Software designed for daytime photography will also rework the data and making it stay purely linear and not apply any processing can be difficult.
6. The camera's internal gain (e-/ADU) for each ISO value points toward limiting the ISO to 400 and not using higher values. Both in theory and in practice, using higher values limits the dynamic range and does not let you pull out fainter details from the noise (even if they look brighter).

We might ask ourselves at this point why the data get rescaled in the first place. Canon, after all,

is filled with bright engineers, the cameras are very successful, they take great images, and this practice has been going on in their cameras for some time (at least since the old DIGIC II 350XT). The answer, in my opinion, is that the scaling makes perfect sense. They are compensating for inherent constraints placed on them by the sensor (which in turn gets to blame physics – good luck winning that debate!). Today's DSLRs will do in-camera compensation for all sorts of things, now including lens distortion and chromatic aberration. These corrections get into the raw data and are far more complex than dark current.

The overwhelming majority of users would want the camera to avoid shifting the histogram far to the right as the exposure lengthens. Keep things under control as best as you can to make an image! Surely that's better than having the image get lost by being washed out. Likewise, since many people shoot in 8-bit JPEG, ISO ratings well above what might be optimal for dynamic range are a good thing as they boost the signal into the range of intensity values that work well for the 8-bit, gamma-stretched JPEGs. It lets you get an image and see it there on the screen without going back and using image processing software to stretch the raw data. So what if a bit of dynamic range has been lost – you have an image!

These are perfect engineering arguments (as is the one to have an IR blocking filter cut off the H-alpha line). The Canon DSLRs do very well in astrophotography. But, they're not designed from the ground up for this. They're designed for a different market and their engineers make different choices as a result. Some of the choices impact how well the camera works for astrophotography.

Hopefully, we now know a good bit more of what we're up against. I, now, at least have some information that will let me determine better means of doing dark-correction in software. For the times when I will grab my DSLR for astro work, I also know more about the ISO setting and how to optimize it for my images.