Ian Fretheim

WATER ACTIVITY IN SPECIALTY GREEN COFFEE:
A LONG TERM OBSERVATIONAL STUDY

Cafe Imports

Ian Fretheim
WATER ACTIVITY IN SPECIALTY GREEN COFFEE:
A LONG TERM OBSERVATIONAL STUDY

Cafe Imports

Contents

Introduction

Green coffee as it is currently conceived at the specialty level is both relatively new and highly dynamic. It is unlikely that the standards and perceptions that pertained in coffee quality even 10 years ago hold true today. Levels of quality acceptance do not simply increase as coffee buyers become more accustomed to a highly competitive baseline of production. Quality acceptance also changes in shape and dimension with trends and new developments—for example in post-harvest processing and cultivar hybridization. Flavors that were once rejected, such as those seeming to stem from coffee cherry in the processing of Washed coffee, become more differentiated and can eventually come to be accepted by specialty-coffee tasters.

While the terms of the challenge have varied, one consistent obstacle has been ensuring the as-purchased quality arrival of coffees. The full dynamism of the specialty industry comes to bear on this point and there are numerous areas that need to be addressed in discussing the issue. A primary area of focus in discussing the stability and instability of specialty green coffee are the physical characteristics (and methods of measurement of said characteristics) through which changes can occur.

A relatively new measurement to specialty coffee, though well established in other industries, is that of water activity. While moisture content is a measure of the amount of water in a system, water activity is a measure of the energy status of that water. The implications of this difference are very important and will become clear as the paper progresses. Briefly, water activity (the energy status of water rather than its quantity in a system) is the mechanism behind why foods mold and spoil, why sea water and fresh water freeze and boil at different temperatures and rates, and why moisture might move from one part of a system to another.

We began monitoring water activity in green coffee at Cafe Imports in 2012. At the time we hoped that it would provide a more insightful measurement than moisture content, specifically with regard to predicting the shelf life of coffees. We hoped to answer (or at least partially answer) the question of why some coffees arrive at the same quality as when they were purchased and some do not.

In this paper we will discuss numerous aspects of water activity in specialty green coffee on the basis of an ongoing multi-year observational study, as well as a number of small-scale trials that we set up along the way. These smaller trials, while neither perfect nor exhaustive, were set up to test our observation data as well as our own and others' theories. We ran these trials against the backdrop of collecting our larger 25,000 point water activity in green coffee data set.

Since beginning our study, we have accumulated a large amount of data. I do not know of a larger or even comparable data set for water activity in green coffee. This is important as interest in water activity has increased over these last six or seven years to the point where we are beginning to see people and organizations chime in on the possible value and use of water activity data points. Unfortunately these voices have come with little supporting evidence, occasionally even on the basis of flawed assumptions, while purporting to be demonstrated by small and ambiguous trials.

We will cut to the chase here and say that the establishment of any industry standard for water activity in specialty green coffee at this time is premature. Until practical data beyond general principles can be shown to support such a standard, and until water activity can be shown to offer a significant improvement over less-burdensome measures, standards for water activity in specialty coffee can serve only to reinforce the power structures by which the coffee-buying world (at least the those parts of it that can afford to run water activity testing) dominates the coffee-producing world.

Why is this so? Without getting too far ahead of ourselves, and as an example, there is currently a water activity (Aw) standard of 0.7000 Aw in specialty coffee. This is so lax as to be meaningless. Specialty coffee as it currently is prepared and exported all over the world rarely reaches this level. Further, the duration of unmonitored and uncontrolled shipment conditions combined with the observed changeability of water activity over this period strongly suggest the inclusion of a significant buffer into any Aw standard point.

The basic science stated behind the 0.7000 Aw standard is correct. However, it is incomplete in that it omits knowledge of the real behavior of the population under examination: water activity in specialty green coffee. This population and its behavior is the primary focus of our essay.

If a lax standard were merely meaningless, that would be one thing. Unfortunately, it can be worse than that. As soon as we establish a standard that is based on a physical measurement, that measurement must be made and those who are incapable of making that measurement lose agency. If a standard is neither meaningful nor accessible, in particular if there is a reasonable, established, and accessible substitute, it can be regressive. Standards that are intended to establish and improve upon the best practices current in our industry, if they are not fully vetted, can end up by eroding trust, limiting accessibility and reducing proper utilization. We need to carefully weigh the costs and the benefits when creating and implementing such standards.

Let's say for a moment that you are a coffee producer. You have been measuring moisture content at <12% as part of your quality program. Now you are told that you need to measure water activity at <0.7000 Aw. You find a way to get together the many thousands of dollars for a water activity meter and you start taking measurements. What do you see?

You see that measurements taken in the morning are different than those taken at night if you're not in an environmentally controlled space or using a very expensive device. Which of these readings should you use? If you select the wrong measurement, will you ruin your coffee? You also see that at 0.7000 Aw, your coffees are well above 12% moisture. In fact, they are generally much closer to 14%. Do you rejoice because now you can pull your coffees out of drying 2% sooner? Will buyers accept these? Or do you wonder why you were told to go out and spend considerable money on a water-activity meter to measure something that you had already been doing for years, 100 percent of the time and with significant margin for error?

If a paper towel company or the forestry lobby claims that commercially washed cotton towels are worse for the environment than paper towels, you better look at the data. If a decaffeination company claims that higher water activity leads to better roasting, you better look at the data. If a high-end specialty-coffee importer claims that water activity is not ready for prime-time (or that specialty coffee is not ready for water activity), and that lower water activity correlates with marginally better results across a number of metrics, you better look at that data.

The caveat for this entire study is that Cafe Imports is a high-end specialty-coffee importer. The people who sell coffee to us do a remarkable job. They are always striving to produce better coffee. Similarly, we are always striving to buy better, more stable coffee. It is certainly possible that our entire data set is shifted across a lower water activity range than the findings of others on account of the high regard for careful processing that our partners have. It is entirely possible that I am mistaken in believing that we see a representative sample of specialty coffee here at Cafe Imports and that in fact the sampling we see is significantly better-processed than the rest of the industry.

The critique implied in this paper is not limited to just some stated standard and it is not outward facing only. At least in the case of the 0.7000 Aw standard both the science and the reasoning

are sound. The critique is intended more broadly for the pattern in specialty coffee of imposing standards and of publishing papers and making presentations without significant observation. It is intended for those times when we allow our theories to come off like discoveries, or forbid, when we allow our statuses to burnish those theories so that they appear greater in the eyes of our audience. Finally, it is for those times when we allow our own hunger and curiosity for new theories to overrun the requirements of thorough investigation, sound methodology, and critical thinking.

In the last few years we have seen a growing pool of poor information on water activity. Poor information can be harmful and regressive when people act on it, in particular if they feel compelled to do so. I cannot say that my own early enthusiasm for water activity did not cross this line. If nothing more comes from this paper, our hope is that it at least becomes the prevailing perspective that a handful of samples and access to Google does not make a viable trial, let alone study. As our managing editor, Meister, says, "We need to remember that we are not scientists." I will only add that that does not mean that we should not act like them, only that we should do our best to restrain ourselves in the face of discovery.

All claims made in this paper are supported by our research and data. We make no recommendations that we have not already put into practice—at the risk of loss to ourselves as a company. Our work is ongoing. The huge number of variables to test and control in coffee require a substantial data set indeed to allow for detailed parsing and still be considered sufficiently thorough. In this vein, much more so than mere critique, we invite all interested readers to replicate and improve upon our findings.

Water Activity in Specialty Green Coffee.

Water activity theory has been covered extensively elsewhere and this information is readily available. In short, water activity describes the energy status of moisture in a system. It is this energy status and not the quantity of that moisture that determines the rate and likelihood of a large number of transformations and processes. For example, the growth of various molds is dependent on the correct water activity rather than on their being "enough" moisture present. This is a simplistic take, but sufficient as an introduction. More details of water activity theory will become clear as our discussion of water activity in green coffee progresses.

This is important because in order to talk about the use of water activity in specialty coffee, we need more than just a theoretical

understanding: We need to understand water activity in our specific product. Water activity is measured from 0 to 1. What are the Aw parameters for specialty coffee? How do you know? If you do not know, how can you talk about high or low Aw in coffee? We know that water activity applies to us generally because it applies to everything. How does water activity apply to us in particular in coffee?

In the six years that we've been measuring water activity at Cafe Imports, we've taken around 25,000 water activity measurements. Of those measurements, 10,000 have been longitudinal, meaning that they have been taken from the same coffees over time.

We take the larger set as a reasonable sample estimate of the population "specialty green coffee." We are not aware of a comparable sample size for water activity measurements in specialty coffee. This data set covers numerous harvests, origins and regions, varieties, processing methods, grades, preparations, altitudes, and other variables commonly found in specialty coffee.

The primary selection pressures on this data set are those exerted by what people have decided to sample to us and, by extension, the choices we make in weighting our purchases and inventory. Every sample that is sent to us gets cupped, and every sample that is sent to us gets its Aw measured. The low end score range of this larger, inclusive set is what most in specialty coffee would assess around 70 points.

A few notes on imbalances in this data set are in order. It is heavily weighted to Colombian coffees, with around 19% of the total representation. Nearly 66% of our sample volume consists of pre-shipment samples (PSS). Between 65 and 70 percent of the samples are Washed. If coffees from any heavily weighted category

have a unique water activity behavior or characteristic, this may skew the data. To check on the impact of these biases we'll compare category (origin, type, etc) samples against one another.

The longitudinal data set includes a full range of coffee grades and types while excluding as noise coffees that were not deemed approvable. The operative exclusion parameter here is contract fidelity meaning that coffee contracted to 83 points needed to cup at 83 points. This has the benefit of giving us a data set that looks at exactly what we're interested in seeing—approvable coffees and their volatility as relates to water activity. Unfortunately, this also introduces a potential selection bias that must be checked.

There are numerous considerations when setting up studies like this, just as there are when reading about them. Our primary interest has been in finding what will help us buy and sell better coffees. That we see relatively high sample volumes (approximately 5,500–6,000/yr) covering a very broad spectrum of coffee quality allows us some room to extrapolate our findings to the larger industry.

Methods and Materials

Samples and Handling

The samples used in this study come from 44 countries and include most commonly practiced processing techniques. We omit both Monsooned and decaffeinated coffees, as well as Robusta.

Samples are delivered to our office every day via delivery service. Shipping around the world through various channels likely introduces an unavoidable variable. We have made no attempt to either measure or mitigate this.

Samples are transferred from delivery bags to PET plastic storage tubes to await analysis, roasting and assessment. Low-priority samples can take as many as three or four weeks to reach the cupping table during the busy season. Most samples are cupped

within one or two days. Our average turnaround time for sample assessment is three to four days, including weekends and holidays.

Measurement and Preparation

All water activity measurements were taken on an AquaLab 4TE Duo Dew Point Moisture Analyzer. The 4TE Duo is a direct–measure type Aw meter that works by measuring the dew point temperature of a sample.

Moisture content was analyzed by a Sinar AP6060 moisture analyzer, a capacitance type instrument.

Samples were roasted one day prior to being cupped. Our sample roasts over the course of the study varied little, reaching first crack between 7:15 and 8:00 and concluding roughly 60 seconds after first crack had been established.

In 2015 we began using Stronghold S7 roasters for our sample roasting. These roasters have allowed us significant increases in the replicative accuracy of our sample roasts. S7s are hot-air roasters that allow the user to develop and save roast profiles for replication. The replication functionality of the S7 is highly accurate. Even with charge weights and temps poorly calibrated to the selected profile, the S7 roasters quickly find the line and then follow it. Since implementing the S7s, our sample roasts crack at 7:45 +/- 5 seconds and are completed at 8:45 +/- 5 seconds.

We us an Agtron M-Basic II to quantify roast degree. Agtrons "operate in the near–infrared energy band" to measure the reflectance of roasted whole bean and ground coffee.

Because of the S7's profile replication, we can be sure that differences in reflectance are produced from the same baseline

of temperature at time as a function of a coffee's response to said temperature at time, as opposed to personal decision making in the roasting process.

Grinding is done with a Mahlkönig EK43.

Brewing is done in identical cups using identical dosage of coffee and water. Water is heated uniformly in digital kettles to 95ºC. Crusts are broken at the same rate at which they were poured, beginning after 4 minutes. Cups are skimmed 3 minutes after being broken and are allowed to cool for 17 minutes (from pour) prior to tasting.

Water is filtered through a reverse-osmosis machine made by Global Customized Water. The filtered water is then remineralized to 130ppm +/- 10ppm with general and carbonate hardness solutions from Global Customized Water.

Assessors taste all cups in multiple rounds and various orders. Scoring is informed by explicit process-based standards (Washed, Natural, Brazilian/Indian, and Wet-Hulled) and done independently. These standards state the scoring ranges for various flavors and attributes as they derive from different coffee processing methods. All contributing scorers are experienced coffee tasters and have demonstrated the ability to identify and rate the scored attributes.

Samples are presented with four cups placed in front of a placard indicating the sample's code. Sample codes are random three digit numbers. Process standards are identified in the codes by the simple addition of a signifying letter (e.g. 386N indicates a Natural processed coffee).

Cuppers do not see roasted samples either before or during the cupping process.

A note on cupping as part of study methodology:

One of the biggest challenges in a study of this type, let alone in carrying out formal experimentation, is determining the measurements to be used and then verifying their accuracy. We can talk about the science of water activity and look at its relation to other instrumented tests. Eventually we want to know how it matters in the product itself. The way we find that out is by tasting coffee and recording the results: cupping.

Just as we cannot assume that every difference (change in score or flavor) that we register is due to water activity, we also cannot assume that a given difference is due to the coffee it's found in. We must be cognizant that cuppers and cupping itself are susceptible to error.

There is an enormous amount of work to be done in specialty-coffee sensory analysis. We can open this discussion here only so far as it allows us to describe the procedures that led to our results. Suffice it to say that one would be hard pressed to prove much on the basis of a cupping methodology that relies on subjective or poorly defined terms, or that fails to control for biasing factors and errors.

Be that as it may, we all currently use some recognizable form of the practice known as cupping. Generically, cupping is the process of tasting and assessing coffee samples on the basis of stated standards and frequently with use of some type of cupping form, or questionnaire. In any sensory analysis endeavor, before we ever get to the sensory, there are numerous variables and opportunities for error that we have to identify and control for. Many of these relate to sample preparation and presentation. We can improve our practice and therefore our data by addressing issues of preparation and presentation.

In the preparation and presentation of a cupping we aim to ensure as far as possible that differences perceived and recorded on the cupping table are attributable to the coffees and not their preparation or presentation.

On the one hand, we want to best be able to accurately identify differences, rate, and describe coffees. On the other, we want to remove doubt about the origin of those differences, ratings and descriptions. To the extent that doubt may be raised with regard to preparation or presentation being the origin of a perceived difference (etc.) in cupping, the decision power or relevance of the data itself is called into question.

Analysis

In the longitudinal data set, when an arrival coffee (ARR) is cupped, the differences from PSS for both cup score and water activity are calculated and attached as data to the PSS (e.g. a PSS may have a score diff of -2, which would indicate that the ARR of that coffee scored 2 points lower than the PSS). Similarly, when a coffee is cupped SPOT from our warehouse, the difference from the ARR scores for that coffee are calculated and attached to the ARR entry.

Another construct that we use pertains to labeling samples on the basis of their PSS. For example, PSS samples with high Aw are labeled with an H. This H stays with the lot for all future sampling, even if the Aw drops. The reason for this approach is that our driving interest has been to see if water activity measured at the PSS level can predict quality arrival and storage.

Regression analysis is used when appropriate.

Population Estimates

Given some background in water activity theory and specialty coffee, we can begin building a model for their interaction.
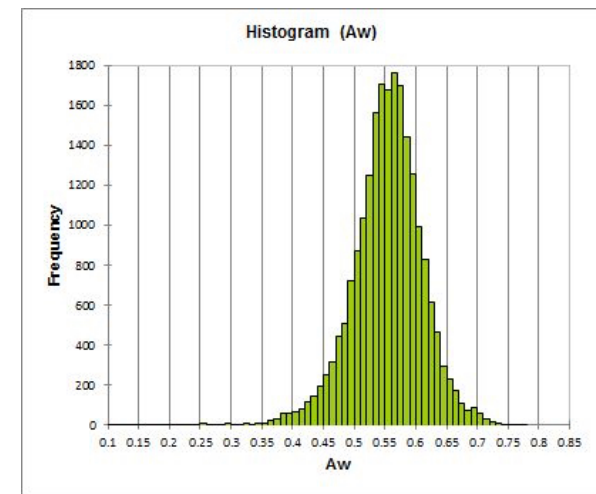
Variable	AW	Mean	0.554
Observations	21916	Stdev	0.057
Minimum	0.188	Median	0.556
Maximum	0.761	Mode	0.558

One way to look at a large amount of data is to create a histogram. The one above divides nearly 22,000 measurements into small groups, in this case 0.05 Aw, and then counts the frequency with which each group is observed. This can give us a broad picture of our entire data set, all at once. We add some basic statistics to this to help deepen our understanding.

The minimum and maximum are just the lowest and highest observed values. The mean value is found by adding all individual values and then dividing by the number of values added together. This is commonly called the average. The median is the value that occupies the exact middle location in the data set, and the mode is the most frequently occurring value.

Finally, the standard deviation (Stdev) is a measure of dispersion or spread. In distributions like ours that are normal or nearly normal, the standard deviation essentially tells us where most of our values are observed. A rule of thumb says that 68 percent of normal data will be within 1 standard deviation of the mean, 95 percent within 2 standard deviations, and 99.7 percent will occur within 3 standard deviations. This allows me at a glance to understand that coffees with water activity above 0.668 (2 standard deviations above the mean) or higher are very uncommon in this data set.

Specialty green coffee is distributed fairly normally through the middle of the 0–1 water activity range. Microbial proliferation doesn't much begin until 0.600 Aw. This is nearly a full standard deviation (0.611) above our observed sample mean (our estimate for the population mean). Given the near normal distribution (demonstrated below by the normal density curve overlayed on our population estimate histogram) of this population estimate, we know that as Aw increases towards the second standard deviation at 0.668, the frequency of observations decreases dramatically.
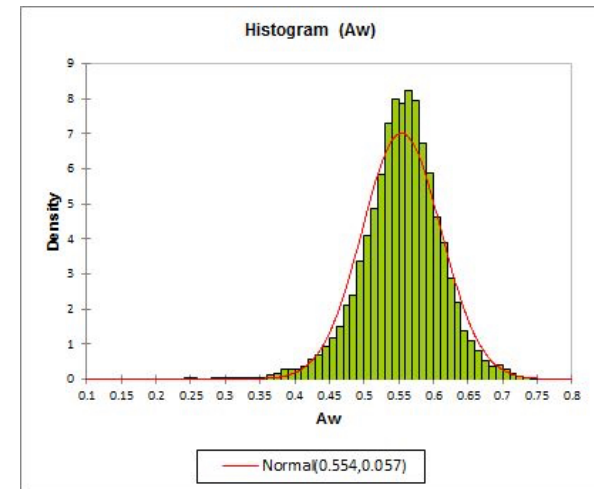
Water Activity and Moisture Content

Moisture Content: the amount of water in a system.

Water Activity: the energy status of the water in a system.

A coffee bean with some given moisture content can have a range of water activities. A coffee bean with some given water activity can have a range of moisture contents. The amount of moisture in a coffee bean is usually reported as a percentage by weight. This is just the amount of moisture in the bean. The water activity of that bean, or the energy status of the moisture in that bean, can vary depending on a number of variables (e.g. temperature and chemical composition). This being said, within a given product the relationship between moisture content and water activity is often fairly regular.

The relationship between moisture content and water activity is product-specific and follows what is known as a moisture sorption isotherm.
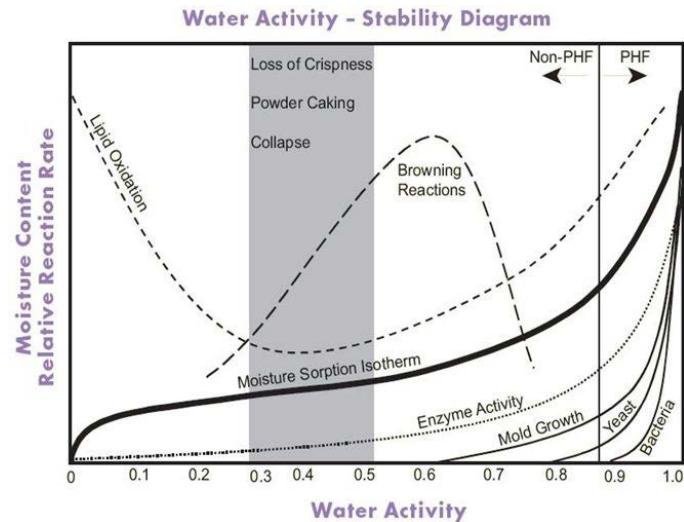

Despite our occasional descriptive flourishes, coffee is really coffee. In terms of water activity it seems that we can broadly take coffee, or at least specialty coffee, as a single product with a relatively narrow range of parameters. Raisin bran cereal is commonly used as an example to explain how water activity works. This is because the very different properties of each component along with the dramatic results (soggy flakes, hard raisins) of mismanagement are clear and relatable. For us we can say without a doubt that coffee is much more coffee than raisins will ever be bran flakes. In a raisin-and-bran-flake-system we have very different moisture levels that we want to maintain, as well as very different Aw values. These

different water activity values will actively change the localized system moisture levels from our targets. This is a good example of water activity causing adverse outcomes below most of the microbial thresholds.

If we blend two lots of coffee with different Aw levels, moisture will migrate from one lot to the other until the water activity comes to equilibrium. However, in most practical situations it is hard to imagine an impact in coffee as drastic as what we see with raisins and bran flakes.

Water activity equilibrium is accomplished by the migration of moisture. As the lower Aw coffee takes on moisture its water activity will increase. As the higher Aw coffee releases moisture its water activity will decrease. Note that Aw is not determined by the total quantity of water in a sample, but only by that which is least tightly bound. Within a given product, the amount of moisture that can be "bound" is limited. When we talk about the relationship between moisture and water activity in coffee, we are talking about a knowable relationship within a given product.

As you can see in the chart below, the observed correlation between Aw and moisture content in specialty coffee is not linear. In 20,993 samples the correlation between water activity and moisture content was 88.4 percent. The R^2 was 78.1 percent, indicating a less than perfect fit to the linear model shown below, but also indicating that 78 percent of the variability in Aw can be explained by the variability in moisture. The 95 percent confidence interval for the observations is just under 1 percent on either side of the MC (moisture content) model line. We will build on this later. For now suffice it to say that moisture content is neither a perfect nor a terrible predictor of water activity.
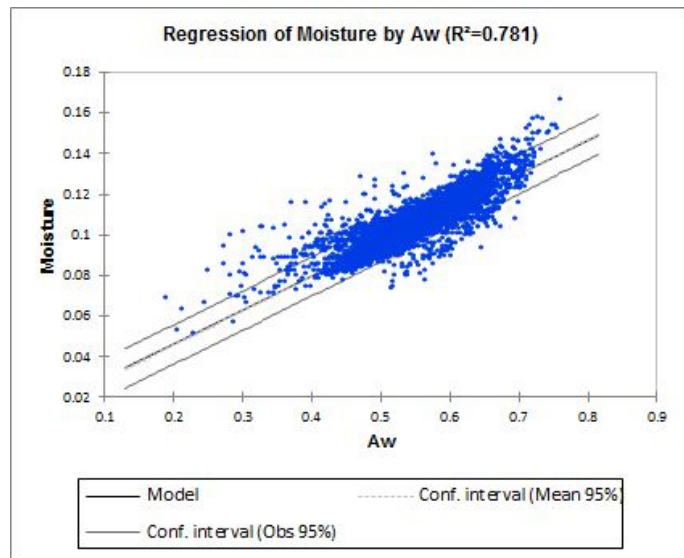
Look back at the Water Activity Stability Diagram.

The line labeled "Moisture Sorption Isotherm" is a generic line describing the most common shape of the relationship between moisture content and water activity in food products. This line is generic because different products will have different placements (note how Moisture Content is not delineated on the Y axis). You can see the approximation of the isotherm in our regression above, in particular with the concave form becoming apparent above 0.650.

For all the time we spend talking about how special and unique different coffees are, coffee itself is a fairly uniform substrate. This has implications for how we study coffee physically. It also makes sense. Humans look wildly different, unless you're a river otter. Then the gross genetic expression of Homo Sapiens Sapiens starts to stand out a little more. Bipedal. Lacking fur. Poor swimmers.

Moisture Content Revisited

If we cut off coffee purchases at 12% moisture, we will cut off many Aw-related problems. The model gives an error of about 1 percent. If we pull the criterion back to 11% the problem region for Aw in specialty green coffee is reduced.

This does not mean that all of these coffees arrive as they pre-shipped and then last forever. We have not addressed any problems arising from change in water activity post sampling. It just means that the majority of known Aw-caused problems are increasingly unlikely to apply because in green coffee processed for specialty export it is uncommon to find 11 percent moisture coupled with Aw greater than 0.610 (below which very little can be attributed to Aw). Remember, in water activity coffee is just coffee and moisture

is an imperfect-not-terrible predictor of water activity. The general uniformity of the coffee substrate, at least as far as moisture is concerned, means that moisture interacts with different coffees in much the same way.

> While people stress that it is water activity and not moisture content that determines the rate, direction and likelihood of a large number of transformations in a substrate, this does not mean that there is no relation between moisture and water activity. In general, in a given system, more moisture will mean higher water activity.

The likelihood of a higher than 0.610 Aw when using 11%–11.1% MC (moisture content) as a criterion is small. While this is still around 7 percent, the main point here is that we can limit our exposure to water activity related problems by anticipating the relationship between Aw and MC. Fifty-one of the 703 coffees shown below with 11%–11.1% MC had Aw higher than 0.610. Looking more closely at 11% moisture content samples we see the following:

Moisture	Aw	σ	Aw + σ	Aw + 2σ	Count of Aw
11%–11.5%	.5882	0.0232	0.6114	0.6346	3124
11%–11.1%	.5799	0.0235	0.6034	0.6269	703

This suggests that even if we miss our criterion and accept a coffee slightly above 11% moisture, that coffee will have to be more than a full standard deviation above the mean Aw for that specialty-coffee moisture level in order to be greater than 0.610 Aw.

Early on we hoped that water activity would allow us to learn to safely buy coffees with higher moisture content. Technically, this is the case. A coffee with 12.5% moisture and 0.550 Aw can be maintained as such with environmental controls. Such a coffee would not be subject to Aw-controlled degradations that occur above 0.550. Of course, a coffee with that level of moisture is very unlikely to have such a low Aw.

It is important to remember that, like moisture content, an Aw reading is just a single Aw reading. It does not tell you if the Aw is stable or moving. If moving, it does not tell you the direction. It does not tell you about either the past or the future. It does not tell you about the entire bag, let alone lot of coffee. A water activity measurement tells you the water activity of your sample at the time of measurement.

In our observations, we've found water activity to be more changeable than moisture. In drying trials that we performed in Costa Rica and Colombia, we found that moisture content declines relatively steadily from day to day, while water activity fluctuates as it declines. In the charts below the wildly fluctuating blue and gray lines are water activity measurements taken of different coffees. The steadier orange, yellow and green lines toward the bottom of the charts are the corresponding moisture contents.

Colombia: Aw and Moisture Through Drying Measured 2x per Day

Costa Rica: Aw and Moisture Through Drying Measured 3x per Day
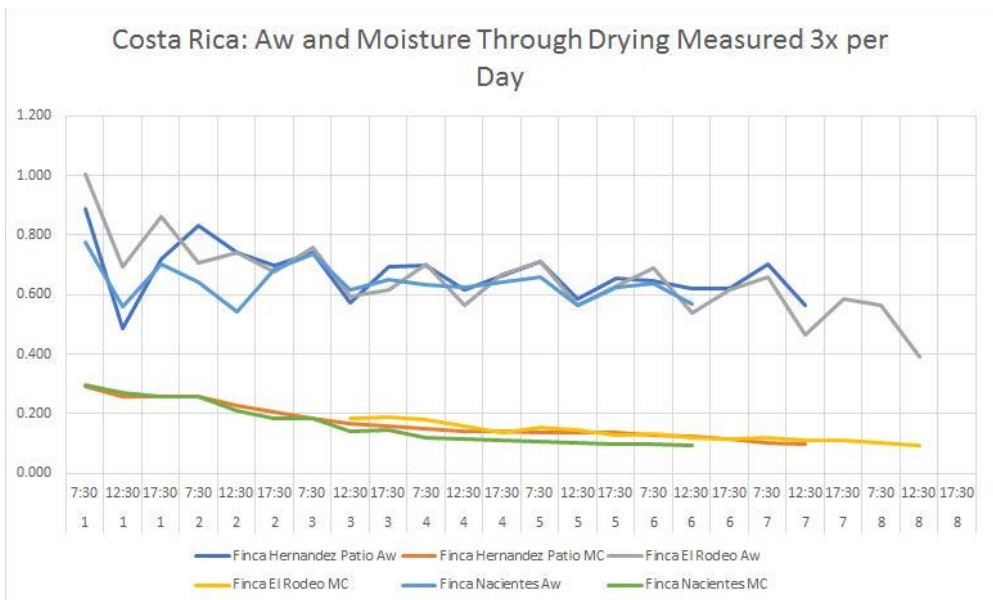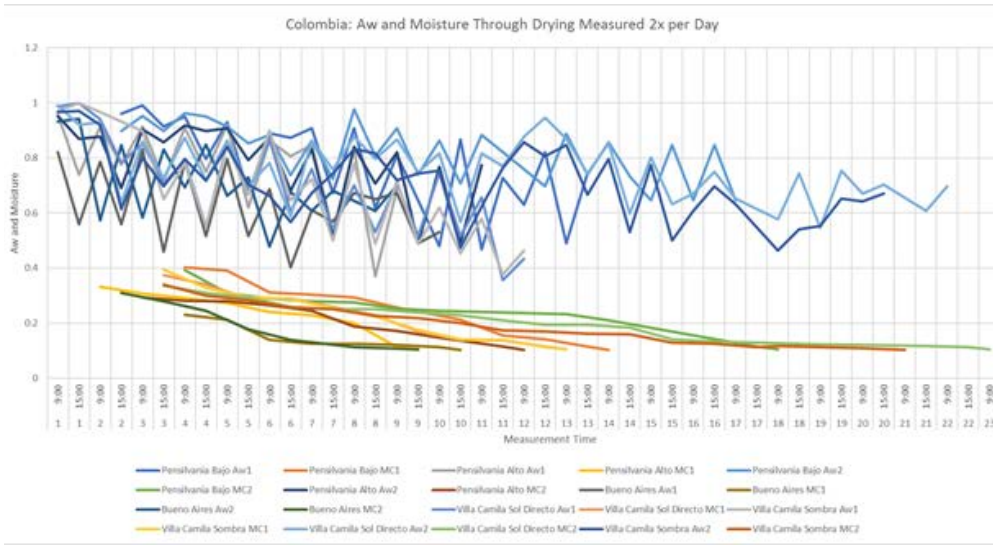
In addition to these, we were able to run a corresponding trial with coffee from Finca Hernandez in a mechanical dryer. We unfortunately missed a series of measurements due to events beyond our control (and in the process learned a little about arranging trials), but nevertheless did get enough readings at least to get a glimpse of Aw and moisture behavior as a coffee progressed through a mechanical drying process.

Costa Rica: Aw and Moisture Through Mechanical Drying Measured 1x per Hour
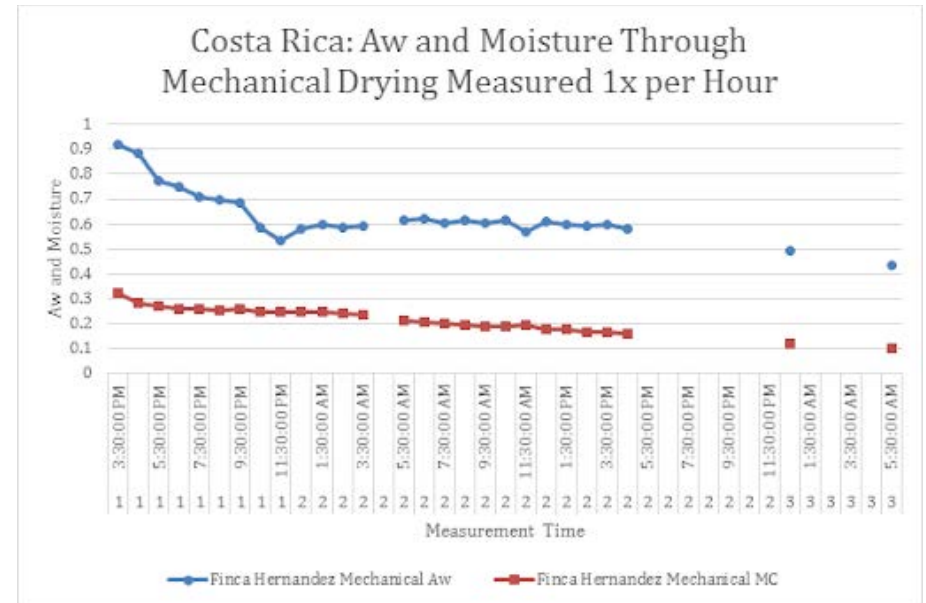
Even hourly measurements pulled from a mechanical dryer show fluctuations within the larger downward trend in water activity. We also measured relative humidity (RH) and temperature in the above drying trials. In general, the RH would increase overnight as the temperature dropped, and then would decrease as the temperature rose during the day. While the water activity of drying coffees did

decrease as the coffees dried, it tended to follow the same daily pattern as the relative humidity.

Today it is clear that what water activity has really done for us is reinforce and push back the criterion that needs to be drawn for moisture content. While the standard cutoff of 12% moisture is reasonable (actually, in terms of Aw it is a bit high around 0.624), it is water activity and not moisture content that we are often concerned with. In some ways the assumed use of moisture content is already as an estimate for water activity. That we may not have known that prior to measuring water activity does not negate the fact. We can remember as well that in some situations, moisture content is just exactly what we want—as when we want to know how much of the mass, and therefore price, of a given shipment is water.

We have seen that water activity is more reactive than moisture content to environmental cues. We will see below that this volatility carries through to the shipment period as deltas can be significant. Despite being more sensitive and capable of making close calls on upper end or very specific Aw levels (Aw measures Aw directly, thus a given Aw target can be aimed at and reached within a very close tolerance when compared to estimating that target with moisture content), and in particular given "black box" shipping periods of one to four months, water activity is probably *less well-suited* to making those close calls.

These conclusions about the relationship between moisture content and water activity in green coffee are further supported by analysis of the longitudinal data. We will dig more deeply into this data below. While there is an error in estimating water activity from moisture content, we will show that this error becomes much less significant when taken in the context of and compared to the changeability of water activity itself observed over the shipment period.

As an estimate of Aw in specialty green coffee, moisture content is not perfect. This means that if we determined that 0.610 was our Aw criterion, and that we wanted to use moisture to estimate it rather than Aw to measure it, we would need to do more than just find the correlation moisture level (~11.7%) for that water activity. Because the correlation is not perfect, a coffee measuring 11.7% moisture may have anywhere along a range of Aw readings (~0.610 +/- 0.028). Our job would be to determine the MC at which we are satisfactorily unlikely to go above 0.610 Aw.

There are two further topics that we should discuss before moving on to examine our longitudinal data for specialty coffee: browning reactions and lipid oxidation. Looking back to the Water Activity Stability Diagram, we see that the rates of both of these reactions increase as we leave the normal water activity range of specialty green coffee on the high end. Lipid oxidation also increases at extremely low Aw.

Aw and Browning Reactions in Specialty Green Coffee

It is easy to see the term "browning reactions" and assume that this is in reference to what's occurring in the coffee roaster. The Maillard reaction, for example, is a type of nonenzymatic browning reaction that occurs during coffee roasting. In many cases we (Homo sapiens sapiens, lacking fur, poor swimmers) very much enjoy the organoleptic results of the Maillard reaction.

However, the Maillard reaction can also be a problem with regard to food spoilage, as depicted in the Water Activity Stability Diagram. "Pokorny et al (1975) mentioned that the Maillard reaction can sometimes start in green coffees stored at relatively high

temperatures and humid medium, producing a browning reaction" (Flament, 38). The Maillard reaction and caramelization processes that occur during roasting occur under relatively [high heat and low water activity](#).

Browning Trials

We set up two trials, as well as analyzed our cumulative dataset to further investigate the role of Aw in the roasting process of specialty green coffee. In one trial we took samples of various coffees from the roaster at 290, 300 and 310 degrees Fahrenheit. Once cooled, we measured the Aw of these samples to see if the Aw of coffees just prior to Maillard in roasting significantly rearranged relative to their original Aw.

The second trial that we conducted dealt with manipulating the water activity of multiple samples of a single coffee. Here we were curious to see what might happen when the coffee variable was controlled and the Aw was changed.

We begin our look at water activity and browning in roasting by looking at correlation data between Agtron output and Aw input.

The density distribution of our ground Agtron scores is shown below with a normal curve overlay. The frequency distribution is also shown. The data is near normal with a very slight positive skew.
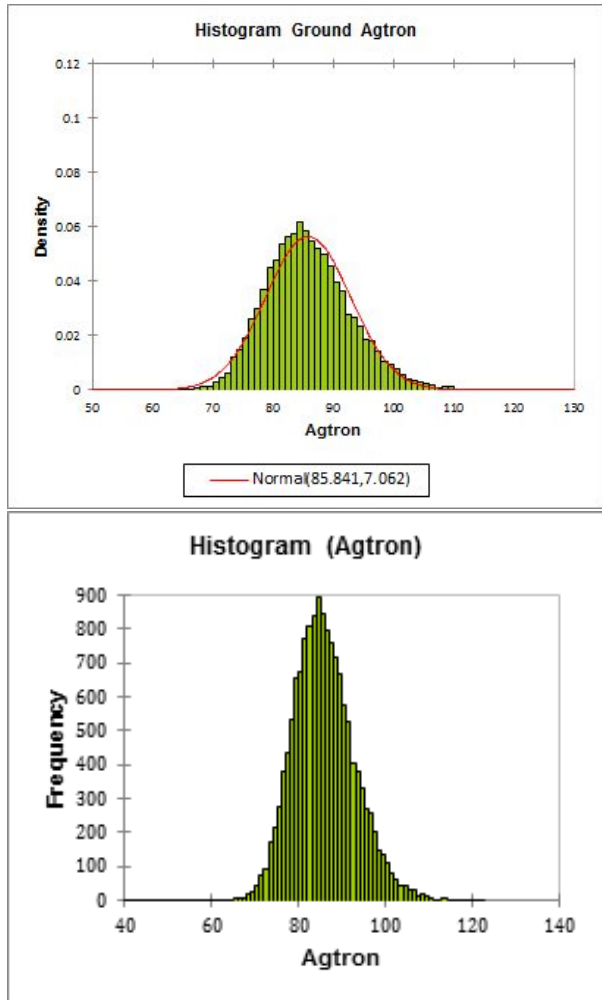
Finally, we look at the regression of Agtron by water activity. The correlation between water activity and Agtron is 23.5 percent. The positive correlation indicates that to a small extent as water activity increases, Agtron does too. Generically, this means that as water activity increases roasts become lighter.

The R^2 for the plot tells us that very little of the variability in Agtron (~ 5.5 percent) is explained by water activity. While there is a slight positive correlation, the linear model clearly is not a good fit for the data.
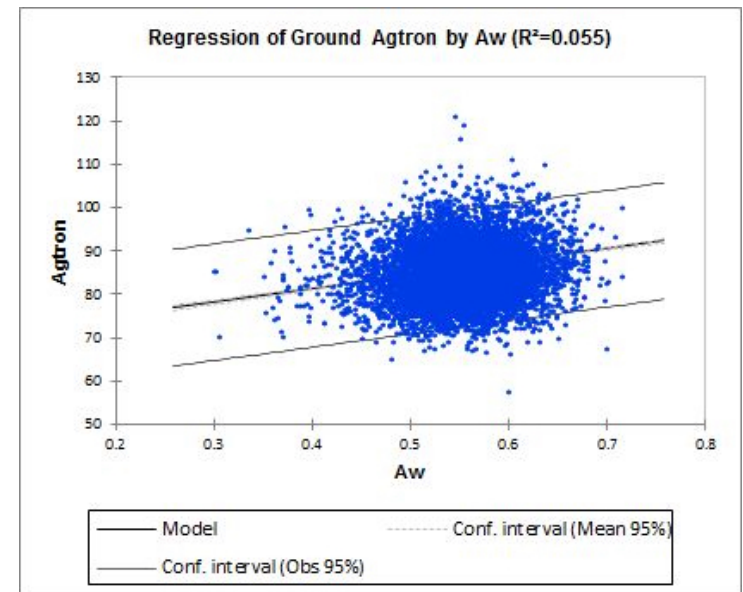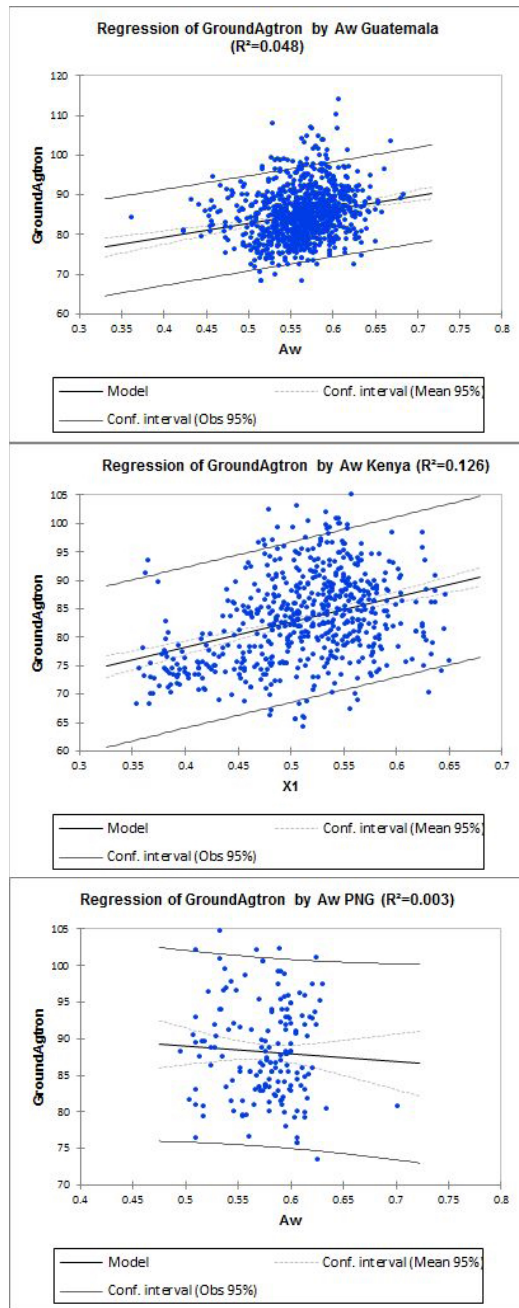

Looking at individual origins we see some variability, though overall things look much the same. Papua New Guinea is the only origin that we observe to have a negative relationship (and a very slight one at that). Even in the PNG plot, the dashed lines indicating the confidence interval of the mean are strongly curved. If you look

closely, all CIM lines are somewhat curved. These give us the range for the possible best fit (model) lines through the data.

Returning to the PNG plot, we can see that even though the model line tracks slightly negative, the plot is clearly random and the strong curve of the mean confidence bands allows for both positive and negative orientations of the model. Kenya is the only origin observed where Aw seems to explain Agtron by more than 10 percent (12.6 percent). Realistically, the Kenya plot appears to be quite random, with the stronger correlation and model fit being driven by a loose cluster of low Aw–low Agtron samples.

Regression of GroundAgtron by Aw Costa Rica ($R^2=0.066$)

Regression of GroundAgtron by Aw Ecuador ($R^2=0.037$)

Regression of GroundAgtron by Aw Brazil ($R^2=0.094$)

Regression of GroundAgtron by Aw Colombia ($R^2=0.042$)

Regression of GroundAgtron by Aw Ethiopia ($R^2=0.091$)

Regression of GroundAgtron by Aw Guatemala (R²=0.048)

Regression of GroundAgtron by Aw Kenya (R²=0.126)

Regression of GroundAgtron by Aw PNG (R²=0.003)

We can see in these charts that given very similar roast inputs across a wide range of coffee variables (origin, processing, variety, age), Agtron output and Aw input have very little relationship. While the scatters for each origin can occupy relatively unique X,Y zones in the plot, each of the Origin regressions overall looks similar to the original Regression of Ground Agtron by Aw.

We should keep in mind that this data is pulled from uniform time-temperature roast profiles. What this means is that the roasting machine's BTU application is merely sufficient to cause a given coffee to follow a predetermined temperature-at-time series. This does not test how coffees behave given the same roast input, but how they behave when required to reach the same roast temperatures at the same roast times. The automation and accuracy of the S7 means that we do not actually know how much more energy is required to keep these higher water activity coffees on the profile line.

Aw During Roast

We wanted to get a look at what happened to a coffee's water activity when subjected to large and rapid increases in temperature. If input water activity had an impact on nonenzymatic browning during the roasting process, whether that pattern was to increase either the degree or rate of browning or decrease it, we expected that we would see a pattern take shape across our data set. We have been unable to observe any significant pattern in roast degree or color output relative to Aw input when following very similar roast profiles.

Failing to find significant correlation between water activity in room-temperature coffee and its subsequent degree of browning under controlled roast conditions, we considered the possibility that the radical change from room temperature to hot roaster was such a shock to the coffee–Aw system that it caused a realignment in

water activity levels. Under this theory a "true" browning-inducing Aw level would be arrived at just prior to the beginning of roaster Maillard reactions. Because higher moisture tends to equate to higher water activity in specialty coffee it was considered very unlikely that this level would be approaching 0.600 or higher, in particular relative to green measurements. However, it was also considered possible that coffees that ultimately showed increased browning under controlled roast conditions would be the those that had higher Aw just as browning began.

Methods and Materials

The trial and therefore the data that I can present are imperfect. For one, we are unable to measure water activity while coffee is at roasting temperatures. Secondly, our Aw measurements were not taken after exactly the same duration after being removed from the roaster due to each reading taking longer than each roast. Finally, this trial looks at only 12 samples.

Samples were selected by convenience from a wide variety of origins with many water activity and moisture levels, processing methods, grade levels, and coffee ages represented.

We again used our Stronghold S7 Pro roasters on a set profile. For this trial we roasted full pounds of coffee. The S7 is capable of profile matching down to 150 grams, which is what we use for sample roasting. It performs best in its replication function above 300 grams and most easily allows for early tryer sampling above 450 grams.

We took three successive samples from each roast at 290 (143ºC), 300 (149ºC) and 310 (154ºC) degrees Fahrenheit. These temperatures were chosen to mark the commencement of the Maillard reaction. The trial roast drop temperature was 424F (218ºC).

Samples were collected, cooled, and stored in small, tightly closed sample cups until water activity measurement. Storage lasted between two hours and two days due to the necessities of our regular sample volume.
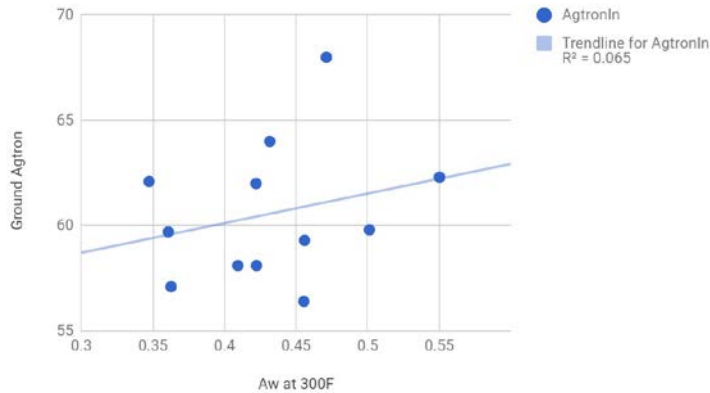
Results

As coffees move from room temperature (68ºF/20ºC) to the Maillard phase of the roast (290ºF) some Aw levels remain relatively flat while others drop significantly. The highest Aw sample, Colombia P10939, drops rapidly before leveling off and finishing in the middle of the group. The second-highest Aw sample, Sumatra P11086, maintains relatively high water activity through 310ºF before ending with the lowest of the group, coincidentally tied with Brazil P9849, the third-lowest initial Aw coffee.


Starting Aw vs Ground Agtron
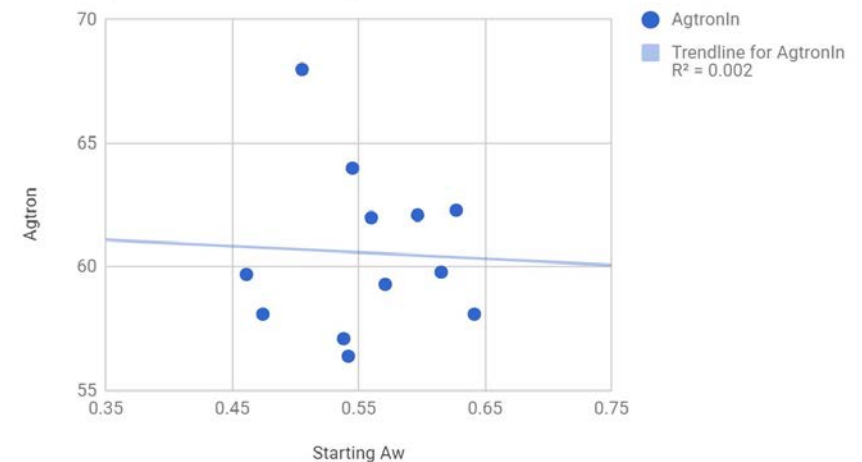
In this trial we did get a slight negative correlation between initial water activity and ground Agtron post roast. At -4.5 percent the correlation is minimal. The model fit is essentially nonexistent at .002.

Aw at 300F vs Ground Agtron

By 300ºF there is some rearrangement and positive correlation with Agtron, but at only 25.5 percent that correlation is hardly significant. Agtrons below 60 resulted from a large range of 300ºF water activity levels, as did Agtrons above 60. All Agtrons (save one with middling water activity at 300ºF) ended up within about an 8-point band, given the same roast profile.

Initial Aw and Aw at 300F
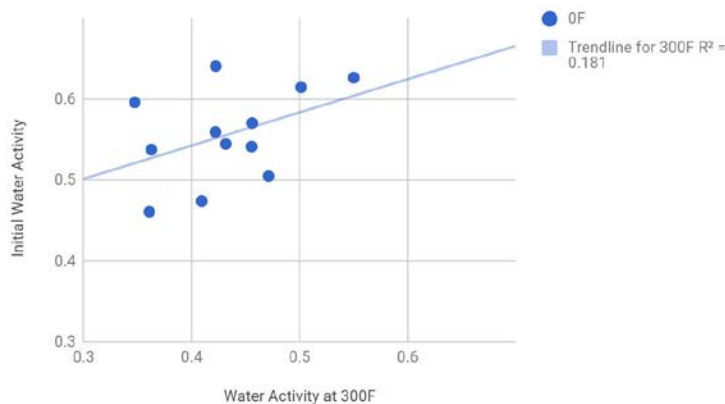
A coffee's initial water activity does not provide a strong indication of what will happen to water activity during the drying phase. Correlation between initial water activity and water activity measured on samples just prior to browning was still only 42.5 percent. Considering the above plot in quadrants, we can see instances of high initial Aw all along the 300ºF Aw axis. Similarly we can see high, middling, and low initial Aw yielding low 300ºF Aw. The only thing we do not see is low initial Aw with high 300ºF Aw.

The emptiness of this lower-right quadrant (low initial Aw, high 300ºF Aw) increases the overall correlation of the model, though this is perhaps misleading. By this point in the paper it should be clear that this lower-right quadrant is probably an impossibility. Whereas the model reads this absence as a lack of randomness, we can see that within the actual possible outcomes

(high-high, high-med, high-low, med-med, med-low, low-low), the plot is essentially random.

Stated differently: At a given initial water activity level (high, middling, or low), with only 12 samples, all water activity outcomes possible to that initial level were observed.

We should be reluctant to draw conclusions from this trial alone. There were only 12 samples in the trial, and our measurements were not taken during the roast but only after samples had been partially roasted and then cooled.

Both the lack of substantial rearrangement as well as the near-random decrease in the initial Aw of these observations does agree with our larger data set in suggesting that room-temperature Aw has little impact on browning produced by roasting.

Water Activity Manipulation Trial

To further test the impact of Aw on roasting coffee we set up a small trial in which we altered the Aw of a single coffee in various ways.

Methods and Materials

In this trial we used four conditions: Control, Bath, Heat, and Dry.

The Control condition was just the coffee itself.

For the Bath condition we put coffee samples in trays into loosely closed containers that had water in the bottom. The coffee samples were not allowed to contact the water.

For the Heat condition we used the same translucent plastic containers to house samples in the same trays and set them in a sunny window.

For the Dry condition we used the same setup as for the others, but this time used a layer of a desiccant (diatomaceous earth) in the bottom of the container.

Coffee samples were left in their respective conditions for one to two weeks.

Samples were then tested in six ways: moisture content, density, water activity, "constant roast profiling," whole bean Agtron, and ground Agtron.

Moisture, water activity, and Agtron have already been discussed. For density measure we use the same Sinar instrument as for moisture, along with the Sinar volumetric cylinder.

"Constant roast profiling" means that we set the inputs on our S7 roasters to the same level and ran the various samples with no changes. Rather than roast the samples on profiles, we tracked the time-temperature profiles that resulted for each sample with this constant heat application. With this we sought to observe how the manipulated coffees would respond to identical roast inputs. Would the higher Aw manipulation become darker? Would it progress through Maillard more quickly?

We began by testing 24 "base" samples of our coffee to establish a baseline for each test.

Each condition was then run with seven samples. We also created seven additional control samples, this time stacking them in the

same trays and containers as the others, but away from light and heat and without any manipulation.

The next step in this trial was the roasting of the coffee samples. Coffees were left in the S7 with no changes made for 9 minutes.

Results:

Brazil	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8
	MC	Aw	Density	Turn Temp.	Drop Temp.	Agtron out	Agtron In	Agtron Dif.
Control	10.4%	0.5436	68.0	126.2	187.4	75.7	91.4	15.7
Bath	13.4%	0.6636	67.2	124.4	181.6	90.6	113.5	22.9
Heat	10.6%	0.5421	68.2	128.7	190.6	71.4	85.0	13.6
Dry	9.7%	0.5107	67.7	126.8	186.8	74.2	91.7	17.5
Stdev Control	0.07%	0.0053	.28	1.50	2.13	4.08	6.16	3.74

The Baseline samples matched the Control group in our tests and we can look at the standard deviation of the Base/Control samples to help get a picture of how much each of the other treatments impacted the coffee. The Bath condition was by far the most impactful on the test coffee's moisture and water activity levels. We got slight decreases in both with the Dry condition. The Heat condition was possibly restricted by being run during a cloudy period. The Heat condition could be improved by the use of a heating pad or sunnier period of time. The Dry condition could be improved by the use of a stronger desiccant, or possibly by the use of a greater quantity of diatomaceous earth.
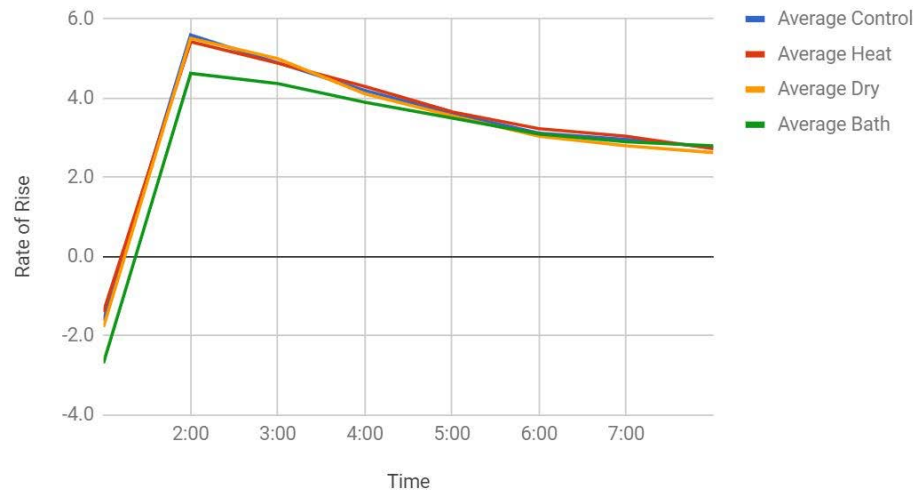
For the Bath condition we found that the rate of rise was lower than it was for the other samples (which all performed similarly, though the Heat condition did turn around more quickly and rise

just slightly faster through the Maillard phase of the roasts) for the first 6 minutes of the roast. For the 6th, 7th, and 8th minutes of roasting, the rate of rise of all samples was similar.

## Roast Profiles for Control, Heat, Dry and Bath Conditions
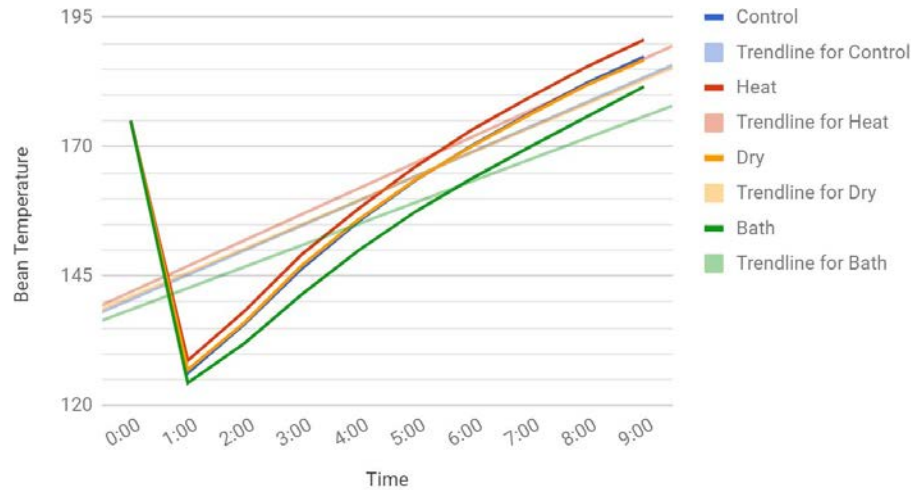
Rate of Rise at Constant Temperature Input

We found that not only was the rate of rise throughout the Bath roast slower than in the other conditions, but the Bath samples specifically were about a minute slower through the nominal Maillard phase (from around 135o–175oC) than the other samples.

In this case, average Bath condition Aw was 0.6636 (0.6850–0.6406). This should have been right in the prime zone for either increasing the rate or the degree of nonenzymatic browning. It could be that the corresponding increase in moisture slowed the browning progression, and that the Aw was coincidence. In order to test for this we would need a way to increase Aw independently of moisture.

Whole Bean and Ground Agtron
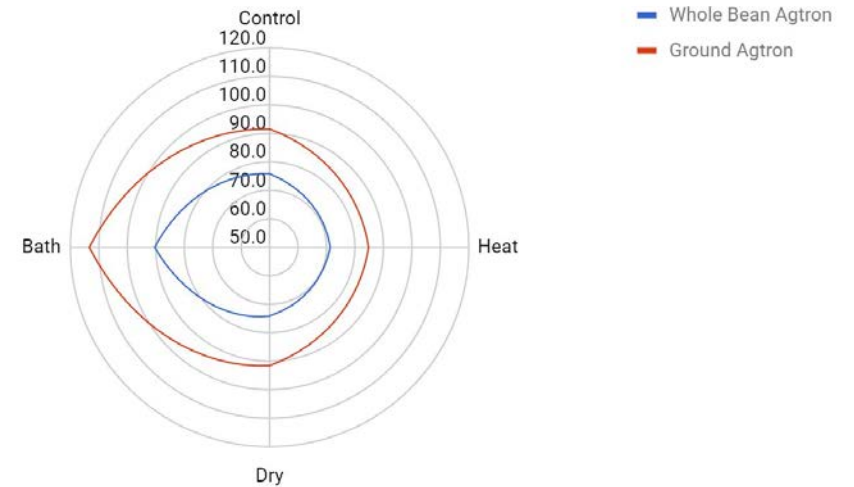
The total browning of the Bath samples was considerably less than in the other conditions. The Bath samples were less browned than the others, both when measured whole bean and when measured ground. The Bath samples also had the greatest

difference between whole bean and ground measurements. The rate of rise and temperature charts show the higher moisture and Aw samples maintaining a cooler roasting environment, despite being subjected to the same inputs. This could be due to an evaporative effect.

The lighter roast output for both Bath Agtron scores, along with the significantly larger gap between those scores for the Bath samples points to a change in the larger roasting dynamic of these samples. The Bath samples roasted differently. The most likely explanation is that there was enough moisture present to impact the roasting environment, as seen above. However, we cannot rule out the possibility of the higher Aw having a more basic impact on the coffee itself. Reaction rates for things like oxidation and enzymatic activity increase at higher Aw values in storage. These could directly act on roast browning compounds in coffee while in storage. Further and more complex testing would be required to explore this hypothesis.

Given the same roaster *input* conditions, it appears that higher Aw and moisture coffees may brown more slowly and to a lesser degree. Given the same roaster *profile* conditions, it is difficult to see a pattern. What pattern does appear seems to suggest that coffees beginning with higher Aw and moisture may resist browning. Looking at both, it may be that these coffees tend to require more energy to achieve the same degree of browning as lower Aw and moisture coffees.

Aw and Lipid Oxidation in Specialty Green Coffee

Lipid oxidation is important in green coffee because the oxidation of lipids can produce off flavors. A common compound in this category is called Trans-2-nonenal (T2N). T2N is produced by the autoxidation of linoleic acid, the "most important acid in the lipids in coffee" (Flament 118). Trans-2-nonenal is a very potent odorant that is described as smelling like paper, wet cardboard, cedar, or wood. At high concentrations it can smell like cucumber or green melon.

We see in the Water Activity Stability Diagram that the rate of lipid oxidation increases in a distorted U shape above and below roughly Aw 0.400. Keep in mind that this diagram is generically applicable to all products and is not merely specific to green coffee. In the table below we use our longitudinal data set to investigate the correspondence between water activity and Trans-2-nonenal type flavor descriptors.

		Cedar		Wood		Cardboard		Age		Paper	
Sub Pop.	Count	Count	%	Count	%	Count	%	Count	%	Count	%
Pop.	4055	554	13.7%	256	6.3%	50	1.2%	582	14.4%	419	10.3%
<0.550	1744	127	7.3%	55	3.2%	13	0.7%	111	6.4%	101	5.8%
<0.610	3522	369	10.5%	177	5.0%	37	1.1%	393	11.2%	331	9.4%
>0.610	533	185	34.7%	79	14.8%	13	2.4%	189	35.5%	88	16.5%
ARR if PSS > 0.610	194	89	45.9%	42	21.6%	8	4.1%	81	41.8%	42	21.6%
ARR if PSS < 0.550	452	55	12.2%	23	5.1%	7	1.5%	38	8.4%	42	9.3%

We see much higher occurrence of Trans-2-nonenal descriptors in coffees that have water activities above 0.610. We also see a much higher occurrence of these flavors if we limit ourselves only to the arrivals of coffees beginning with water activities above 0.610.

It is important to note that we also have the common Trans-2-nonenal flavor descriptors in coffees that do not have high Aw and that never did. Water activity does not explain everything. Water Activity is not the only factor at play in green coffee. Additionally, we should keep in mind that our findings are observational. We are not measuring, let alone verifying Trans-2-nonenal content. Trans-2-nonenal is detectable at low concentrations with certain aromatic characteristics. The identification of these characteristics is not necessarily the identification of Trans-2-nonenal .

We have observed a higher occurrence of Trans-2-nonenal type flavors in high Aw coffee. Trans-2-nonenal is formed by the autoxidation of a primary coffee lipid, linoleic acid. The rate of lipid oxidation increases at Aw above 0.450. We have observed dramatic increases in these flavors in coffees that are or that have

been measured at Aw higher than 0.610 and moderate increases in occurrence from 0.550 to 0.610. Coffees below 0.550 Aw have the lowest occurrence of Trans-2-nonenal type flavors. We will revisit these numbers with reference to moisture content later in the paper.

While this does not mean that the lipids in coffees with Aw above 0.610 have oxidized, the conceptual path from high Aw to Trans-2-nonenal type aromas in coffee is easily followed. A more conclusive study would require the direct measurement of Trans-2-nonenal (and related compounds) in relation to water activity and sensory testing. The cumulative observations of our study seem to support this transformation occurring in specialty coffee and suggest further, direct research into lipid oxidation.

Aw in Specialty Green Coffee Longitudinal Shelf Stability Study

Finally, we come to the real question that we set out to answer way back when we started measuring water activity. Does Aw reasonably predict shelf life (storage and shipment stability)? This question itself is poorly framed. We mean to ask: Does Aw measured remotely on a sample pulled from a larger lot of specialty green coffee reasonably predict the post-shipment arrival and storage holding characteristics of that coffee's cup score?

We begin by looking at the same broad water activity categories as we did for the Trans-2-nonenal query. In this case we substitute average and absolute score change as the target descriptors. We

will look at these both from the perspective of the life of the coffee (PSS > ARR > SPOT1 > SPOT2 > etc.), as well as just zeroing in on the change from PSS to ARR. Average score change (Δ) can give us a general impression of how a large category performs overall. Absolute score change takes all change as positive, ignoring the difference between positive and negative (coffees that gained points and coffees that lost points). This tells us about the degree of change within a category.

	Average Score Change (Δ)	Absolute Score Change	Count
Population	-1.27	1.93	3900
< 0.550	-1.05	1.74	1743
< 0.610	-1.22	1.87	3486
>0.610	-1.67	2.40	414
	Average Score Change	**Absolute Score Change**	**Count**
ARR if PSS > 0.610	-1.53	2.16	149
ARR if PSS < 0.550	-0.96	1.64	443

We can see that samples below 0.610 Aw experience less overall score drop, as well as less general volatility. The coffees above 0.610 lost more points on average and experienced more volatility than the others. If you remember, 0.610 was roughly one standard deviation above the mean (0.554) for our population.

The distribution of our longitudinal set (that is, coffees for which we have multiple measurements taken at time intervals) is very similar to our population estimate. There is a very slight negative skew that may be the result of increasing selection bias for lower Aw coffees as our work with water activity progressed.

If we limit the search to look at the score change that occured only between PSSs and their ARRs, we see a similar pattern. The coffees that began with lower Aw lost fewer points on average and also were less volatile. One more statistic that we can look at for volatility is standard deviation of score change. This gives us a view of how diverse the score change values are around our mean score change.

For the coffees beginning above 0.610 Aw the standard deviation of score change is 2.66 points while for those beginning below 0.550 Aw it is 1.99 points. This means (score change is normally distributed) that roughly 68 percent of the samples > 0.610 Aw have score changes between +0.99 and -4.33 points. Roughly 68 percent of the samples < 0.550 Aw have score changes between +0.94 and -3.04 points. Both average score drop and score change volatility are greater for coffees that begin with water activity above 0.610.

PSS to Arrival

Looking at just the score change from PSS to ARR, coffees beginning above 0.610 drop an average of 1.53 points with a standard deviation of change (σ of Δ) of 2.26 compared to 0.96 and 1.89 for those beginning below 0.550. Around 68 percent of coffees beginning above 0.610 Aw arrived between +.73 and -3.79, while those beginning below 0.550 arrived between +.93 and -2.85 points off their PSSs.
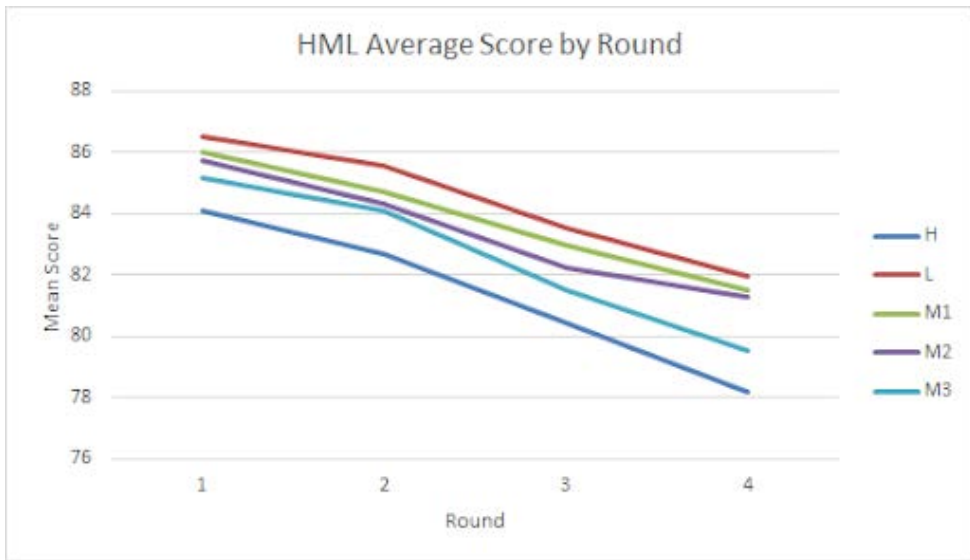
We observe a half point difference in change, on average. The larger standard deviation of score change for the higher Aw samples can magnify the half point average difference in change.

Aw	PSS	Δ	σ of Δ	Δ+	Δ-	ARR	ARR+	ARR-
< 0.550	86.21	-0.96	1.89	0.93	-2.85	85.25	87.14	83.36
> 0.610	84.04	-1.53	2.26	0.73	-3.79	82.51	84.77	80.25

All Lots

We can take a step back to look at water activity and overall shelf life from a wider perspective. We use a few labeling systems in our longitudinal data set. One of those is for the round in which a coffee is cupped, and another serves to categorize the PSS water activity level. Round 1 is always the PSS, round 2 is the ARR, and rounds 3 and 4 are subsequent SPOT cuppings, spaced by about three months. The three-month spacing for SPOT cuppings is on the one hand a matter of anecdotal periods of change, and on the other tends to mirror the initial shipment period. With this system we'll see a single coffee four times in a year. The PSS water activity categories (HML for short) are shown in the table below.

HML Average Score by Round

Tier	1	2	3	4	5
Score Range	90–100	87–90	84–87	81–84	<81

HML Aw Average Score by Round by Tier

HML	L	M1	M2	M3	H
Aw Range	< .400	.401–.560	.561–.590	.591–.610	> .611
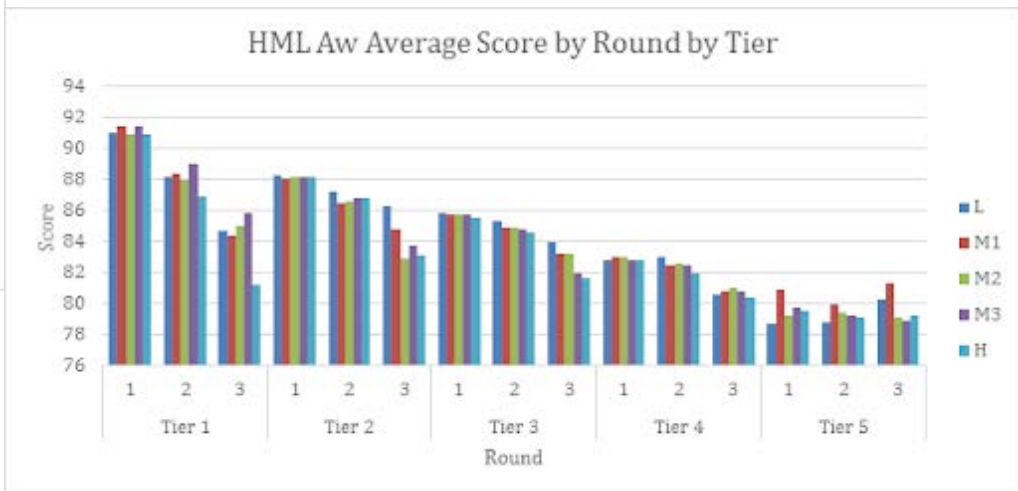
While it is difficult to see at this scale, the primary difference between the water activity groups as shown here are the minor alterations in slope from one line to the next. Note how the spacing from L and M1 to M3 and H increases slightly with each round (M3 actually arrives as well as L, losing around 1.1 points, before declining more rapidly).

We can break the above chart into subgroups by initial score tier. Because relatively few coffees make it all the way to Round 4 cuppings, the extra division of the data into score tiers means that we lack sufficient data to include Round 4 in the analysis below. We again see that the lower Aw samples tend to have more shallow slopes.
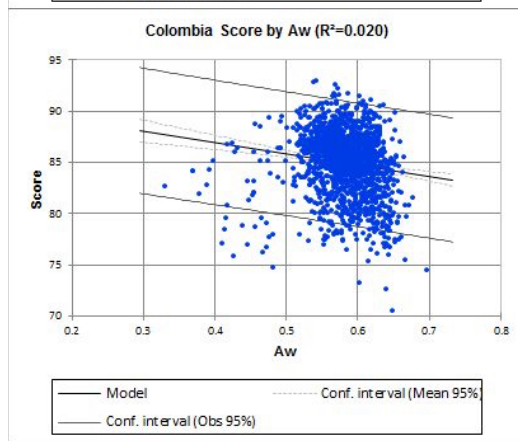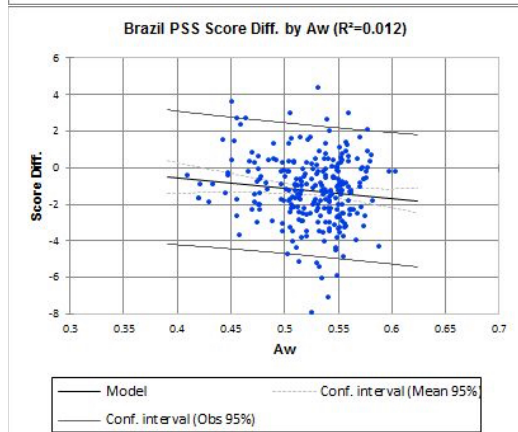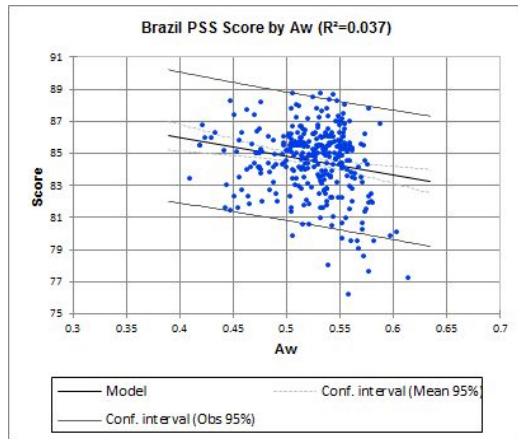
Origins

The basic observed relationship between score and Aw is slightly negative, as is that between score change and Aw. Neither is strong. While we do observe lower volatility with lower Aw coffees, it is important to get a larger look at the data that yields those averages. We can do this with scatterplots. Here we see plots of all scores against Aw, all score changes against Aw, pre-shipment scores against Aw, and pre-shipment score changes against Aw.
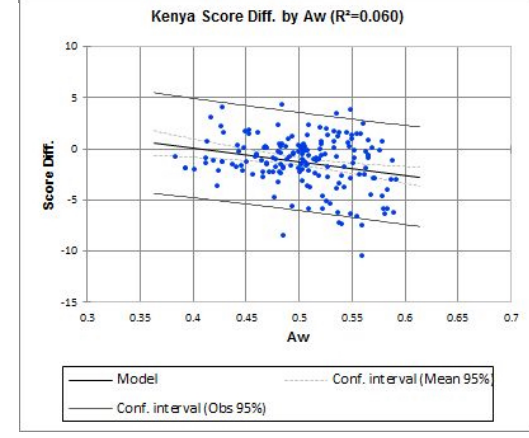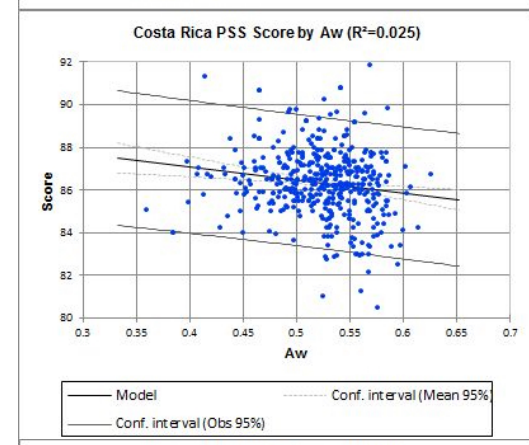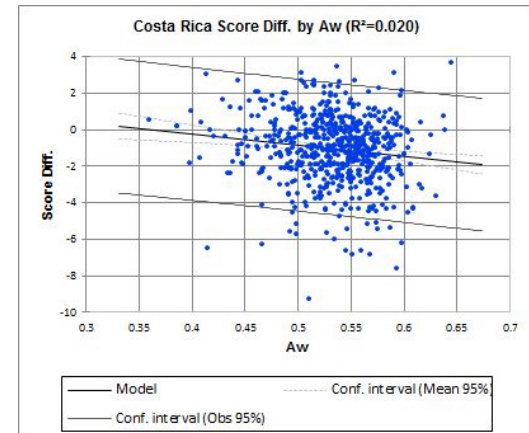

Regression of Score by Aw (R²=0.086)

Regression of Score Diff. by Aw (R²=0.009)

Regression of PSS Score by Aw (R²=0.080)
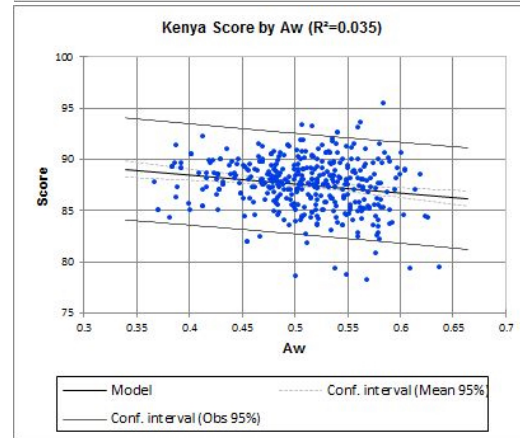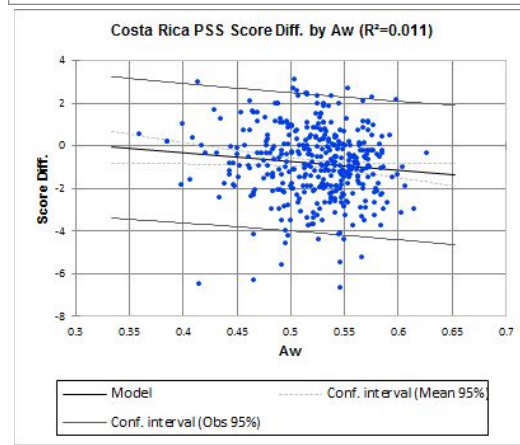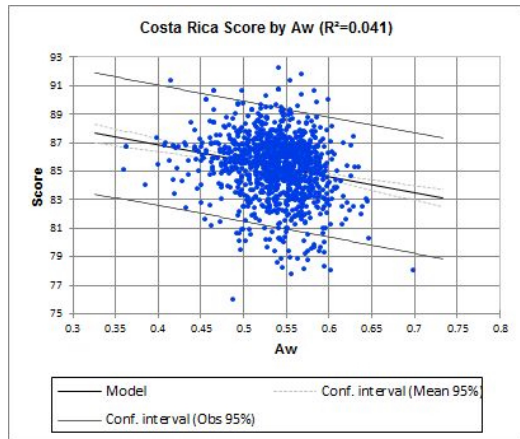

Regression of PSS Score Diff. by Aw (R²=0.013)
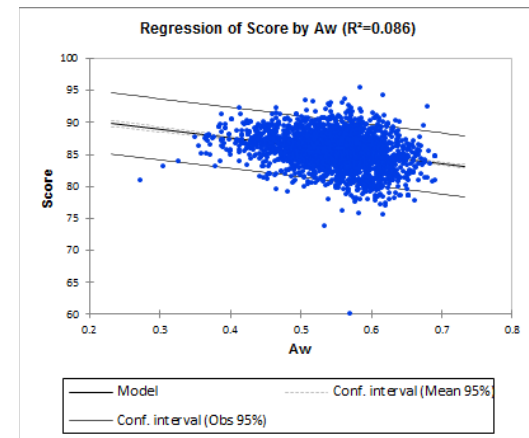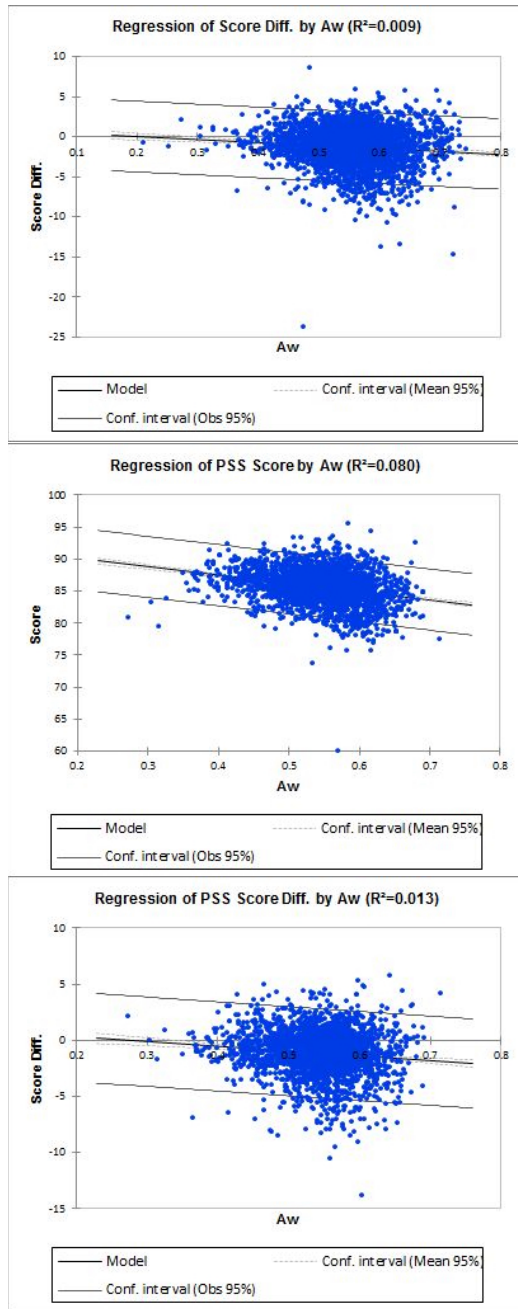
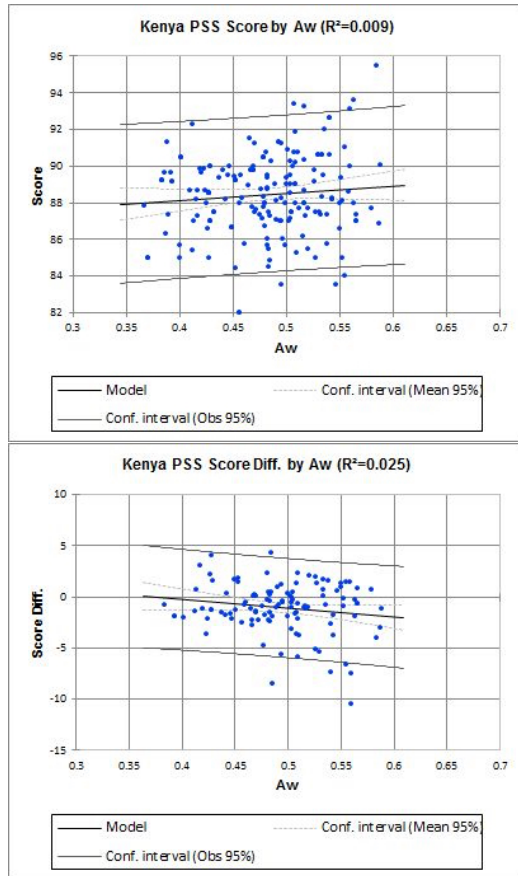Broken down by origin, and within origin by PSS-ARR, the (X,Y) placement can shift, but the overall patterns do not.

Brazil Score by Aw (R²=0.055)

Brazil Score Diff. by Aw (R²=0.007)

Brazil PSS Score by Aw (R²=0.037)

Colombia Score Diff. by Aw (R²=0.010)

Brazil PSS Score Diff. by Aw (R²=0.012)

ColombiaPSS Score by Aw (R²=0.028)

Colombia Score by Aw (R²=0.020)

Colombia PSS Score Diff. by Aw (R²=0.007)

Costa Rica Score by Aw (R²=0.041)

Costa Rica Score Diff. by Aw (R²=0.020)

Costa Rica PSS Score Diff. by Aw (R²=0.011)

Costa Rica PSS Score by Aw (R²=0.025)

Kenya Score by Aw (R²=0.035)

Kenya Score Diff. by Aw (R²=0.060)

Kenya PSS Score by Aw (R²=0.009)

Kenya PSS Score Diff. by Aw (R²=0.025)

Aw Over Time

We can also look at how water activity itself behaves over time.

HML Average Aw by Round

Here we see that there is a tendency for coffees to condense between around 0.525 and 0.550, given enough time. This specific range is likely a function of our warehouse conditions more than anything intrinsic to coffee in general.

Averages over large sample sets can be deceiving. Given the above chart, we might conclude that only coffees that PSS above Aw 0.610 (H) need be of concern.

Round	Aw	Ave	σ	Δ	σ of Δ	est. ARR	ARR +σ	ARR -σ	ARR +2σ	ARR -2σ
1	H	0.663	0.020	-0.021	0.050	0.611	0.661	0.561	0.711	0.511
1	M3	0.599	0.008	-0.014	0.047	0.585	0.632	0.538	0.679	0.491
1	M2	0.574	0.009	0.003	0.038	0.577	0.615	0.539	0.653	0.501
1	M1	0.511	0.037	0.033	0.044	0.544	0.588	0.500	0.632	0.456
1	L	0.375	0.026	0.065	0.084	0.440	0.524	0.356	0.608	0.272
2	H	0.622								
2	M3	0.590								
2	M2	0.576								
2	M1	0.542								
2	L	0.551								

The standard deviation of change in water activity (σ of Δ) is quite high in general. If our 1 standard deviation range for Aw in green coffee is from 0.497 Aw to 0.611 Aw, we have a primary working range of 0.114 Aw. The average standard deviation of change (σ of Δ) from PSS to ARR is 0.053, or about 46 percent of the primary range of water activity values that we're working with. When we look at a statistic like average change we want to check and see if

it is composed of many uniform changes or if it is composed of a more diverse collection of changes that cancel each other out. Even though the mean outcome may be the same in both scenarios, the specific volatility may yield insights that a simple mean can miss.

Estimates of ARR Aw based on average change and outside average change (one standard deviation) show that the water activity measurements that we're looking at are very changeable. These estimates for PSS arrival can be checked against the actual values shown for Round 2 in the bottom half of the table (highlighted). The L category PSS average Aw is skewed low due to a handful of very low Aw samples. It is notable that the L arrival Aw level is higher than the M1.

One reason that this is important is that the ill effects that we observe in green coffee that are due to water activity are not reversible. Once lipids have oxidized, they do not un-oxidize just by bringing water activity back into an ideal zone. Molds may be re-inhibited by reductions in water activity, but they are neither killed nor removed by Aw reduction. Just because a PSS had water activity in an ideal zone does not mean that the coffee the sample represents will remain there.

A second reason that this is important is that it undermines the value of a single point water activity measurement taken of a sample sent to Minneapolis from Colombia, for example. In the long run, we've seen that lower Aw coffees have performed better in some ways and to a certain degree, but this does not mean that any particular low-Aw pre-shipment coffee will perform as modeled.
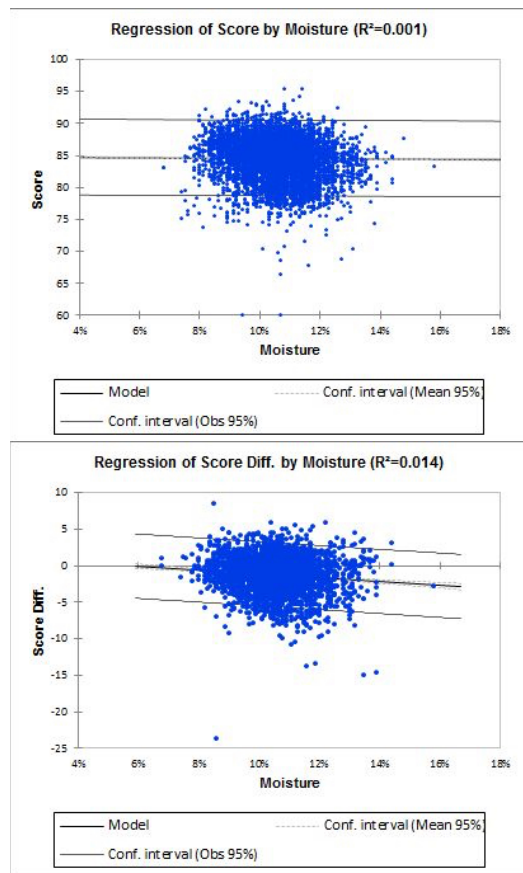
Moisture Content, Revisited

Returning to moisture content, we can run MC through some of the same tests.
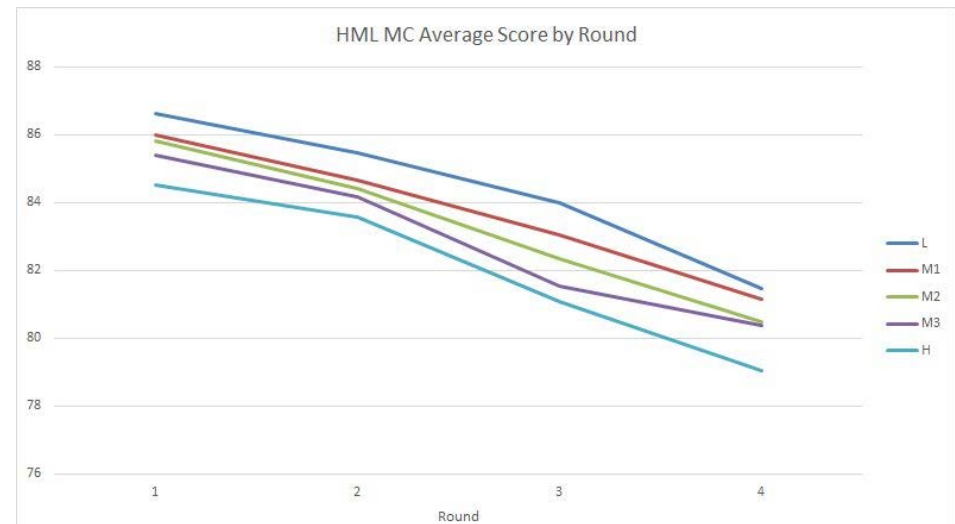
This is a very similar pattern to the Aw Average Score by Round chart. Here MC ranges are selected to correspond to Aw (H>11.6%; M3 10.8%–11.5%; M2 10.2%–10.7%; M1 9.5%–10.1%; L <9.4%). There are a few sharper angles, but overall we see the same hierarchy and general pattern.

The regression of score by moisture is flatter than that by water activity, while the regressions of score change by water activity and moisture content are very similar.
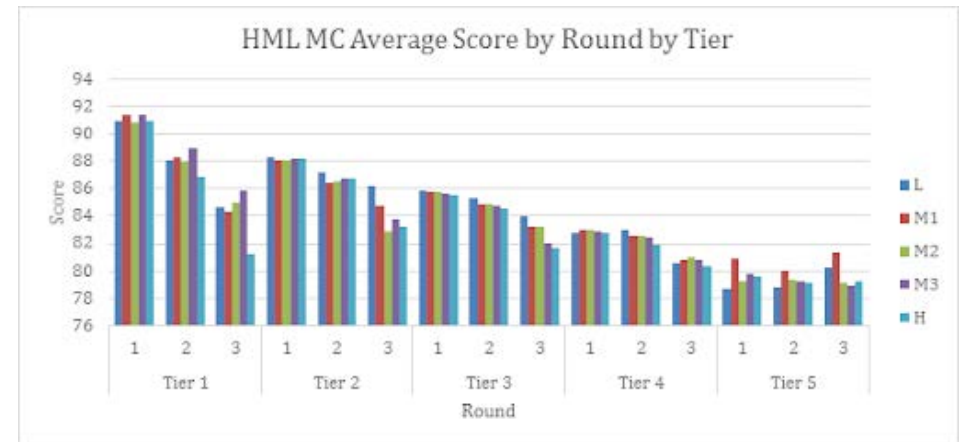
Here again the overall pattern is very similar to Aw. This is not surprising given the correlation between Aw and MC. In general, if there is more water in a system, the water activity will be higher.

Conclusion

There is a lot of excitement around water activity and, as we have seen, there are uses for the measurement in specialty coffee.

We began our study by seeking a novel application for water activity in specialty green coffee. Specifically, we sought to predict and account for the shelf life of highly nuanced coffee flavors on the basis of water activity's known application to general food-quality preservation. Over the course of our work we gathered and organized a very large amount of data and learned much about the uses and misuses of Aw. This paper has been a summary of those findings most directly related to water activity.

While we did not find our novel application, we also did not proceed so haphazardly as to either come up empty handed or, I

hope, to mislead people about the practical efficacy of water activity. As shown, there are regions of use for water activity in specialty green coffee and there are also regions that warrant further study. At Cafe Imports we continue to take water activity measurements, deepening our study, and we continue to consider Aw as one part of our larger green assessment protocol. We do this on the basis of the findings described in this paper.

A Maillard spoilage reaction may occur in green coffee with elevated water activity under hot or humid storage conditions. Testing could be arranged to detail the specific parameters for this reaction. The normally recommended storage conditions (cool, clean, and dry) for green coffee should be sufficient to ward off this problem.

One notable use for Aw has been in predicting off flavors likely due to lipid oxidation. We have observed correlation between the sensory perception of compounds produced by the oxidation of coffee lipids and high water activity in green coffee. These observations suggest further research into the coffee lipid oxidation process in higher end specialty level coffee. Because water activity is concerned with a system's interaction with its environment, these observations also suggest the importance of maintaining a controlled storage facility, and generally one that is below 60 percent RH and 65º or 70ºF.

Another more general application comes in the observation of marginally lower score change and volatility for coffees with lower water activity.

Many of the primary problems associated with high water activity do not impact us in specialty coffee. Most microbial activity is already inhibited at the Aw levels at which coffee is normally sampled and shipped. In theory, non-enzymatic browning of shipment and

storage coffee could occur at high shipment or storage temperatures in samples with an elevated water activity level. However, as it stands right now, there is no urgent need to adopt water activity across the industry.

One major challenge both to specialty coffee and to our desire to use water activity is the shipment period. While we very rarely see the upper and lower quarters of the water activity range (<0.250, >0.750) in specialty coffee, this does not mean that coffees subjected to extremes of heat or humidity (or drastic changes in either) during milling or shipping do not experience them. It only means that we do not observe them. There can be as much as four months time between pre-shipment sample (PSS) and arrival (ARR). The water activity of a PSS can only go so far in ensuring safe arrival. While water activity monitoring may be untenable during shipment, an understanding of the relationship between coffee Aw and environmental temperature and humidity may be used to inform the development of a best practice system for coffee shipment. This point warrants further investigation.

Our observations suggest that 12% moisture content in specialty green coffee may be too high of a criterion to reliably avoid water-activity related problems. Something between 11% and 11.5% would be a better generic criterion for the avoidance of water-activity related problems. It must be remembered that the water activity of a coffee is a single variable among many and that it can change over time. A change in one of these variables can result in a change in water activity, just as a change in water activity can result in a change in another variable. The real basis for water activity's use is in anticipating how these changes may occur given particular environmental conditions. As such, lower Aw is not inherently better than higher Aw. Lower Aw is better only when given a particular

or likely array of conditions (e.g. temperature and RH that are not at odds with a given Aw) and a certain set of goals (e.g. insulation from mold and lipid oxidation).

The acquisition and use of scientific instruments is not, in itself, science. Any presentation of a new process or device in specialty coffee should be accompanied by thorough research and data. Let us be done with the days of publishing presentation papers on the basis of a neat idea, a handful of data points, and access to Google.

Water activity meters are no different. The decision to utilize a water activity meter should be entered into with an understanding of what it can accomplish, in particular because of the expense of the instrument and the complexity of the science. Our hope is that this paper will aid people to this end. Given the current expense to own and operate them, along with current utilization restrictions with regard to shipping, we can only recommend water activity measurement in a supporting role within a larger, comprehensive green-coffee assessment program.

Pushing water activity to become an industry standard would very much be putting the cart before the horse. It would impose a poorly understood and expensive measurement with limited application utility and a reasonable substitution on people who cannot afford it, do not currently need it, and would not likely be able to properly utilize it. Better would be to use what we know about Aw in green coffee to help refine MC standards and hopefully also develop applicable best practices in processing, drying, shipment, and storage.

Standards can come later, once processes like water-activity measurement are well studied and thoroughly understood. Those of us who are fortunate enough to have access to such instruments

would do well by not seeking to impose them on others prior to reaching these milestones. As long as these sorts of instruments are burdensome to own and operate, we should use them to improve our use of established and more broadly applicable tools.

Acknowledgments