IMF RM-60 Inlet Bubble Analysis

Ryan Love-Riner

June 3, 2025

Background

After producing drum-roasted specialty coffee for almost 10 years, Iconik Coffee Roasters of Santa Fe, New Mexico invested in a state-of-the-art IMF RM-60 air roaster, more than quintupling potential output and providing new avenues towards paramount quality and consistency. Prior to the IMF, those metrics were entirely in the hands of roaster technicians. With the automated nature of the IMF's roasting protocols, consistency is a relative guarantee and quality is driven purely by recipe optimization.

One key to recipe optimization is the accurate notation of a roast's first crack, which initiates the rapid development of the delicate inner bean. There are a variety of established methods used to identify first crack, most of them relying on a roast technician's senses. Technician sensory based identification is, however, limited and easily inaccurate. During the later stages of its drum roasting program, Iconik began to identify first crack by using the environmental temperature probe inserted near the top of the drum face, noting that as the interior moisture of the beans escaped, the thermal flux of the drum's environment stalled, causing the environmental temperature's rate of rise to drop.

This technique, however, does not apply to the IMF, as its mechanics no longer involve environmental temperature measurement per se. (For a more detailed explanation of the IMF's roasting mechanics, consider Scott Rao's comparative analysis of the IMF and Loring air roasters.) Instead, the sensory data output only includes inlet temperature, bean mass temperature, and percentile airflow settings, as well as the rate of rise of each of those metrics.

While returning to the use of technician senses has worked as a stopgap measure during our transition to air roasting, we wish to pursue a more refined and accurate means of identifying first crack based on the IMF's sensory data. This study aims to develop potential tools for such a technique.

Introduction

In an exploratory analysis of our roast data, we have noticed that, with some regularity, the inlet temperature rises above that dictated by the later stages of the recipe, particularly when the inlet temperature is held constant for the duration of the roast. An example can be seen below; note the bubble present in the right hand side of the upper and lower orange lines, which are the inlet temperature and inlet rate of rise (RoR) respectively.
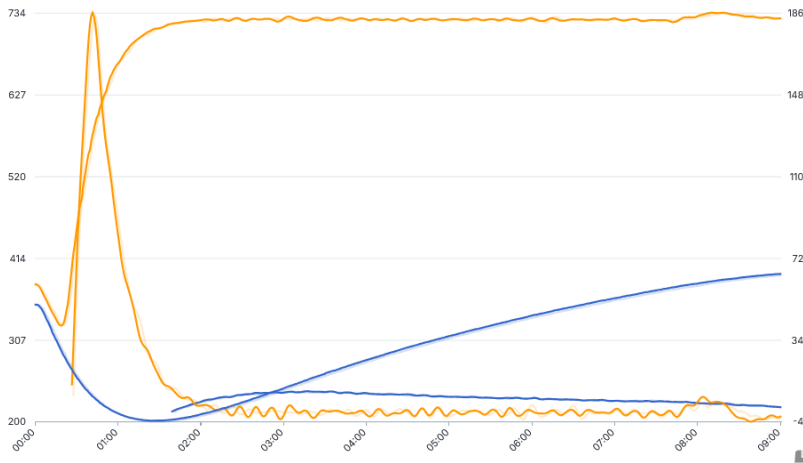
Based on our previous use of environmental reactions to first crack and an exploratory analysis of a data subset, we suspect that this so-called inlet temperature 'bubble' is caused by the IMF's software attempting to compensate for the sudden addition of moisture to the environment and a resulting stall in thermal flux. Should this be the case, the bubble may be a potentially accurate and bean-specific indicator of first crack. In order to make this connection, we will begin by analyzing the variance of our available environmental metrics across bean-roast groups. This distinction is important–although green beans and roast recipes are potentially independant factors, they are made artificially colinear in this study by being tied to one another (each bean is only roasted with a specific recipe). We cannot, with this data, develop a causal relationship between the coffee bean and the inlet bubble. Even so, if we can demonstrate that the inlet bubble is statistically associated with each bean-roast group, we will have strong evidence that the group dictates the nature of the bubble, meaning we can use it as a means of post-hoc first crack identification.

Hypothesis

We propose that the available inlet temperature bubble metrics–`peak temperature time`, `peak temperature difference from recipe`, `peak inlet RoR time`, and `peak inlet RoR`–have statistically significant associations with individual bean-roast groups.

Method

We will use boxplot visualizations and analysis of variance tests, along with some nonparametric alternatives. Model assumptions will be tested visually and formally.

Data

The data used in this analysis consists of 63 roasts performed on the IMF RM-60 over three separate roast days. All roast batches were of a consistent weight. 10 different green coffee groups are represented. Roast data was collected using Cropster before being cleaned and organized in Excel and read into R as .csv files. Please contact the author for further information and/or access to the data, which is property of Iconik Coffee Roasters, LLC.

Analysis

Time of peak inlet temperature   Based on our exploratory analysis, the time of the inlet bubble's peak during roast (`peak temperature time`) appeared most consistent within bean-roast groups. We'll begin by looking at a box plot of the data.


Peak Temperature Time by Bean–Roast Group
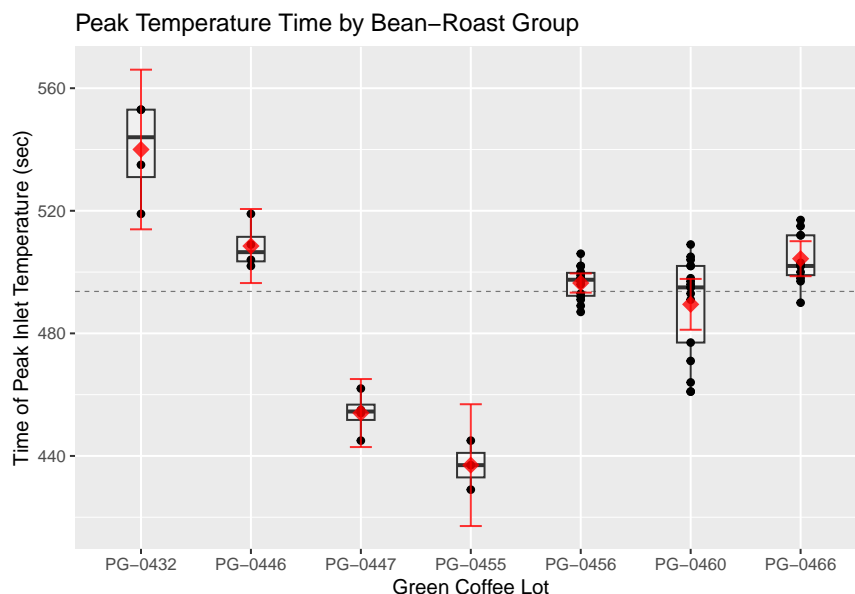
If the boxes were all roughly aligned with one another along the dashed horizontal line, it would mean that the average time of the inlet peaks (red diamonds) were all about the same, regardless of the bean-roast group. Instead, the population mean times do appear different across groups, indicating the potential for useful associations. Unfortunately, we can also see less than ideal variances present (and/or sub-optimal sample sizes of some groups), demonstrated by some wide confidence bands (thin red lines) about the means of some groups. We'll use an ANOVA to test the hypothesis that population means vary across groups.

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## 'Green Id'    6  27059    4510   34.59 3.71e-16 ***
## Residuals    50   6519     130
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The test yields a p-value of 3.17e-16, far smaller than even a rigorous alpha value of 0.05. This suggests a statistically significant difference in means across groups. But the ANOVA test relies on two model assumptions (a normal distribution of residuals and a homogeneity of variance across groups) that must be tested before we can accept this result.

We'll first plot the residual densities to look for a normal distribution.

Density vs Residuals (peak temperature time)

Blue = Kernal density curve
Red = Normal distribution
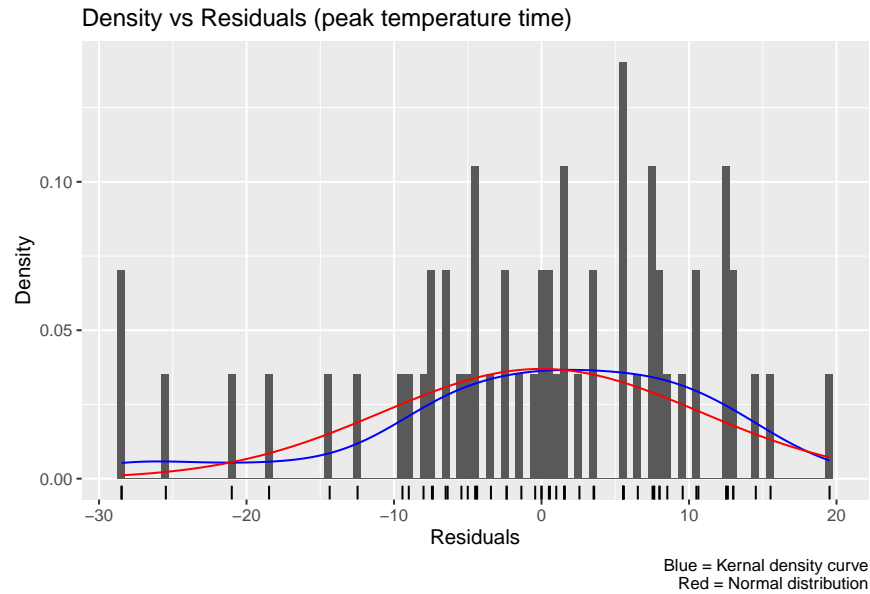
In the histogram plot above, the bars represent the kernal density of residuals. The red line traces the shape of what would be considered a normal distribution of residual densities, while the blue line shows the actual distribution. We want them to line up as closely as possible, but they don't appear to match well. We can also check a QQ plot of the data to confirm our suspicions.

QQ plot of peak temperature time
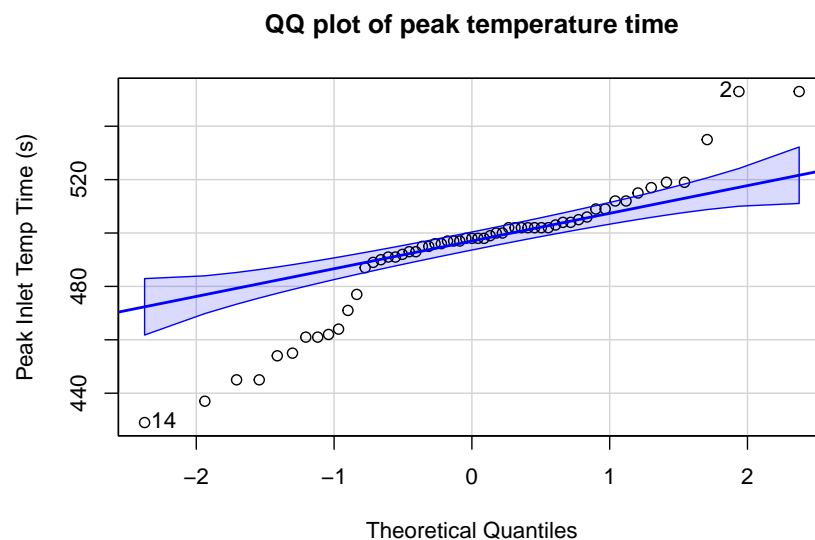
Observations at the extremes fall far outside the normal distribution. We'll use the Anderson-Darling test for normality to confirm this formally.

```
##
##  Anderson-Darling normality test
##
## data:  inlet_peak_aov$residuals
## A = 0.73762, p-value = 0.05154
```

This represents an unusual case! Since the p-value yielded by the test is $0.0515 < 0.1$, we must reject the null hypothesis and conclude that the distribution of variance density is not normal, violating one of the model assumptions and invalidating the results of our ANOVA test. But not to worry! There are some alternatives to which we have recourse. First, let's check the model for homogeneity of variance using Levene's test, which does not rely on normally distributed residuals.

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  6  1.7133 0.1373
##       50
```

With a p-value of $0.14 > 0.05$, we can reject the null hypothesis and conclude that the population variance of observations are homogenous, validating that model assumption. Even with homogenous variance across the population, we can't trust the results of the ANOVA test above. We could attempt to transform the data, but since we don't plan to make projections based on this analysis, we have a shortcut at our disposal: the Kruskal-Wallis test. It doesn't rely on a normal distribution of residuals, and instead of examining the difference in means across groups, it measures the group medians.

```
##
##  Kruskal-Wallis rank sum test
##
## data:  Peak Inlet Temp Time by Green Id
## Kruskal-Wallis chi-squared = 34.868, df = 6, p-value = 4.572e-06
```

Given a p-value of $4.57e\text{-}06 < 0.05$, we can reject the null and finally conclude a statistically significant difference in observation medians across groups. What this means is that the time at which the inlet peaks occur are statistically associated with individual bean-roast groups, an excellent first step in using this metric to identify first crack. Still, we have several other metrics at our disposal, which we will test similarly.

Peak Temperature Difference From Recipe   Next, we'll examine the temperature difference between the inlet peak and the prescribed inlet temperature (`peak temp diff`), proceeding much as we did above.


Peak Temp Diff by Bean−Roast Group
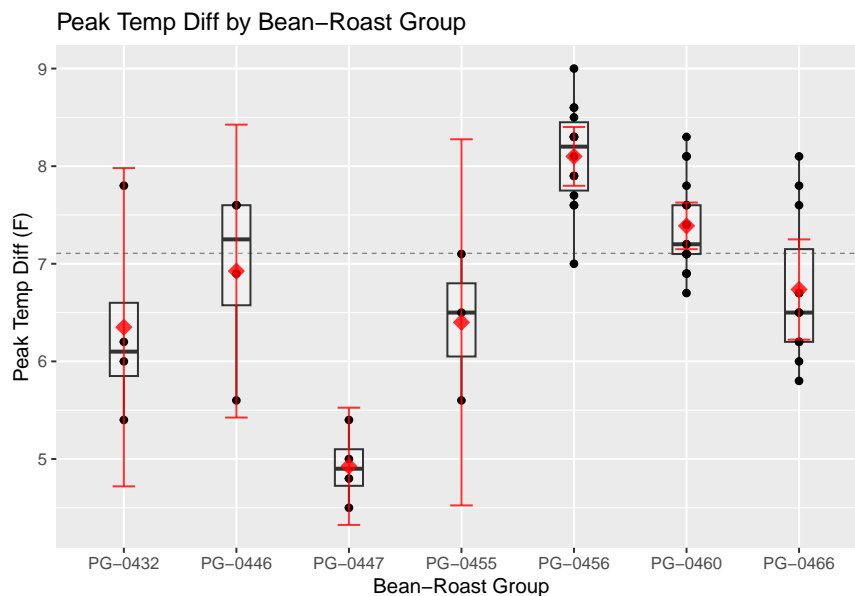
Again, there does appear to be significant variation in group means. The confidence intervals on some bean-roast groups look unfortunately large, however.

```
##               Df Sum Sq Mean Sq F value  Pr(>F)
## 'Green Id'     6  39.63   6.605   16.33 2.66e-10 ***
## Residuals     50  20.23   0.405
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Our ANOVA test on this metric yields a p-value of 2.66e-10 < 0.05, causing us to reject the null hypothesis and conclude a significant difference in means between groups. We'll need to check model assumptions, starting with a normal distribution of residuals.
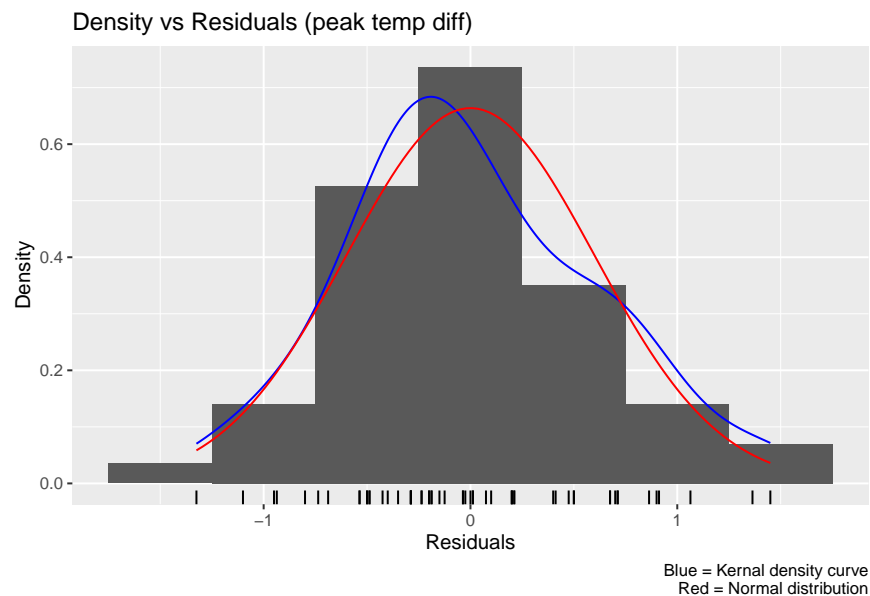
The distribution of the residuals looks far closer to normal here. Let's check a QQ plot.
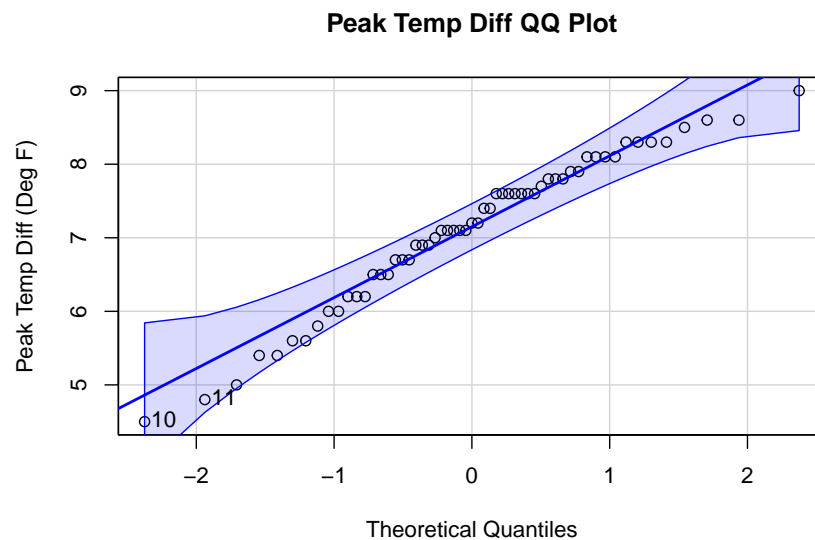
This looks even better: all observations are contained within the confidence bands, demonstrating a normal distribution of quantiles. Still, we'll use the Anderson-Darling test to be safe.

```
##
##  Anderson-Darling normality test
##
## data:  inlet_peak_aov2$residuals
## A = 0.38166, p-value = 0.3888
```

Since the p-value of 0.39 is over the alpha value of 0.05, we can conclude that the data is normally distributed. Next, we'll check variance across groups for homoscedasticity.

We used Levene's test earlier, since it doesn't assume normality. This time we'll use Barlett's test, which does assume normality.

```
##
##  Bartlett test of homogeneity of variances
##
## data:  Roast Recipe Inlet Difference at Peak by Green Id
## Bartlett's K-squared = 7.9596, df = 6, p-value = 0.2411
```

It yields a p-value of 0.24 > 0.05, which fails to reject the null and allows us to conclude homogeneity of variance. Since we have shown that both model assumptions have been met, we can rely on the ANOVA test results.

This is a powerful demonstration that `peak temp diff` means are significantly different across groups. Not only does this support what we found concerning the association between inlet peak time and bean-roast group, it emphasizes a more direct association between the inlet bubble and the green bean itself. Since the represented bean-roast groups employ a variety of recipes, something other than the recipe is causing a difference in these inlet temperature peaks. More analysis into this relationship is recommended.

Peak RoR Time   For now, we'll move on to our third metric, the time of peak inlet temperature rate of rise (RoR). This data employs an RoR cycle time of 15 seconds, meaning that each observation of (`peak inlet RoR time`) represents the time at which the maximum inlet temperature degree change in Fahrenheit per 15 seconds occurs.
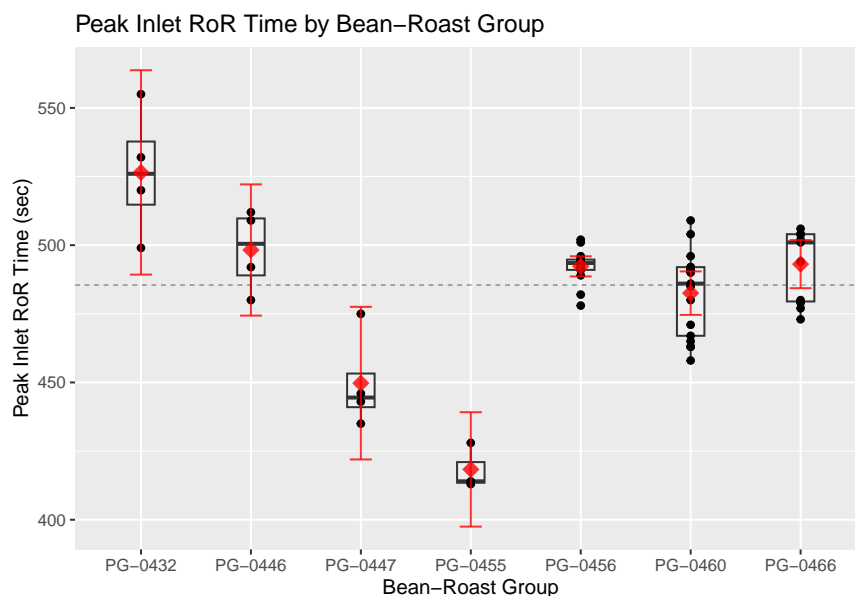
This plot looks less promising than the first two. Some of the group means appear different, while others appear quite close to the population mean.
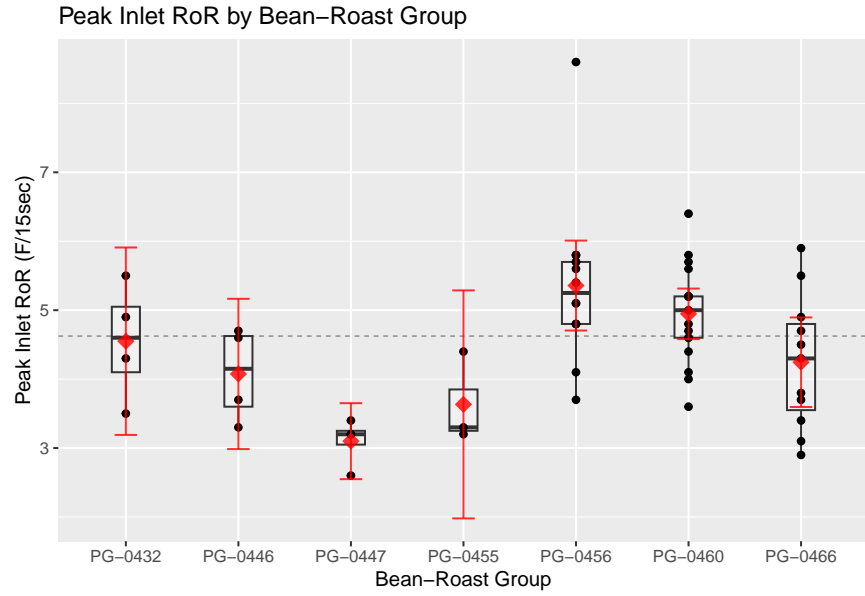
```
##                Df Sum Sq Mean Sq F value  Pr(>F)
## 'Green Id'    6  27449    4575   24.36 2.88e-13 ***
## Residuals    50   9389     188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The ANOVA test (p-value 2.88e-13 < 0.05) suggests a significant difference in group means, pending a model assumptions check. First, the residuals.

Blue = Kernal density curve
Red = Normal distribution
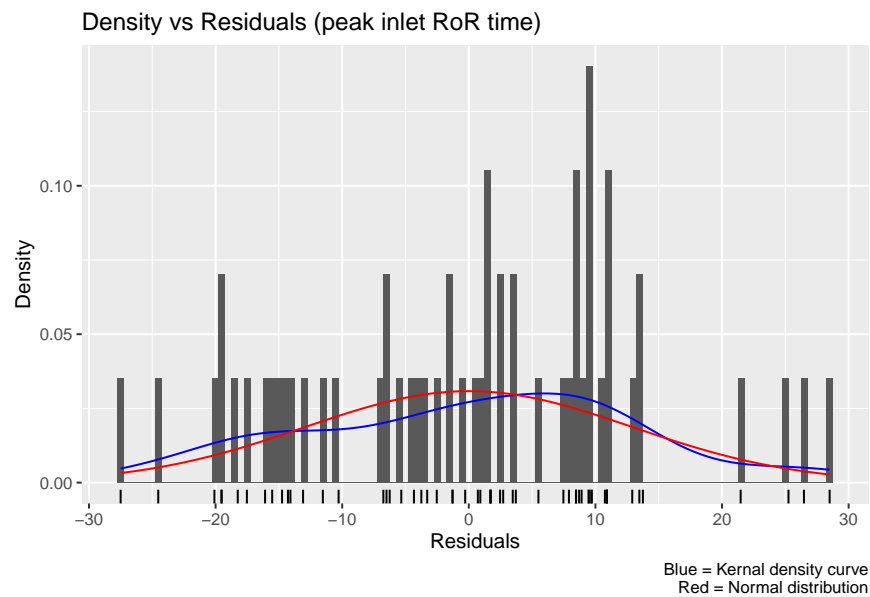
We can immediately see some prominent left-skewedness in the blue curve, meaning the residuals are not normally distributed. Although we could perform transformations on the data to achieve more normally distributed residuals, they may be complicated and we have already established two powerful metrics. We will pass over this one for now.

Peak RoR Temperature Finally, we can look at the actual peak inlet temperature RoR (`peak inlet RoR`)for each roast.

Peak Inlet RoR by Bean–Roast Group

The group means here do look fairly different.

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## `Green Id`    6  24.34   4.056    5.21 0.000313 ***
## Residuals    50  38.93   0.779
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The ANOVA test confirms this with a p-value of $0.0003 < 0.05$, causing us to reject the null hypothesis and conclude a significant difference in group means. Next we'll check residuals.

Density vs Residuals (peak inlet RoR)

Blue = Kernal density curve
Red = Normal distribution
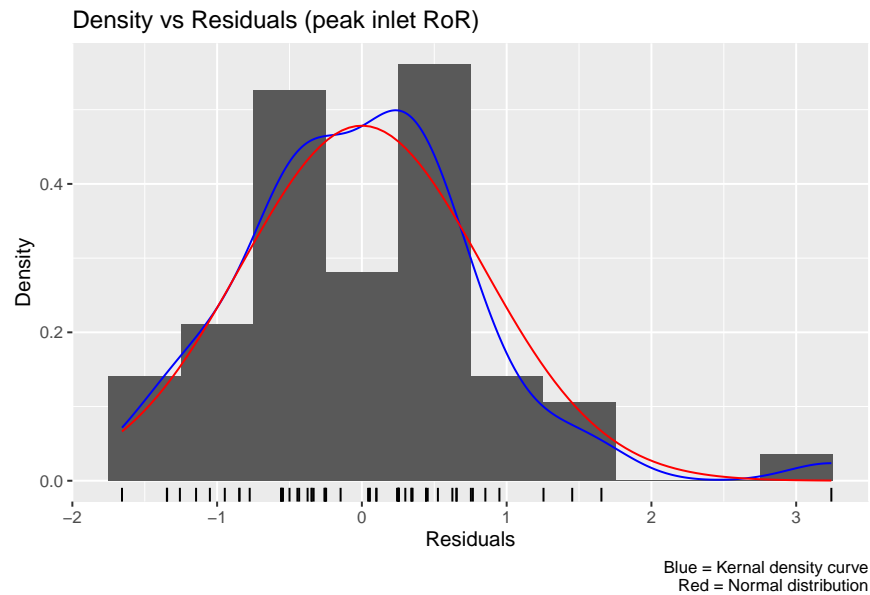
The theoretical and real distributions of variance trace reasonably well, although the real (blue) curve does show some bimodalism. We'll check it formally to be sure.

```
##
##  Anderson-Darling normality test
##
## data:  inlet_peak_aov4$residuals
## A = 0.53313, p-value = 0.1656
```

The Anderson-Darling test rejects the null with a p-value of $0.17 > 0.05$, allowing us to conclude a normal distribution of residuals. With that normality in hand, we can use Bartlett's test for homogeneity of variance.

```
##
##  Bartlett test of homogeneity of variances
##
## data:  Peak Inlet RoR by Green Id
## Bartlett's K-squared = 6.3865, df = 6, p-value = 0.3813
```

Since it yields a p-value of $0.38 > 0.05$, we fail to reject the null and conclude homogeneity of variance. With both model assumptions met, we can conclude that the ANOVA test is sound, meaning that the group means of `peak inlet RoR` vary. This represents a third piece of powerful evidence that the values of the inlet temperature bubble are associated with the roast being executed.

Conclusion

We have shown that three of our available IMF environmental metrics–`peak temperature time`, `peak temp diff`, and `peak inlet RoR`–have averages that are statistically associated with bean-roast groups, confirming our hypothesis. From this we may conclude that the inlet bubble phenomenon is correlated with the roast being executed and we may confidently use this bubble as a means of identifying first crack on a post-hoc basis, specifically by counting back a calculated amount of time from the median value of `peak temperature time` within each bean-roast group.

Limitations

Causal relationships have not been established. Artificial colinearity may exist between green beans and roast recipe. Hidden variables in ambient roasting environment are unlikely, but may exist. Number of observations are limited, especially within certain groups. Results apply exclusively to the IMF model RM-60 air roaster.

Further Research

While not necessary for identification purposes, demonstrating a causal relationship between first crack and the inlet bubble would greatly strengthen the conclusions of this study, as well as provide valuable insight into the physics of industrial air roasting and the mechanics of the IMF. We recommend designing an experiment that roasts the same green coffee bean using two or more different recipes and comparing the bubble metrics. This would isolate the green bean factor as the only variable contributing to the bubble.

This study also has valuable implications for the process of grading green bean consistency. Iconik has noted a growing concern with inconsistent quality in green bean lots, noting a heightened variability in roasting behavior within groups. The box plots produced in this study point to this issue as well. Although roasting consistency is not synonymous with flavor optimization, it is a significant contributing factor in product consistency. Analysis of variance in these and other roasting metrics may prove useful in determining the viability of long-term sourcing relationships.