# Aquarius Version 5 Salinity Product Tutorial

Aquarius Instrument-Based Calibration

Dr. Sid Misra (NASA/JPL)

Hi, my name is Sid Misra. I work at the NASA Jet Propulsion Lab in Pasadena, California. What we're going to show you here is a video tutorial that gives you a general overview of how we're implementing our calibration for the Aquarius data.

This particular video tutorial is going to deal with a specific aspect of the radiometer calibration and we're calling it the "instrument-based calibration." This correction is one in a series of improvements we're making toward the Version 5 release.

If you want to learn more about the Aquarius Mission itself, please go to aquarius.nasa.gov. Aquarius data is available on the PO.DAAC. PO.DAAC stands for Physical Oceanography Distributed Active Archive Center. It's hosted by JPL and the website is podaac.jpl.nasa.gov. And you can get your Aquarius data and other interesting oceanographic data from the same website.

If you're not familiar with the Aquarius Version 4 data, you can always go to the website that you see on your screen and take a look on how we went from radiometer data all the way to salinity, with all the steps in the middle. The data was released in 2015 and it was very successful data but there's always improvements that we can make. And that's why we're going from this Version 4 to Version 5, making those tiny little adjustments to improve the data even further.

In order to get salinity measurements from our radiometer antenna temperature measurement, you have to go through various steps. If you look at this flow chart on your screen: at the bottom you see "salinity" and at the top you see "counts." You have your salinity measurement that's coming out of the ocean. Now the ocean is a bit rough, so you have to add roughness to that factor. Then that brightness temperature that comes out the ocean has to go through the atmosphere. The atmosphere adds its own microwave thermal energy to the data. So that's something they have to take into account. It has to go through the ionosphere. Aquarius operates at 1.4 GHz and the ionosphere actually affects Aquarius data significantly because it takes all the polarizations of the Aquarius data, which is something that we can to keep into account.

So fine, we go through Faraday rotation in the ionosphere and we reach the antenna of Aquarius. Now at this point, the antenna of Aquarius is not only seeing thermal energy from the ocean, it is also seeing [reflections from] emissions from the galaxy, emissions from the sun, the moon. It's also seeing reflections of the sun going from the ocean, back into the antenna. All these combine together to form something called "Total Antenna Temperature."

Now that is what goes in to the radiometer. The radiometer then converts that in voltages. The radiometer chain is basically a series of amplifiers well calibrated, in very precise thermal condition that convert your antenna temperature to voltage to something that's called "uncalibrated counts" at the end.

Once we get this data from Aquarius, we have to do all the steps in reverse. We take those counts and then go forward, try to get antenna temperature from those counts, "Total Antenna Temperature." We try to get "Earth Antenna Temperature," so on and so forth. We get the antenna temperature at the top of the ionosphere, top of the atmosphere, to the ocean, to remove the roughness, get the brightness temperature of the ocean and from that value, we try to get the salinity.

What I will be talking about here is only the first couple of steps of that whole process where we go from counts to salinity. I will be talking about you go from counts that are measured by the radiometer to the "Total Antenna Temperature," that is right at the antenna of Aquarius. There is a very simple formula, technically, that can represent this relationship between counts and antenna temperature. It's simply a gain and offset formulation. The antenna temperature times the some gain, which is the gain of the radiometer plus some offset, which is added by the radiometer, is equal to your counts. Even though calibration itself is very complicated, there are various moving parts, at the end of the day all we care about is getting a good number for "G" and a good number for "O" and how it varies over time. And that's what we concentrate on. We'll basically demonstrate how we're working with those two variables to get a better product out.

As I said we will concentrate on only the parts from counts to antenna temperature. Why are we concentrating on that? This is because when Aquarius launched, at the very start we noticed some calibration anomalies, specifically two types of calibration anomalies. One was a drift and the other was a pseudo-random oscillation. The Aquarius folks lovingly like to call those pseudo-random oscillations "wiggles." So from now on, I will be using the term "wiggles" interchangeably with pseudo-random oscillations. If you look at the plot what you're seeing is the difference between the <u>measured</u> antenna temperature of Aquarius and the <u>expected</u> temperature of Aquarius for the v-polarization Beam 2. What's the expected temperature of Aquarius? It's basically something that we generated using a salinity model, an atmospheric model, an ionosphere model, and an antenna model to get, more or less, an expected temperature.

Now if there were no calibration anomalies, what would we see? We'd probably see a flat line with a bit of noise around it because these expected antenna temperatures, remember basically they're our best guess and they're based off models. The plot that you see: the blue curve, you see, it's actually not flat. Right at the start there's a slow exponential decay in the Ta minus the expected. That basically tells you that Ta is slowly decreasing with respect to what you expect. We call that anomaly a "drift." And it's represented by the red line.

The other anomaly is the wiggles, which are these sort of sinusoidal oscillations that you see on the plot. Right at the start you see the random oscillations. The interesting thing about these oscillations is that they were different for all six channels. We tried to factor out common factors between all six channels — beams 1, 2, 3; v-pol and h-pol** — and found out that they weren't common. Every instrument had a different periodicity for these wiggles. So those are the two main issues that we're going to address.

So let's first understand how these wiggles and drifts are currently calibrated. To compensate for any instrument calibration anomalies, Aquarius salinity measurements are obviously measured. But, on the other hand, we have this model called the HYCOM model. The HYCOM model gives us global salinity measurements — as best as it can — and what we do from that is we create antenna temperatures from the modeled salinity measurements. That's what your "Ta expected" measurement is [that we were talking about a couple of minutes earlier]. So from the HYCOM model, we get "Ta expected." And now we do a long running average filter that is seven days. And we also do a global mean over the ocean off this "Ta expected." So basically taking the "Ta expected" over the ocean, averaging it over the whole globe and then doing a running average over seven days. And then you compare that heavily averaged number to your measured “Ta’s.” And that clearly starts showing you these wiggles and drifts.

Now once you have these HYCOM Texpected measurements, you take them and you basically apply corrections to your original Ta’s. You apply corrections to the noise diode of the radiometer. What’s the noise diode of the radiometer? It’s a component that helps you measure the gain, the “G” term of the radiometer, and any other term you also apply the correction to and the offset of the radiometer, which is the “O” term.

Now it’s important to understand these are done over large time scales and the global mean, which basically means that Aquarius itself is still getting you that good data that no model can ever produce. But over large scales these models are good enough to at least constrain your Aquarius calibration. So our purpose here is: (A) Let’s go back and find the root cause of what those calibration anomalies are; and (B) Let’s try our best to reduce our dependency on the HYCOM model. Even though the dependency exists on these large time and spatial scales, we still would like to reduce our dependency.

So first I’ll talk about the drift correction. If you look at the plot, all six channels indicated a drift when we compared it with respect to HYCOM expected antenna temperatures. And you’ll also notice that this drift was a lot more significant during the initial months post launch and then Aquarius basically settled down for the last couple of years. A general theory, which is accepted widely, is that this drift was probably caused by outgassing of the Aquarius instrument. So what is outgassing? Many of these microwave components when they’re on Earth they trap gasses inside them. When you launch into the vacuum of space, these gasses start escaping these components. It could have many effects, it could cause condensation on your instruments, it could basically degrade the performance of your instruments. Or, in our case, it could change the performance of the instrument. So, for example, let’s say the noise diode was going through an outgassing event. Pre-launch we measured an Excess Noise Ratio (ENR) of the noise diode. Post launch, outgassing happens and that ENR value slightly changed, which means that the noise diode temperature slightly changed. We use that noise diode temperature to finally get the Ta’s. So if you don’t know the noise diode temperature that we calculated pre-launch well enough, you start seeing these slowly varying errors in Ta’s over time. So that’s what outgassing is and it can have many effects on your radiometer.

So now that we have the drift and we suspect it’s outgassing, the first thing to do is figure out is this drift a gain drift or an offset drift? I’ll go back to that equation [that we had]: the “G” and “O” terms. The drift can be in either one of them and when you compare it with respect to the ocean, it will look like that. Now if you correct the wrong term, what it would do is mess up the antenna temperature calibration over land, let’s say, or maybe over ice. It would be good for the ocean but it would be wrong for everywhere else. You have to get these things right. So the first thing we did try to figure out if this would be the gain or offset. In order to do that, we needed an external source not present inside the radiometer. You can’t calibrate the calibrators of the radiometer with instrument parameters itself. You need go outside and look at a stable source outside the radiometer. And, in this case, we chose two of them. One was the ocean, which we already know… we already have a model for it. Very stable target and, to a certain degree, we know how it behaves.

And the other one we chose was Antarctica. For Aquarius we chose this one particular region of Antarctica, which we noticed was temporally and spatially very stable. And we used L-band measurements of that region to figure out what Texpected should be over the ice. Remember we have Texpected over the ocean. We want the similar Texpected over ice. But the reason we chose ice and Antarctica is because at L band (1.4 GHz), which is what Aquarius operates at, ice is extremely stable. In front of you, you see a plot that's depth with respect to time and the color represents the temperature of ice. At the surface of ice over summer and winter, the temperature variation is pretty high. Over time, that variation very slowly propagates deeper and deeper into the ice. And L band just being at such a low frequency, it can actually penetrate very deep into the ice. So what that means is, interestingly enough, we're actually looking at temperatures that were a couple of seasons ago. But not only that, the amplitude variations deeper into the ice become smaller and smaller, which means the deeper you're looking into the ice, the temperature variation over seasons is very stable. And this is something that we can model. We took Antarctica, we modeled it. If you look at this plot for Aquarius v-pol horn 3, you can see the blue curve represents what Aquarius saw over Antarctica and the black curve is what we modeled. So the black curve is our Texpected and our blue curves are Ta. So now we not only have a Ta minus Texpected over the ocean, we also have a Ta minus Texpected over Antarctica. And this will give us a straight shot on figuring out whether it's a gain drift or an offset drift.

Now just imagine if this were an offset drift. What does that mean? An offset drift is something that's added on to your Ta. It doesn't care what Ta is. Ta could be looking over the ocean, which is 100 Kelvin the drift would be the same. It could over the ice, the drift would be the same amount. It could be over the Amazon, which is 300 Kelvin, the drift would be the same.

On the other hand, if this is a gain drift, it does not behave the same way. Aquarius is calibrated with respect to a reference load at 300 Kelvin. All gain calculations are made with respect to 300 Kelvin, which means if there's a gain draft at 300 Kelvin, looking over land or the Amazon, you would not see it. At 200 Kelvin, you would see some of the drift. At 100 Kelvin, you would see twice the amount of drift you saw at 200 Kelvin, and so on and so forth. So all we have to do is compare the Ta minus Texpected that we measured over the ice and the Ta minus Texpected that we measured over the ocean. And see if the drift is the same or if it's scaled between the two to figure out what kind of drift this is.

And that is exactly what we did. If you look at the six plots we compared the ocean with the model, which is the green curve. The blue curve is the drift measured with Antarctica and scaled. We had to scale it by a factor of approximately two to match over the ocean. And that immediately gave us the clue that this was definitely a gain drift. If you notice all six plots, the blue curves and the green curves match up so well on top of each other that we're fairly certain that this is nothing else but a gain drift in all six channels. So now that we know that this is a gain drift, we know how to correct for it. To correct for the gain drift, we have to correct that "G" term. The "G" term is derived off the noise diode temperature that I was talking about a bit earlier. So all we have to do is vary the noise diode temperatures at the same rate as the drift as the gain drift is occurring and, lo and behold, we basically fix the drift. So for Version 4, what we did is we used the HYCOM model to fix the gain drift. Once we figured out it is a gain drift, we the ocean model itself to fix the noise diode temperature. For Version 5, are working towards using the Antarctica model only for the gain correction and not use the HYCOM model. We would like to try our best to remove ourselves from the HYCOM model and be completely independent of it.

00:37:33 - 00:38:57 Restatement of last bullet on Slide 9… not as good as previous or next option

Previously we had shown you how to go from “Antenna Temperatures” to “Counts.” There was your “G” term and your “O” term. This equation is basically the reverse of what we need to do once we have the radiometer counts and we need to go back towards antenna temperatures.

The four terms of importance are CA, CR, G, and TR. CA is the count that the radiometer measures when you’re looking out the antenna, at whatever scene it might be. It might be the ocean, land, cold sky, ice. CR is an internal source. It’s called a reference load. It’s there for exactly what you think it is: it’s supposed to give the radiometer a reference as to what the actual temperature should be. The reference load, in general, is kept around 300 Kelvin. We know that all of the time we have engineering data on the reference load.

So from time to time the radiometer, instead of looking out the antenna, looks at the reference load. And then we use that reference load to calibrate our antenna counts. Of course, that’s only the counts and that only helps you with offsets. You still need to take into account the gain of the radiometer system itself. Remember antenna temperature gets converted into voltages, it goes through a bunch of amplifiers, and then it gets converted into counts. So those amplifiers give the radiometer a particular gain. In our case, we measure it in terms of counts per Kelvin. That is something that we measure pre-launch, characterize it pre-launch with respect to various factors including temperature. Once Aquarius launched, we verified this term by doing occasional calibration with respect to the ocean, cold sky, Amazon, ice, what have you.

And the last term is TR, which is the reference temperature, the actual physical temperature of the reference load counts. And that term is there to give you the offset that the radiometer might introduce. So just imagine in this term if antenna counts were zero — that there were no antenna counts — so CR over G is typically would be equal to TR. And those two terms would cancel out. But if for some reason there is some offset [this], those two terms would give you that offset term. So, all in all, that equation is a great term to go from counts to antenna temperatures. And we will be using this equation a lot in our description of wiggle corrections.

All right, so now that we’ve covered instrument drift, we can now move on to the wiggles correction that occur for Version 5.

Now, wiggle correction is an interesting thing and it’s not that easy to grasp, But it all ultimately boils down to the TA=CA-CR equations. The wiggles occur due to errors or offsets introduced into the term CR, the reference load counts. We’ll see how that happens.

You have your antenna temperatures coming into your antenna, all your reference load temperatures coming in while looking at the reference load. Both of them get converted into voltages and go through the radiometer gain change, through all of these amplifiers, then it goes through a device called the Voltage to Frequency Converter, the VFC. The Voltage to Frequency Converter is exactly what it sounds like. It takes in the voltage and converts it into an output signal to the particular frequency. So the voltage goes up, the output signal goes out at a higher frequency. That frequency then goes into this counter, so the more oscillations there are, the counter counts higher. And, lo and behold, the counts are directly proportional to your frequency, which is directly proportional to your voltage, which is basically proportional to your antenna temperature.

The wiggles, unfortunately, happen due to a weird quirk that occurred in the Aquarius VFC. It's called locking. What happens is, as the voltage increases, at some point the VFC suddenly locks at a certain frequency. So even though the voltage is increasing, the frequency stays the same. If the frequency stays the same, then the count value stays the same. So, at the end of the day, even though the antenna temperature is increasing or decreasing, for certain regions your counts don't budge, they stay the same.

We can clearly see this in the histogram of counts. The histogram you are looking at is a histogram of all six channels of Aquarius. Different colors represent the different channels. Now, this is a natural, geophysical signal. This is the histogram of counts going over oceans, land, everything. Ideally it would be smooth, but you notice there are these spikes in the histogram. Why are those spikes there? All it means is those count values are being seen more than they should. The count values are stuck at those spike values more than its neighbors. Which, in the end, translates to the fact that those count values are due to the frequency locking, and the frequency locking is due to the quirk in the VFC.

Now what is that quirk? It's a bit too technical to get into, but just imagine your VFC is nothing but a system of op-amps – imagine a clock coming into it, or imagine a clock from some other system leaking into your VFC. And that clock is operating at a certain frequency and what it does is basically phase locks your VFC at a certain output frequency. So it's not your normal noise, it's some sort of periodic noise – which is what a clock is – which is causing your output of the VFC to get locked. And it not only gets locked at those frequencies, but also harmonics of those frequencies, which is why you see so many spikes in the histogram.

Another interesting thing to note is between all six channels, the spikes actually occur at the same count point, which means it is very unique to the VFC used, which is the same for all six channels.

So in the plot you see is a histogram of one of the Aquarius channels, that's in blue, and another Aquarius channel is represented in red. These histograms in general should be ideally very smooth. It's looking at ocean, it's looking at land, it's looking at ice, these are smooth transitions, there's no need for these count histograms to have any spikes in them. The spikes are exactly due to the VFC locking that I was talking about. All those spikes mean that a spike is occurring more than its neighbor, a count is being observed more than its neighboring count. Why is a count being observed at more than its neighboring count? It's because the voltage to frequency converter is stuck at that frequency even though the input voltage is changing.

The asterisk, for example, represents the reference load counts. Over the period of the mission, the reference load counts do not stay stable. They slowly drift in and out, go from say 900 all the way down to 700 and then go back up. And that's not a big deal, it's something that us in the radiometry community know and expect, and it's why we have all these calibration terms.

But the problem now is, as the reference load counts are drifting slowly over many, many weeks and months it is going in and out of these locking frequencies for these spikes, and as goes in and out of these locking frequencies it is getting biased in one way or another.

The plot you see on the top is divided into three subplots. One is a histogram. The one in the middle is what your reference load mean would be. The last one is a bias introduced due to these locking frequencies. Now just imagine your reference load count slowly drifting from 855 to 835. Now, as it is drifting, suddenly, at 845, it encounters a locking frequency. What does that mean? That means that the histogram, instead of being a nice Gaussian, has a spike at one end. And it's slowly drifting through that, the spike goes from one end to the other end. If you take the mean of that Gaussian signal, now there is a bias at one end, as it drifts towards the spike it goes back to zero and then gets biased at the other end.

Now imagine for the whole mission, the reference load count is going though many such spike points or locking points. Which means it gets biased low and high and comes back to zero, hits the next locking point and gets biased low, zero, high, zero and so on as it keeps hitting these locking frequencies.

That is exactly, over a long period of time, what results in the wiggles that you see in the TA. So, go back to the equation. The CR is being biased low and high, slowly, over time as it goes through these locking frequencies. Which means the whole equation is getting biased low and high as it goes through these locking frequencies, which is exactly what you see in TA when you look at the wiggles, it looks like a pseudo-periodic random oscillation or wiggle.

The reason that these wiggles were different for all six channels, the reference load counts, as you can notice in the histogram, exist in different regimes. So, some go through no spike counts at all, some go through many locking frequencies. Some go through locking frequencies at a different time, some go through locking frequencies at a different pace, which is why the oscillations don't match with each other with respect to channels. But when you boil it down, down to the count level, it all makes sense.

So now that we know the root cause of the wiggles, we know what to correct, all we need to do is figure out a way how to correct it. We basically need to find a function that is dependent on the value of the count that tells you what amount of wiggle error -- or offset error or mean bias –that we need to correct for, for that particular count, and then apply it to the whole mission.

We actually take advantage of an interesting behavior we observed in Aquarius. We noticed in Aquarius that two consecutive reference load samples were slightly offset from each other. Aquarius, as it's orbiting measures antenna, antenna, antenna, antenna samples, and then suddenly it measures a reference sample, a reference sample and then two noise diode samples, and this is to achieve its calibration. And what we noticed is that the two reference samples were slightly off set from each other in counts. They were slightly offset from each other in counts, but drifting in the same way, up and down, as you can see in the plot.

The blue is reference sample two, the red is reference sample one, and you can see they are slightly off of each other but their main behavior over the four years is the same. What that means is that these two counts are going in and out of locking frequencies at different times. Which means they are affected by the mean bias differently at any given time. The first thing we did is to confirm that. We subtracted the two reference load counts samples with each other. And we noticed that the behavior was very, very similar to the wiggle behavior that we saw in TA-Texpected. So that was another set of confirmation, like, alright, this is something that we can not only use to confirm that the reference load counts are probably responsible for the wiggles but we can actually use the reference load difference to fix this.

Now how do we fix this? So, if you know the mean bias between the slight offset that exists between the two reference load counts, anything extra on top of it, is basically, in mathematical terms, a differential function of the wiggle correction. Just imagine your wiggle correction term being f(x), and your count being x. Your complete term is x + f(x). The other reference load count has something like y + f(y). Now I subtract the two and now I have f(x) – f{y) and x-y. f(x) – f(y) is basically nothing but the differential term of the function f(x), and that's what we need to solve. All we need to do is ~~was~~ to figure out how to get the wiggle function was to solve the differential function, which is nothing but the reference load difference.

Now, going back into our grad school days of math: to solve a differential function, all we need is a starting point. Once we have that starting point we can actually solve that whole differential equation. If you know exactly the amount of wiggle added at a particular point and you exactly know the constant offset between the two, you can basically go back and forth between the two terms and solve this differential equation and get the wiggle correction for the whole series of reference load counts for the whole mission. And that is what we do.

How do we get that starting point? If we think about it, the mean bias is in one way, when your reference counts are drifting though a locking frequency and it goes to zero when the locking frequencies are right in the middle of your histogram, and then it goes through the other way. We know where these spikes are, so we know exactly at those locking frequencies the mean bias is zero. The wiggle correction is zero. And that's our starting point. We take a locking frequency count value, we assign a zero wiggle error correction to it, and from then on we propagate that correction to both ends and solve the differential equation. And that's how we basically solve for the wiggle error and obtain a correction for each count point.

The plot is the result of that differential solution. The result you see is the wiggle correction, the y axis is the wiggle correction and the x axis is with respect to which reference load count value that particular correction term is applied to.

Now that you have this term you can apply it to the whole mission going back to the equation CA-CR, this correction gets applied to the CR term. The blue plot is what we had before applying the wiggle correction. The red plot is what we have after applying the wiggle correction. And the most clear case of it, if you just look around 2012, those wiggles have disappeared. You will notice there are a couple of residuals left especially around early 2013, early 2014 but those are either other calibration issues, or geophysical model functions issues but definitely not wiggle correction issues.

So this is an interesting thing that came out of the wiggle correction: once we went away from the ocean-model correction to an instrument-only correction and applied our drift correction and wiggle correction a couple of residuals started coming out that initially were just corrected by the ocean model. These are tiny calibration errors. We are currently looking into them to take care of them, but all and all you can see it is a significant improvement on the calibration without using any external reference source, which is a huge improvement for us.

This plot here shows the end result of the instrument only wiggle correction for all six channels. The curves show what the actual correction was applied to the data. You can see in certain cases a lot of correction, in most cases actually, is applied at the start. H3 for example did not have wiggles at the start of the mission, but at the end of the mission it did have wiggles which you can clearly see when you look at 3H.

So that's how we applied the wiggle correction, and combined with the drift correction that we are now looking at over Antarctica. This is our way of slowly moving away from the HYCOM ocean reference model and slowly applying a calibration that is based on the instrument only. The wiggle correction is in the data stream and is what's going towards Version 5. We are working towards applying the Antarctic gain drift correction towards Version 5, as well. Future data due to these corrections might actually include some residual biases, which are very small, .1 Kelvin or less. But even these ones, we are trying to track down. All in all, what this does is gives us Aquarius data that is independent of the HYCOM model and is better for you in terms of long term salinity trends.

We certainly hope that this video helped you understand the instrument only correction a bit better. I hope you enjoyed the talk as much as I enjoyed giving the talk. And enjoy the data!