Chapter 5: Empirics of production

Jeffrey Sachs, James Rising,
Tim Foreman, John Simmons, Manuel Brahm
The Earth Institute
Columbia University

October 1, 2015

Prepared for the Global Coffee Forum
Milan, Italy

Part of
The impacts of climate change on coffee: trouble brewing
http://eicoffee.net/

In this chapter, we develop a globally applicable coffee production model, which describes the predictability of yields under current and future weather.

1 Crop modeling approaches

One broad approach to predicting crop development and providing decision-support tools to farmers is through biological process models. These models capture the process of plant development (phenology) at the individual plant level, and are available for many crops as models for widely-used crop modeling systems such as DSSAT and APSIM[1]

Biological process models of coffee production appear to be in their early stages. The most advanced may be the one developed by Rodríguez et al. (2011), with a "tri-trophic, physiologically-based system perspective", capable of studying water and light needs and pest impacts. The next generation of biological models, represented by Dauzat et al. (2014) and Maro et al. (2014) are still under development. This state of the literature motivated our focus on statistical models.

A statistical production model relates high-resolution weather data (such as temperature and precipitation) with observed yields (Schlenker and Roberts, 2009). The most advanced of these estimate the effect of growing degree-days (GDDs) and "killing degree-days"(KDDs) in a non-linear fashion, and account for varying unobserved characteristics that are idiosyncratic to each region, such as management, elevation, and soil properties. Statistical approaches have been used to study individual regions (e.g., Gay et al., 2006; Guzmán Martínez et al., 1999). We use our global coffee production database and to generate a global model which with elements that vary from one country to another.

The statistical techniques we use fall under the heading of econometrics. Econometric techniques allow for the inclusion of many different parameters and a treatment of differences between regions which are not directly captured in our data. They also allow for the careful identification of "causal" relationships, rather than simple correlation. We then extend this statistical model with a new technique developed here, which we call hierarchical modeling. The hierarchical model consists of three levels of hierarchy: sub-national models, national models, and a global model. At each level, different regions are allowed to have different responses of yields to weather, but are also informed by the effect estimated across all regions (e.g., all national models are informed by the global combined model).

These statistical models are estimated using natural experiments, by comparing observed yields in years with different distributions of weather to estimate the effect of weather in general. These experiments completely inform our models of production. The models below include daily minimum and maximum temperatures, precipitation, and humidity. We do not include soil properties in this chapter because it is impossible to do statistical experiments where soil characteristics vary over time, to see its effects. As a result, the statistical model cannot determine the effects of soil.

This approach puts a "black box" around the complicated system surrounding production, and makes no attempt to disentangle the effects of farmers responding to weather, the effects of that weather on the crops themselves, and the effects that these have on the plant's susceptibility to disease. This black box is both a strength and a limitation. It captures realistic relationships between weather and yields, rather than theoretical responses of the crops in an experimental setting. It can capture the environmental determinants of coffee disease spread, and their impacts implicitly. It can also be used to predict yields under climate change and weather events. However, because it cannot distinguish the social and natural causes, it makes an implicit assumption that yields will continue to respond the same way to increasing temperatures over time.

[1]DSSAT is the Decision Support System for Agrotechnology Transfer(?) and APSIM is the Agricultural Production Systems Simulator(?).

The production model can also be used to predict yields months before a harvest. By combining climatological signals, like ENSO, for which there is some predictive skill, with yearly averages (climatologies), it is possible to generate plausible weather patterns to apply to the model. However, these results could only be taken as suggestive: the statistical models we produce only account for about 32 - 38% of the variation in yields across time and space. The biennial cycle of coffee, for example, is not explicitly captured in our model, which considers only effects driven by weather (Bernardes et al., 2012).

In addition to the biennial cycle, there are a large number of factors which drive coffee yields that are not explicitly included in this model: market drivers, evolving technology, changing varieties, and the governance and politics which frequently affect the coffee sector. These are all important. By limiting our analysis to the student of weather and climate change, we can better understand those elements.

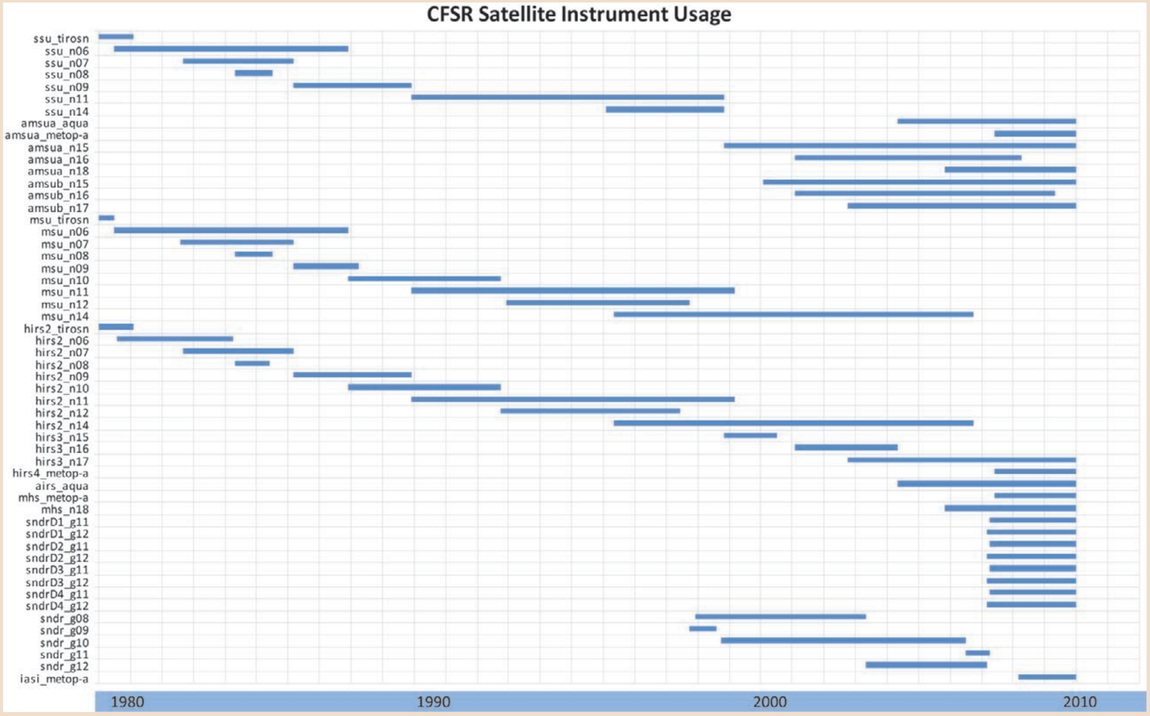
2 Weather and climate data

The current climate is represented by weather records from recent history. We use weather data since 1979 from the Climate Forecast System Reanalysis (CFSR). This data product combines station and satellite measurements using weather models to produce reliable weather estimates at a high spatial and temporal resolution. The spatial resolution is .32°x.32°, a grid with boxes that are about 35 km on a side at the equator. The temporal resolution is hourly, which we use to generate growing degree-days at a daily scale.

Yields and production data are not available in a high-resolution, gridded form. Instead, yields, in the form of production quantities and harvested areas, are reported for political units. High resolution information about coffee producing regions needs to be combined with these low resolution recorded yield data. For example, coffee is grown exclusively in the southwest of Guatemala, in regions that cover 8.7% of the land area, but production data is reported for the entire country. Since we know that the country-wide production is coming only from these regions, we can limit the weather and other data used to infer coffee production relationships. To match the gridded weather data with growing regions, we use our coffee production database to spatially aggregate the weather data.

Climate Forecast System Reanalysis

The Climate Forecast System Reanalysis is a global weather product constructed by NOAA (Saha et al., 2010). CFSR merges the overlapping ranges of satellite products, as they are available across years:



CFSR combines both conventional and satellite data from the following sources:

Conventional: Radiosondes and Pibals, AMMA special observations, Aircraft and ACARS data, Surface observations, PAOBS, SATOB observations, SSM/I ocean surface wind speed, Scatterometer winds

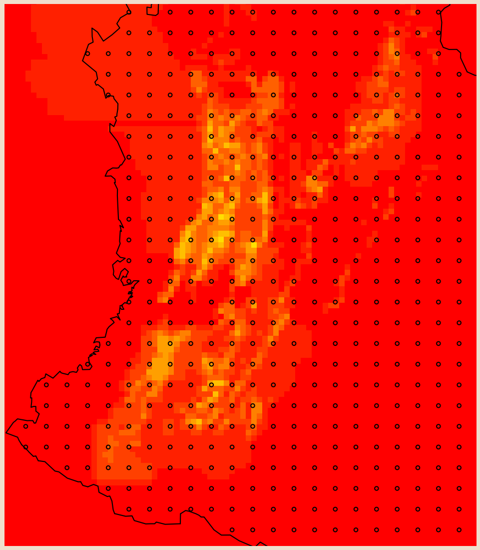
Satellite-radiance: TOVS radiances, Recalibrated MSU radiances, ATOVS radiances, GEOS radiances, Aqua AIRS, AMSU-A, and AMSR-E data, MetOp IASI, AMSU-A, and MHS data, CHAMP/COSMIC GPS radio occultation data.

Spatially-weighting weather

To generate weather observations at the same spatial aggregation as yields, we perform the following procedure. For each political unit,

1. Translate CFSR grid cells into a lattice of points.
2. Find all grid lattice points within a given country.
3. Identify the measure of harvested area in the coffee database nearest to each lattice point.
4. Take the weighted average of weather observations, weighted by coffee harvested area.

An example is shown below for grid cells that fall within Colombia.



Circles show the location of CFSR grid lattice points. Colors show the coffee weights.

3 Brazil case study

The Brazilian Institute of Geography and Statistics (IBGE) provides municipality-level production for coffee in Brazil since 1990. Nearly 2,700 municipalities with coffee production histories are included, and representing an average resolution of less than 40 km. This dataset allows for a broad case study of the impacts of climate change at a high spatial resolution. We refine the structure of our production model for Brazil before applying it globally.

3.1 An empirical model of production

Using the IBGE Brazilian coffee production estimates, combined with high resolution weather from the CFSR reanalysis product, we estimate a physically-based statistical model of coffee production. The model predicts yields using a nonlinear relationship with temperature and precipitation. We base our model on Schlenker and Roberts (2009), and divide GDDs into three groups: beneficial growing degree-days between 0°C and 33°C, killing degree-days above 33°C, and frost degree days below 0°C. We also use the average minimum temperature, which appears to be more significant than frost degrees. This kind of statistical relationship is based on the biological response of coffee to temperature, but puts a "black box" around farmer responses and ecosystem and pest dynamics. If farmers are providing sufficient irrigation and shade to coffee plants, the effect of high temperatures will be mitigated beyond what biological models suggest on their own.

Calculating growing degree-days

Growing degree-days (GDDs) are calculated using a continuous sinusoidal fit to minimum and maximum daily temperatures, as shown below:

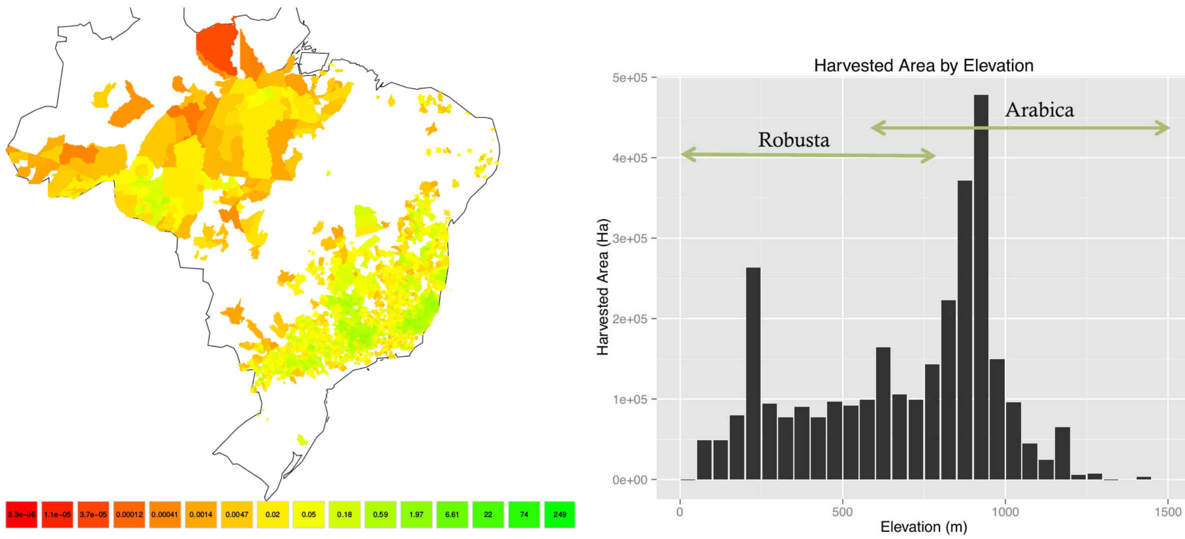
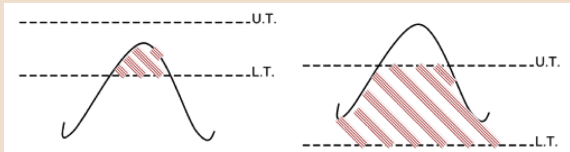


Figure 1: Brazil dataset across space and elevation. **Left:** Density of coffee production, as the average production divided by municipality area. Regions in green account for the majority of production. Most production occurs in the south, however there are coffee producing regions also in the southern Amazon. **Right:** Distribution of coffee producing area, displayed across the average elevation of each municipality. The greatest extent of coffee production occurs in municipalities with around 900 m of elevation, but coffee is also produced in municipalities with a much lower elevation, including a peak around 200 m. The range of typical elevations for growing Arabica and Robusta are shown above the histogram.



Calculations for growing degree-days and killing degree-days. Any temperatures above a given lower threshold (L.T.) are included, up to a maximum of an upper threshold (U.T.). As temperatures shift over the course of a day, fractional growing degree-days are accumulated.

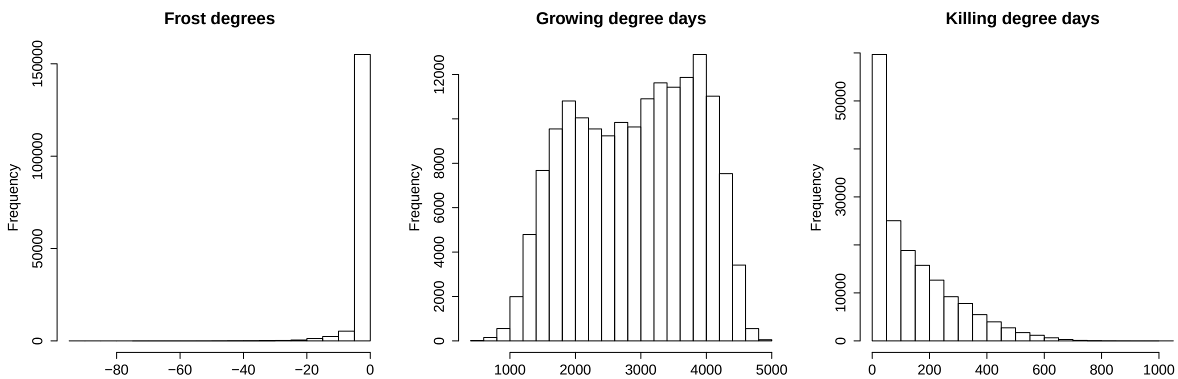


Figure 2: Histograms displaying the number of growing seasons with a given number of frost degree-days, growing degree-days, and killing degree-days. The exponential decays in frost and killing degree days are useful for capturing the impact of extreme events. The broad range of growing degree-days represented in the center histogram allows for accurate estimates of the coffee growth response.

We also include precipitation, as the total accumulated precipitation over the six months before harvest. Precipitation is included as a quadratic, to capture the expectation that both too little precipitation and too much precipitation are harmfully impact yields.

3.2 Optimal temperature range

Guzmán Martínez et al. (1999) suggest that 10°C is the appropriate base temperature for calculating GDDs for coffee. We explore a large range of minimum and maximum temperatures for GDDs, seeking the limits that provide the greatest predictive capacity. See Appendix .1 for predictive capacity of a range of possible limits. We find that a minimum temperature of 0°C and a maximum temperature of 33°C for beneficial GDDs is optimal. This means not only that all days over 0°C are estimated as beneficial, but that higher temperatures up to 33°C are progressively more beneficial. A day above 33°C is not immediately detrimental, but it has a progressively smaller benefit until it becomes negative, and we find that temperatures over about 35°C are detrimental in Brazil.

3.3 Predictive periods

Coffee production is very sensitive to weather during flowering, and the period during which we correlate weather with yields is important. To determine the optimal span of weather for predicting yields, we try out many combinations of starting and ending months. The harvesting period in Brazil ends in September, so we consider months starting with October to predict the yield in the next year. The coefficients of models for each of these periods are shown in figure 3.

A few features are important in these results. In the top graph displaying coefficient values, areas in the upper-left are gray, denoting that models that use only the months shortly preceding harvest do not produce significant results. Second, we expect the effect of GDDs to be positive, KDDs negative, the linear component of precipitation (`precip`) to be positive, and the quadratic component of it (`precip2`) to be negative. This is confirmed for most date ranges, and we want to avoid regions that misestimate these values due to noisy or minor effects. Finally, the t-values figures show the confidence in these values, and are a measure of the statistical significance of the model as a whole. These values generally decrease as the starting month becomes later.

Figure 4 shows the combined t-values for the GDD and KDD coefficients. The highest t-value is for GDD and KDD values calculated just for January and February. The probably reflects a highly sensitive period for the berry production. Nearly as high, and covering a six-month span, is December through May. We will use this as our span for calculating weather impacts.

3.4 Econometric model

The form of the statistical model is,

$$\log y_{it} = \alpha_i + \gamma g_{it} + \kappa k_{it} + \mu m_{it} + \pi p_{it} + \psi p_{it}^2 + P_{3,s(i)}(t) + \epsilon_{it}$$

Above and in the other models below, the observation variables and their corresponding effect estimating coefficients are:

	Var.	Coeff.
Growing degree-days	g_{it}	γ
Killing degree-days	k_{it}	κ
Average minimum temperature	m_{it}	μ
Total precipitation (linear)	p_{it}	π
Total precipitation (quadratic)	p_{it}^2	ψ

where i indexes municipalities, t the years, and $P_{3,s(i)}(t)$ is a state-specific cubic trend to capture shifting productive capacity. We aggregate weather from December to May, and use 0°C to 33°C as the limits for computing growing degree-days.

Interpreting regression tables

Many of the results in this chapter are in the form of multiple regression tables. Each regression is of the form,

$$y_i = \alpha + \beta_1 x_{1,i} + \cdots + \beta_k x_{k,i} + \epsilon_i$$

which describes the relationship between a dependent variable, y_i, taking different values for each ith observation, and a linear combination of independent variables, $x_{1,i}, \ldots, x_{k_i}$. The ϵ_i term represents the remaining error that cannot be explained by the model. In addition, these models use “fixed-effects”, which are parameters unique to each region, so that the model is effectively estimated by

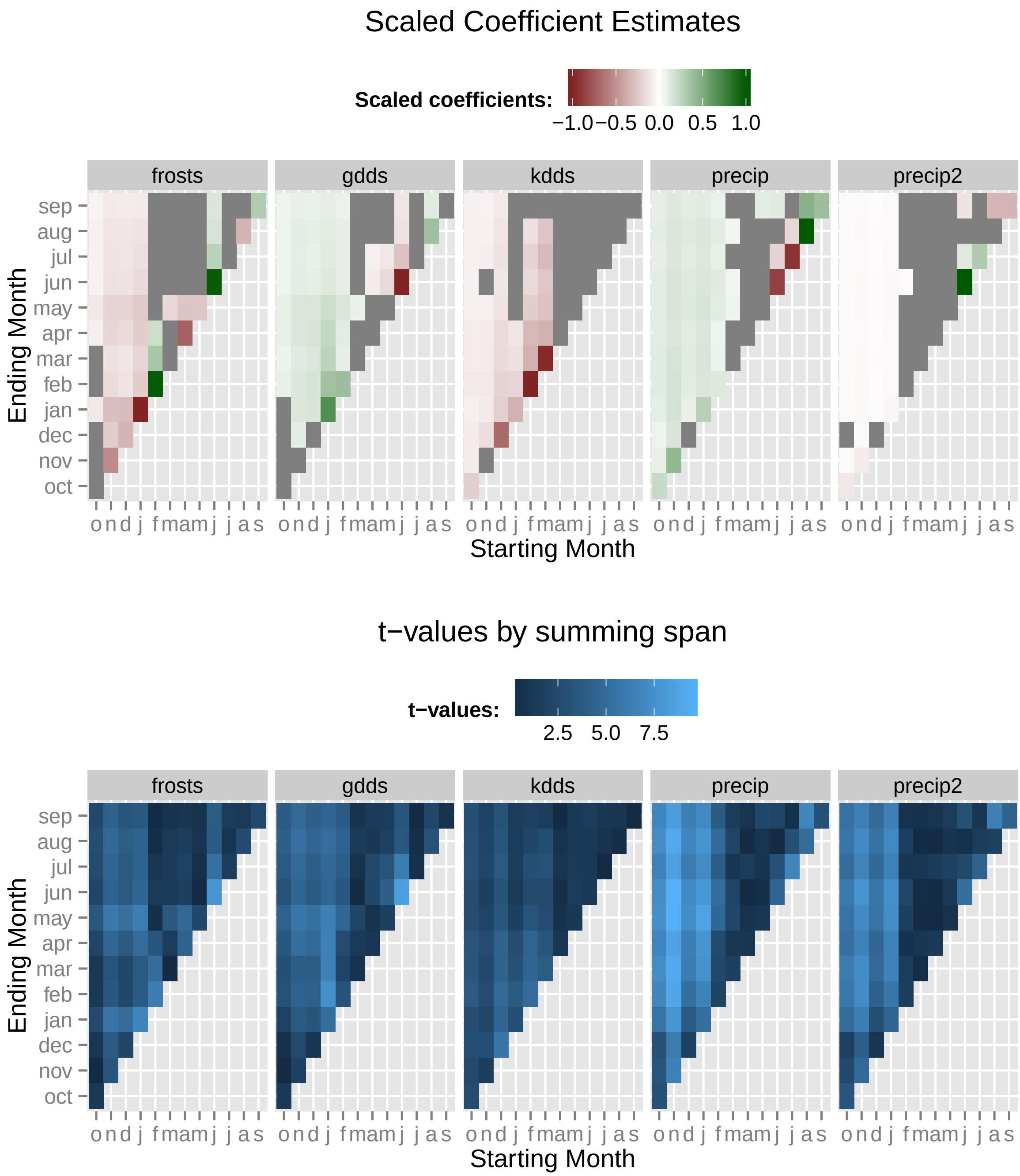


Figure 3: Coefficients from estimating models with different month spans, and the t-values intervals associated with each coefficient. The top 118 municipalities in harvest density were used.

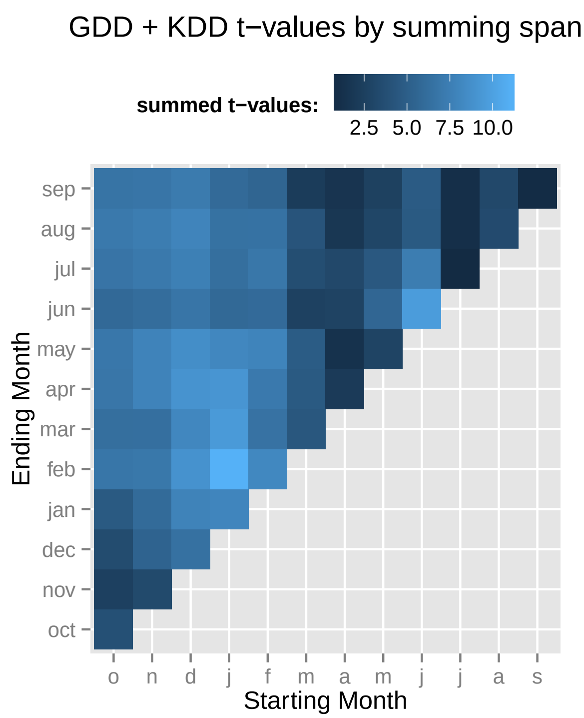


Figure 4: The sum of t-values across the GDD and KDD coefficients, for identifying the most effective range.

considering the effects of changes in the independent variables, rather underlying static differences between them.

The regression tables are mean to be read in columns. The first column specifies the variable for which an effect is reported, and the model columns specify the size of that effect. If a coefficient estimated is 10, that means that the dependent variable increases by 10 for every unit the independent variable increases.

The numbers directly below each effect and reported in parentheses are the values 'standard errors', a measure of the uncertainty of that value. If the standard error is less than half of the value, then there is 95% confidence that the sign of the coefficient in question is correct. This corresponds to the statistical significance of the estimate, and is denoted by asterisks (***).

The results are shown below as a table of statistical coefficients. Table 1 displays the results across all municipalities, and 2 is for the 118 municipalities with the greatest density of coffee harvesting.

		Dependent variable:	
	Means	Log Yields	Harvested Hectares
		(1)	(2)
GDDs / 1000	2.946	0.152***	72.869
	(0.931)	(0.050)	(124.246)
KDDs / 1000	0.149	−2.806***	−2,197.369***
	(0.146)	(0.342)	(555.055)
Avg. Min.	0.944	−0.091***	−25.0
	(3.499)	(0.018)	(34.0)
Precip. (m)	1.421	0.347***	−9.587
	(0.719)	(0.028)	(64.092)
Precip.^{2} (m)	2.538	−0.366***	−8.520
	(2.439)	(0.036)	(84.618)
State cubic trends		Yes	Yes
Observations		43,165	43,185
R^2		0.383	0.655
Adjusted R^2		0.343	0.633
Residual Std. Error		0.535 (df = 40542)	4,300.446 (df = 40561)
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 1: Estimates for statistical models relating growing degree-days, killing degree-days, average minimum temperature, and precipitation to the logarithm of yields, and to harvested area, for all municipalities. Stars (***) represent statistical significance levels, showing that most coefficients appear to have a relationship with production outputs.

We find that increases in temperature below a daily maximum temperature of 33°C limit are beneficial, resulting in higher yields and higher total production (see Appendix .1). Based on this, we compute "growing degree days" (GDDs) as the degree days[2] between 0°C and 33°C. All temperatures above 33°C

[2] An explanation of degree days is at `https://en.wikipedia.org/wiki/Degree_day` and Appendix 3.1 for our method of calculating them.

	Dependent variable:	
	Log Yields	Harvested Hectares
	(1)	(2)
GDDs / 1000	0.475***	1,700.306*
	(0.109)	(976.997)
KDDs / 1000	−2.989**	−23,179.330***
	(1.423)	(8,681.404)
Avg. Min.	−0.183***	−290.009
	(0.0183)	(335.665)
Precip. (m)	0.441***	−1,168.520*
	(0.076)	(677.845)
Precip.^{2} (m)	−0.494***	1,978.722**
	(0.099)	(854.580)
Observations	3,181	3,181
R^2	0.320	0.485
Adjusted R^2	0.290	0.462
Residual Std. Error (df = 3043)	0.364	14,412.800
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 2: Estimates for statistical models relating growing degree-days, killing degree-days, average minimum temperature, and precipitation to the logarithm of yields, and to harvested area, for the top 118 municipalities by production density. Stars (***) represent statistical significance levels, showing that most coefficients appear to have a relationship with production outputs.

are combined into the measure of "killing degree days" (KDDs).

Any days in which the maximum temperature exceeds 35°C have a sharply harmful effect. As a result, even small increases in temperatures under climate change can produce large decreases in yields, particularly in regions where temperatures are currently nearly optimal. This is consistent with other work on the nonlinear effects of high temperatures (Schlenker and Roberts, 2009).

Every additional 1000 GDDs (of which there are about 3000 on average in coffee-growing municipalities in Brazil) increases yields by about 16%. Every additional 100 KDDs (an average year will have only 150 KDDs) decreases yields by 76%. These values are estimated using marginal changes, so the average year is the baseline from which these percent changes are applied.

Figure 5 shows a graphical representation of the growing degree-day production model, with 95% confidence intervals. The assumptions are as described before: growing degree-days and precipitation are calculated using hourly reanalysis data; state cubic trends capture the evolution of coffee production.

There is also a large and statistically significant negative effect on harvested acres. This suggests that in hot years where the crop is damaged, the plants are simply not harvested. As a result, the actual damaging effects of high temperatures on yields are likely to be greater than reported. The yield numbers hide the fact that unproductive plots in poor years can be left unharvested, causing both total production and harvested acres to decrease without as large of decreases in yield.

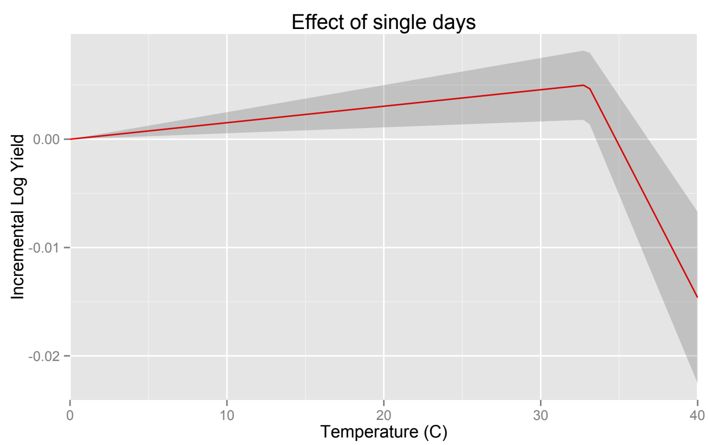


Figure 5: Marginal impact on log yields for an additional day at a given temperature. Up to 33°C, additional temperature results in greater yields. Above 33°C, this effect is sharply diminished and hot days above 35°C result in large decreases in yield. The grey band shows the 95% confidence intervals around the estimated effect for a single day at a given temperature.

3.5 Multilevel Brazil model

Next we extend the model to include "multilevel" effects. The multilevel model studies how the estimated coefficients vary across other characteristics of the municipalities. In this case, we consider how the effect of GDDs, KDDs, and average minimum temperature vary with elevation. Elevation is both an important determinant of coffee quality, and is a proxy for the variety of coffee grown: Brazil grows both Arabica and Robusta coffees, but does not report their production separately (until recent years).

The multilevel relationship is that:

$$\begin{aligned}
\log y_{it} &= \alpha_i + \gamma_i g_{it} + \kappa_i k_{it} + \mu_i m_{it} + \pi_i p_{it} + \psi_i p_{it}^2 + \epsilon_{it} \\
\gamma_i &= \gamma_0 + \beta_\gamma Elevation_i + \eta_{\gamma,i} \\
\kappa_i &= \kappa_0 + \beta_\kappa Elevation_i + \eta_{\kappa,i} \\
\mu_i &= \mu_0 + \beta_\mu Elevation_i + \eta_{\mu,i} \\
\pi_i &= \pi_0 + \beta_\pi Elevation_i + \eta_{\pi,i} \\
\psi_i &= \psi_0 + \beta_\psi Elevation_i + \eta_{\psi,i}
\end{aligned}$$

where the top line is the normal regression relationship, but with separate coefficients for each municipality i. The remaining lines relates all municipality coefficients together according to their varying elevations. The results are shown in table 3 and in a graphical form in figure 6.

Next we consider how the sensitivity to temperature varies with elevation. The parameters that mediate this sensitivity– the positive effect of GDDs and the negative affect of KDDs– are shown in figure 6. We find that temperatures above 33°C at 1000m above sea level are five times as damaging as they are at 250m. These results support the common wisdom: Arabica, grown at higher elevations, is much more sensitive to weather than Robusta. We find that as elevation increases, the potential increased yield from higher temperatures as well as the potential damage due to extreme temperatures increase.

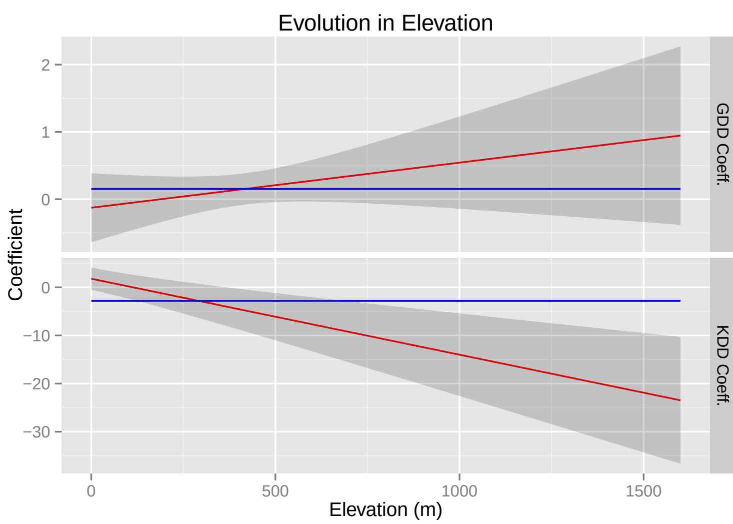


Figure 6: The effect of an additional GDD and KDD as these vary by elevation. As elevation increases, plants become more sensitive to temperatures. The effect of GDDs increases, though very slightly. The harmful effects of KDDs increase quickly.

3.6 Yield estimates under a warmer climate

We can apply the production model to weather produced from climate change. As a proxy for climate change, we estimate yields using historical weather data increased by 2°C. Precipitation values are left

	Dependent variable:	
	Log Yields	Harvested Hectares
	(1)	(2)
GDDs / 1000	0.208***	40.303
	(0.051)	(130.508)
Elev. GDDs / 1000	0.001***	2.110***
	(0.0002)	(0.657)
KDDs / 1000	−6.106***	−4,600.562***
	(0.516)	(725.931)
Elev. KDDs / 1000	−0.016***	−17.054***
	(0.002)	(3.653)
Avg. Min.	−0.183***	−25.750
	(0.018)	(34.334)
Elev. Avg. Min.	−0.00000**	−0.183
	(0.00000)	(0.183)
Precip. (m)	0.358***	−32.650
	(0.030)	(76.846)
Elev. Precip. (m)	0.0001	−0.164
	(0.0001)	(0.285)
Precip.^{2} (m)	−0.391***	−10.825
	(0.039)	(98.941)
Elev. Precip.^{2} (m)	0.0001	0.648*
	(0.0001)	(0.390)
Observations	42,141	42,161
R^2	0.378	0.651
Adjusted R^2	0.338	0.628
Residual Std. Error	0.538 (df = 39582)	4,282.486 (df = 39601)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 3: The effects of GDDs, KDDs, and average minimum, as each varies by elevation. While the estimates are not significant, they suggest increasing sensitivity to temperature in the form of both GDDs and KDDs as elevation increases. All municipalities in Brazil used.

unchanged, since they show an unclear trend. This change produces several effects: it increases the number of GDDs benefiting yields, increases the number of KDDs harming yields, and increases average minimum temperature. The resulting balance between these three impacts is not evident *a priori*. The figure below shows the distribution for municipality yields across Brazil, from observed data, and under climate changed weather predictions.

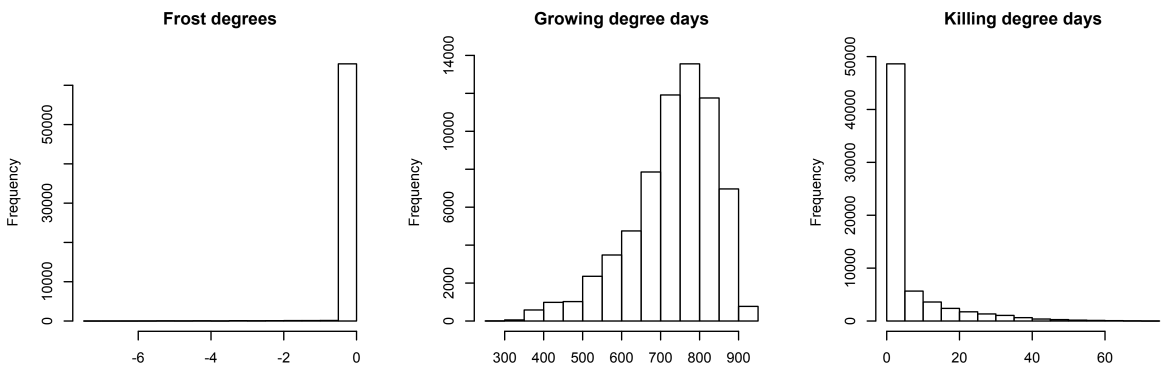


Figure 7: Growing degree day histograms, after an increase of 2°C.

As shown in figure 8, the observed yields show wide variation. The blue distribution is shifted to the left, eliminating some of the most spectacular yields and lowering the average yield. The average yield in the warmer experiment is about 80% of the original yields (see figure 9).

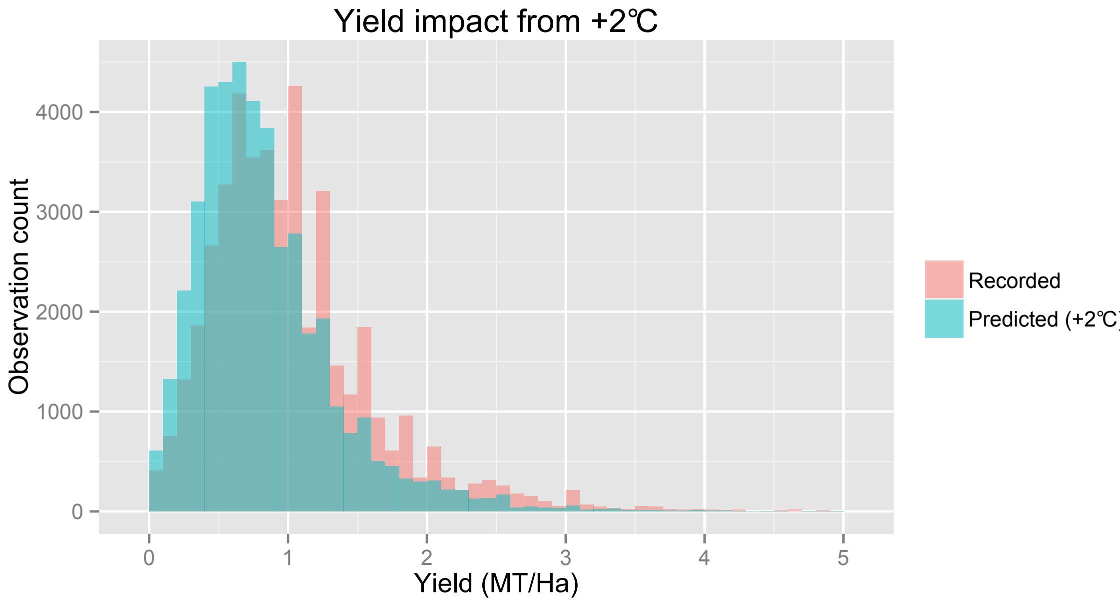


Figure 8: Observed yields over the period from 1990 - 2015 are shown in red, and model predictions under weather with temperatures increased by 2°C shown in blue.

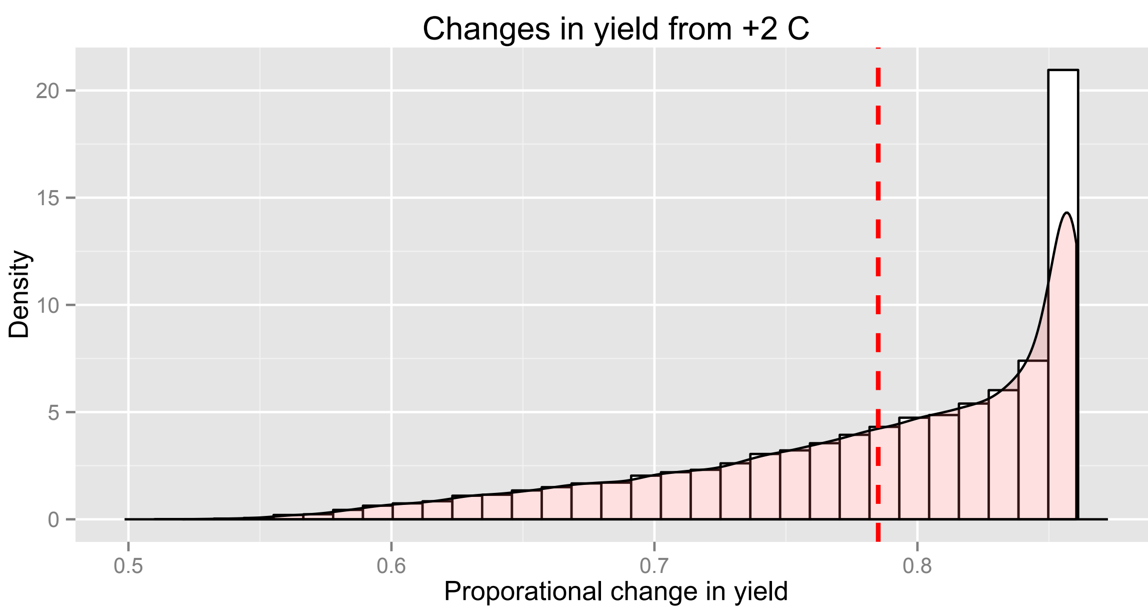


Figure 9: Distribution of the proportional change in yields, with a mean yield 79% of historical yields.

4 Global production

In this section, we estimate the a model like the one for Brazil for all countries. Using the intra-year production estimates in the coffee database, we estimate the relationship between country yields and weather. We use the temperature span of 0°C to 33°C for growing-degree days, as estimated for Brazil.

The first estimate is exactly analogous to the Brazil estimate, in that a single coefficient is estimated across all countries for the global average effect of GDDs, KDDs, frost degrees, and quadratic precipitation. This is reported in table 4 and shown schematically in figure 10.

	Log Yield	Production
GDD / 1000	0.238**	1,710.548
	(0.119)	(5,917.907)
KDD / 1000	−1.935	−3,955.098
	(1.786)	(43,378.870)
Frost Deg.	−0.005	284.772
	(0.008)	(1,550.480)
Year Precip	−3.454	707,932.900
	(12.928)	(483,967.400)
Year Precip^2	14.494	−10,991,768.000
	(135.355)	(6,955,772.000)
FE	*Region, variety*	*RegionVariety*
Trends	*Y*	*Y*
Errors	*Region*	*Region*
Observations	1,945	1,945
R^2	0.684	0.807
Adjusted R^2	0.676	0.802
Residual Std. Error (df = 1896)	0.441	33,325.380
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 4: Growing degree day model, pooled across all countries.

4.1 Hierarchical model framework

It is reasonable to expect different countries to have different effects from temperatures. We could estimate each country independently, and this would be an "unpooled" model. However, we also want the model for one country to inform, to an extent supported by the data, the model for another country. To capture this, we will construct a "hierarchical model", where each country's sensitivity to temperature will be drawn from a common distribution, simultaneously estimating each country's parameters and the distribution across all of them.

Furthermore, we allow varieties in different regions to operate differently, as supported by the data. For example, where plentiful data supports a higher optimal growing temperature for Robusta, the model should represent this. If very little data is available, the predicted response should by default conform to an average for that region and variety. Finally, we want to incorporate higher resolution data where it is available. The municipality data in Brazil informs the same common parameters as the Brazil-specific country-level yield data.

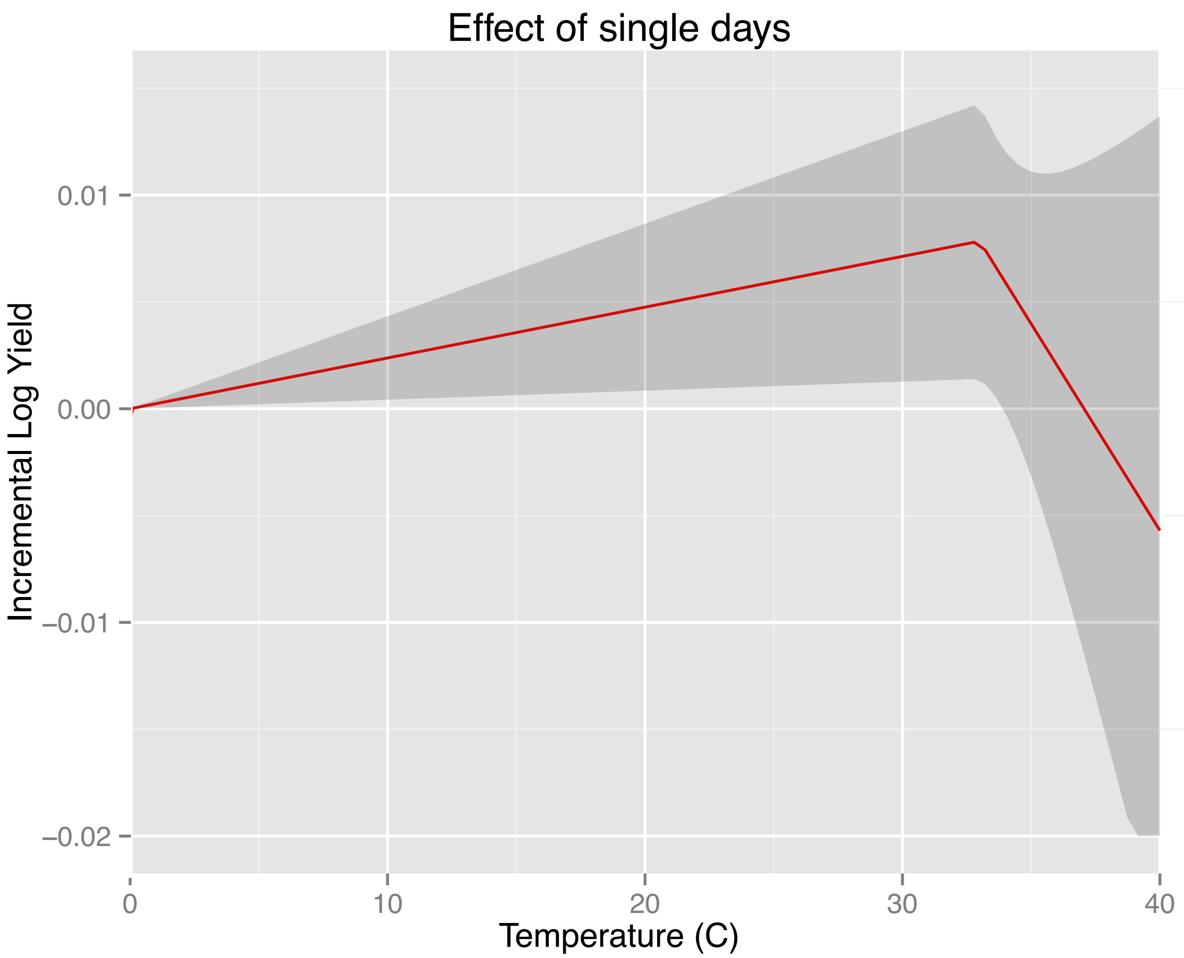


Figure 10: Pooled model growing degree-day plot.

We have developed a technique for allowing this kind of data-driven multiple levels of aggregation and degrees of generalization, based on Bayesian Hierarchical Modeling (Gelman et al., 2014) and Inversion Theory (Menke, 2012). Under this technique, each country and sub-country region has its own parameters, but the parameters are further modeled as being related to each other. The hierarchical model is a direct extension of the statistical production model, which can be thought of as many different production models combined together.

Derivation of the hierarchical modeling system

Formally, we want to allow each variety in each country to have its own model, consisting of coefficients for growing degree-days, killing degree-days, average minimum temperature, and precipitation. The pooled model is as follows:

$$\log y_{it} = \alpha_i + \beta_v + \gamma g_{it} + \kappa k_{it} + \phi f_{it} + \pi p_{it} + \psi p_{it}^2 + \epsilon_{it}$$

while the partially pooled model starts with the unpooled relationship,

$$\log y_{ivt} = \alpha_i + \beta_v + \gamma_{iv} g_{it} + \kappa_{iv} k_{it} + \phi_{iv} f_{it} + \pi_{iv} p_{it} + \psi_{iv} p_{it}^2 + \epsilon_{ivt}$$

Consider the GDD coefficient for country i and variety v, γ_{iv}. To partial pool across countries for a given variety, this coefficient comes from a distribution of possible coefficient values, characterized by an unknown mean and standard deviation for that variety:

$$\gamma_{iv} \sim \mathcal{N}(\gamma_v, \tau_{\gamma_v})$$

Further, we partially pool these 'hyperparameters' as coming from a distribution across all varieties:

$$\gamma_v \sim \mathcal{N}(\gamma, \tau_\gamma)$$

We apply this for each parameter, $\gamma, \kappa, \phi, \pi, \psi$.

Estimating a partially-pooled model

Computationally, estimating this form of model can be very difficult. We construct an innovative framework for doing this using Ordinary Least-Squares matrix algebra.

The Gaussian relationships above, such as $\gamma_{iv} \sim \mathcal{N}(\gamma_v, \tau_{\gamma_v})$, are mathematically equivalent to the OLS-style relationship,

$$\gamma_{iv} = \gamma_v + \tau_{gamma_v} \eta \text{ with } \eta \sim \mathcal{N}(0, 1)$$

Under OLS, error terms are members of a Gaussian distribution, $\epsilon_i \sim \mathcal{N}(0, \sigma_e^2)$. We represent the hyper-model for the γ coefficient with the OLS-style relationships

$$\begin{aligned} \gamma_{iv} &= \gamma_v + \epsilon_{iv} \\ \gamma_a &= \gamma_c + \epsilon_a \\ \gamma_r &= \gamma_c + \epsilon_r \end{aligned}$$

and similarly for the other coefficients. It is then possible to rewrite these and the original unpooled relationship to take the same form, with the same complete set of coefficients:

$$
\begin{array}{rlllr}
\log y_{ivt} & = \alpha_i & +\gamma_{iv} g_{it} & & +\cdots \\
\hline
\log y_{ivt} & = \sum_j \alpha_j \mathbf{1}_{j=i} & + \sum_{ju} \gamma_{ju} g_{it} \mathbf{1}_{ju=iv} & +\gamma_a 0 + \gamma_r 0 + \gamma_c 0 & +\cdots \\
\hline
0 & = \sum_j \alpha_j 0 & + \sum_{ju} \gamma_{ju} \mathbf{1}_{ju=1a} & -\gamma_a 1 - \gamma_r 0 - \gamma_c 0 & +\cdots \\
0 & = \sum_j \alpha_j 0 & + \sum_{ju} \gamma_{ju} \mathbf{1}_{ju=2a} & -\gamma_a 1 - \gamma_r 0 - \gamma_c 0 & +\cdots \\
 & & \vdots & & \\
0 & = \sum_j \alpha_j 0 & + \sum_{ju} \gamma_{ju} \mathbf{1}_{ju=1r} & -\gamma_a 0 - \gamma_r 1 - \gamma_c 0 & +\cdots \\
 & & \vdots & & \\
0 & = \sum_j \alpha_j 0 & + \sum_j u \gamma_{ju} \mathbf{1}_{ju=1c} & -\gamma_a 0 - \gamma_r 0 - \gamma_c 1 & +\cdots \\
 & & \vdots & & \\
0 & = \sum_j \alpha_j 0 & + \sum_j u \gamma_{ju} 0 & +\gamma_a 1 + \gamma_r 0 - \gamma_c 1 & +\cdots \\
0 & = \sum_j \alpha_j 0 & + \sum_j u \gamma_{ju} 0 & +\gamma_a 0 + \gamma_r 1 - \gamma_c 1 & +\cdots
\end{array}
$$

The first line is the start of the original model to be estimated. The second line re-writes this with more systematically, and in such a way that "constant" terms can be set to zero for fictional observations. The remaining lines are fictional observations added to estimate the entire model.

We have built this approach into a tool for the R statistical package which is available at `https://github.com/eicoffee/hierlm`.

Figure 11 shows the effects of partial pooling at different levels. As the level of pooling increases, the range of country-specific values is brought closer together.

The results are shown in table 5. Only the hyperparameter means are shown. Each statistically significant country coefficient is listed in Appendix .4, and the remainder are in an online table at `http://eicoffee.net/`. The first column uses only observations at the country level. The second column places a prior on the Brazil coefficients, conforming to the Brazil municipality estimates above. These more-precise estimates then inform the global distribution for each coefficient, which in turn informs all of the countries, including Brazil.

We now extend the analysis to the global context. To do so, we develop a new technique, which allows each country to have its own model, but for all of these models to be informed by each other. Within each country, Arabica and Robusta are assumed to have distinct model parameters, where the data permits.

The growing degree-day effect is greater for the Robusta variety, while the estimated effect of killing degree-days is greatest in regions that grow both Arabica and Robusta. A useful metric is the "break-even" temperature, the crossing point for which lower temperatures improve yields, on average, and higher temperatures depress yields. This is similar for Arabica and combined countries, at about 37°C. The corresponding temperature for Robusta is 40.5°C.

Figure 12 shows the variation across countries of the effect of killing degree-days and the break-even temperature. South America and Southern Africa show the least sensitivity to temperature, while Indonesia and the lands near it show the most.

Another way to view these results is to compare the break-even temperature to the average maximum daily temperature. This is shown in figure 13. Countries in the lower-left corner have both low maximum temperatures and low break-even temperatures. Since their estimated break-even temperature is over 10°C above their average maximums, these countries are at lower risk of losing their harvests. Countries in the lower-left corner are the most at risk: they have higher temperatures, but remain sensitive to high temperatures. Three countries occupy the upper-right corner: Liberia, Nigeria, and Guinea. These coun-

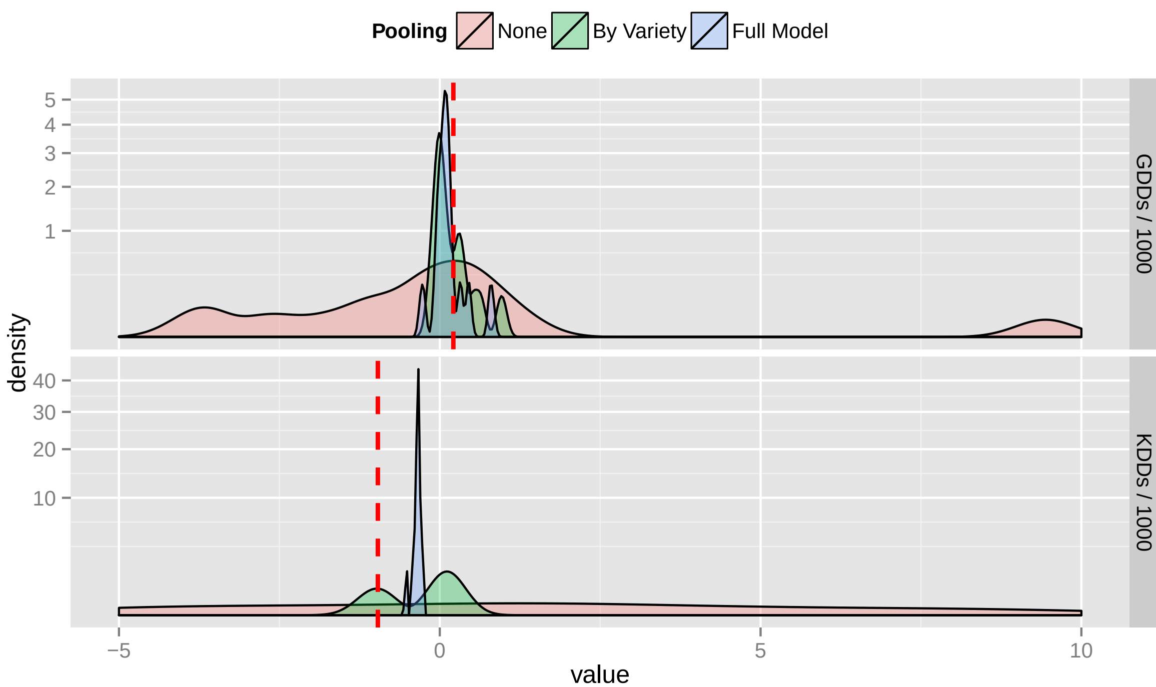


Figure 11: Distribution across countries of values for the GDD and KDD coefficients for different levels of pooling.

	Dependent variable:	
	Countries only	
	(1)	(2)
GDDs / 1000 (Combined)	0.079	0.217**
	(0.123)	(0.095)
GDDs / 1000 (Arabica)	0.131	0.229**
	(0.112)	(0.103)
GDDs / 1000 (Robusta)	0.161	0.401***
	(0.152)	(0.133)
KDDs / 1000 (Combined)	−0.110	−1.801***
	(0.543)	(0.323)
KDDs / 1000 (Arabica)	−0.082	−1.731***
	(0.556)	(0.356)
KDDs / 1000 (Robusta)	−0.157	−1.766***
	(0.543)	(0.348)
Avg. Min. (Combined)	−0.077	−0.108
	(6.344)	(6.248)
Avg. Min. (Arabica)	−0.134	−0.152
	(7.147)	(7.164)
Avg. Min. (Robusta)	−0.114	−0.163
	(8.964)	(8.985)
Precip. (Combined)	−4.285	−2.124
	(5.792)	(2.390)
Precip. (Arabica)	−1.689	−0.156
	(6.058)	(3.254)
Precip. (Robusta)	−1.565	−0.279
	(5.971)	(3.403)
Precip.^{2}(Combined)	5.340	−5.530
	(82.317)	(28.605)
Precip.^{2}(Arabica)	21.749	11.218
	(79.174)	(37.825)
Precip.^{2}(Robusta)	12.794	0.264
	(88.198)	(42.271)
Observations	3,011	3,016
R^2	0.902	0.903
Adjusted R^2	0.885	0.886
Residual Std. Error	0.335 (df = 2561)	0.336 (df = 2566)
F Statistic	52.575*** (df = 450; 2561)	52.962*** (df = 450; 2566)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 5: Hierarchical model results, for the mean of the global distribution of coefficients for each parameter and each variety.

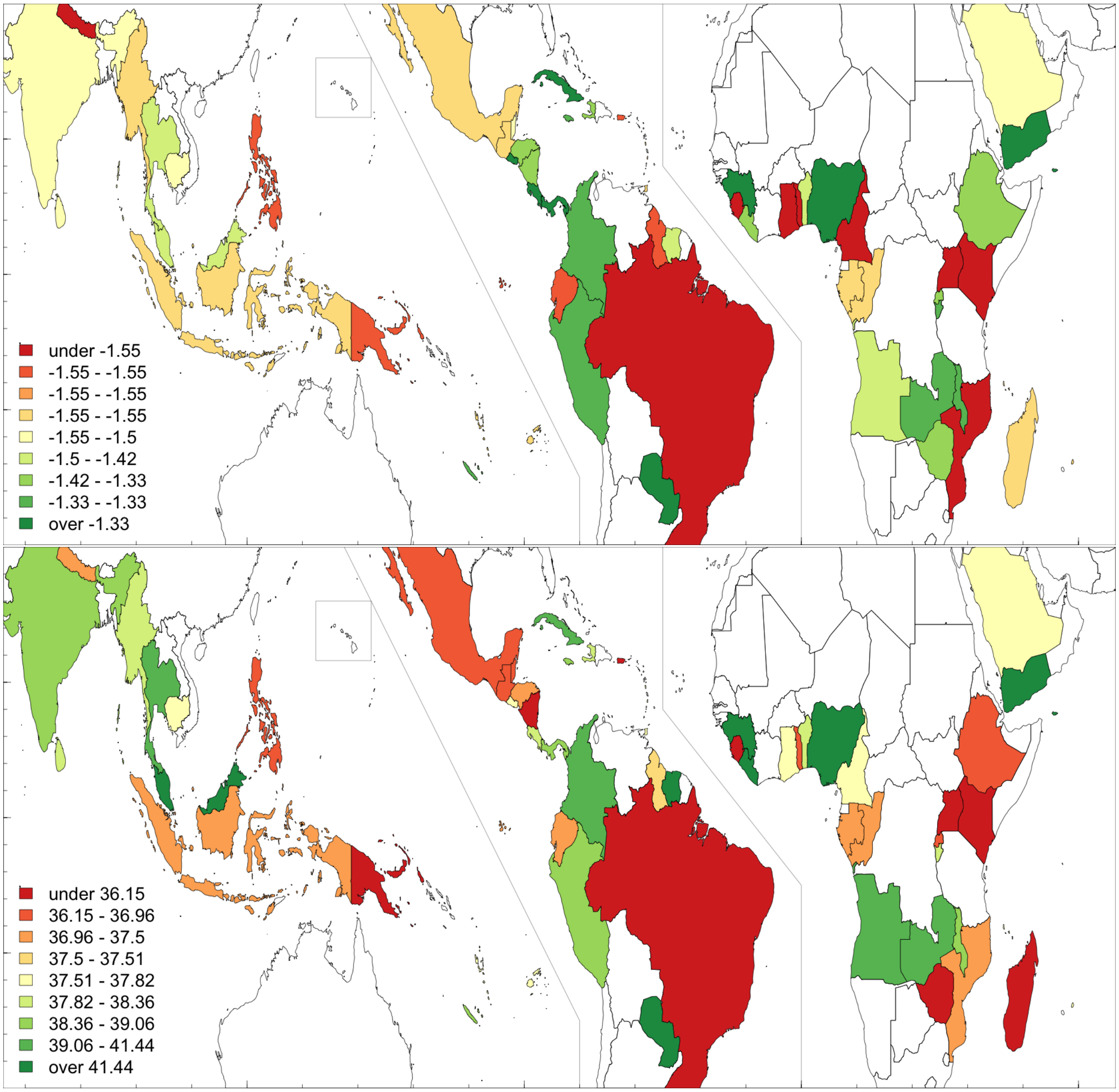


Figure 12: Coefficients of killing degree-days and the temperature at which yields decrease, across countries for the partially pooled model.

tries are estimated to have the highest temperature thresholds for yield losses. They are also amongst the countries with the highest maximum temperatures, averaged over their coffee growing regions, suggesting that these countries have achieved some level of adaptation to their high temperatures.

Other countries with similarly high temperatures have much lower break-even temperatures. That is, adaptation to high temperatures does not come inevitably from the experience of them. One explanation for their low temperature sensitivity is that Liberia, Nigeria, and Guinea all produce largely Robusta crops, which has lower sensitivity than Arabica. However, Benin, Togo, and Sri Lanka, with among the highest temperatures, also mainly produce Robusta, and they remain very sensitive to temperature.

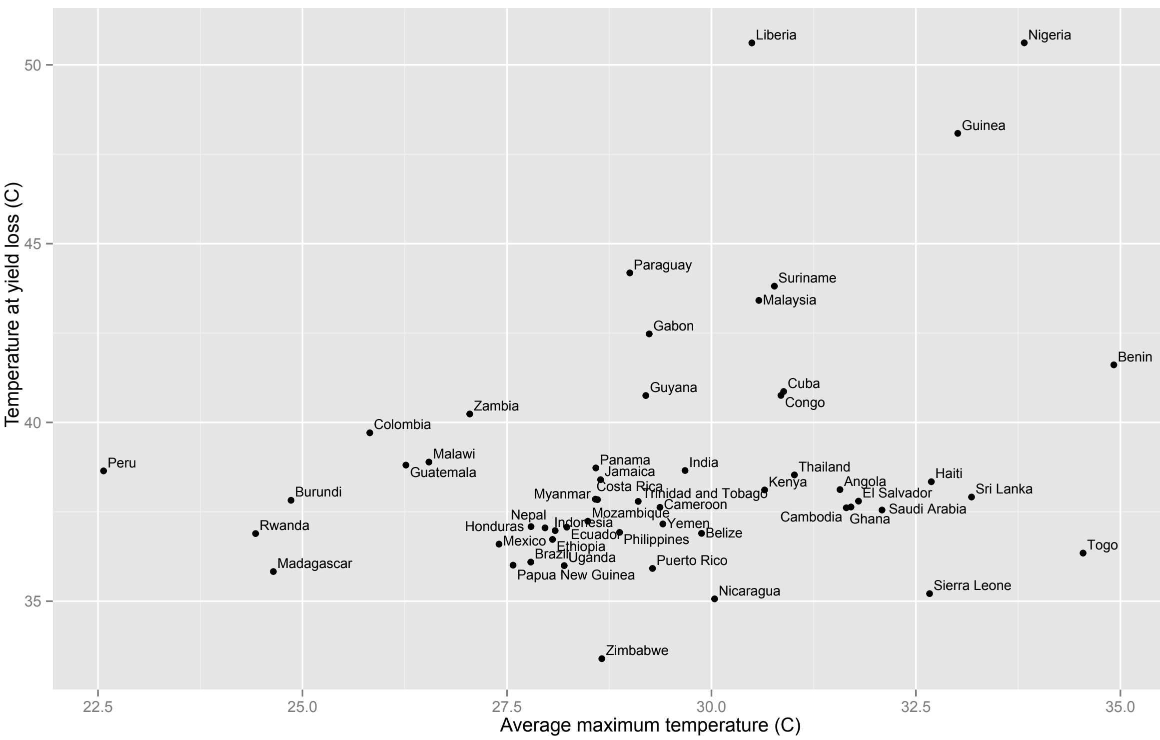


Figure 13: Observed average maximum daily temperature, 2004 - 2009, compared to the temperature at which yield losses are predicted.

4.2 Future productivity

We can use the global hierarchical model to predict yields under the future climate as described by global climate models. We shift temperatures according to the change in the average temperature and apply the average proportional change in precipitation to all daily weather observations from the CFSR dataset.[3]

[3] In section 3.6, we perform a similar analysis for the model of Brazilian yields. There, we leave precipitation unchanged, because of the lack of evidence of trends in precipitation in Brazil. However, agreement across climate models on predicted precipitation varies across countries. Here we include the effect of predicted precipitation changes to most accurately represent the predicted climate of 2050. The effects of climate change on precipitation remain very uncertain, despite occasional agreement across models, and the predicted patterns shown in section **??** should not be construed as confident forecasts.

Then we calculate GDDs, KDDs, average minimum temperatures, and total precipitations, and apply them to the model. Figure 14 shows the result.

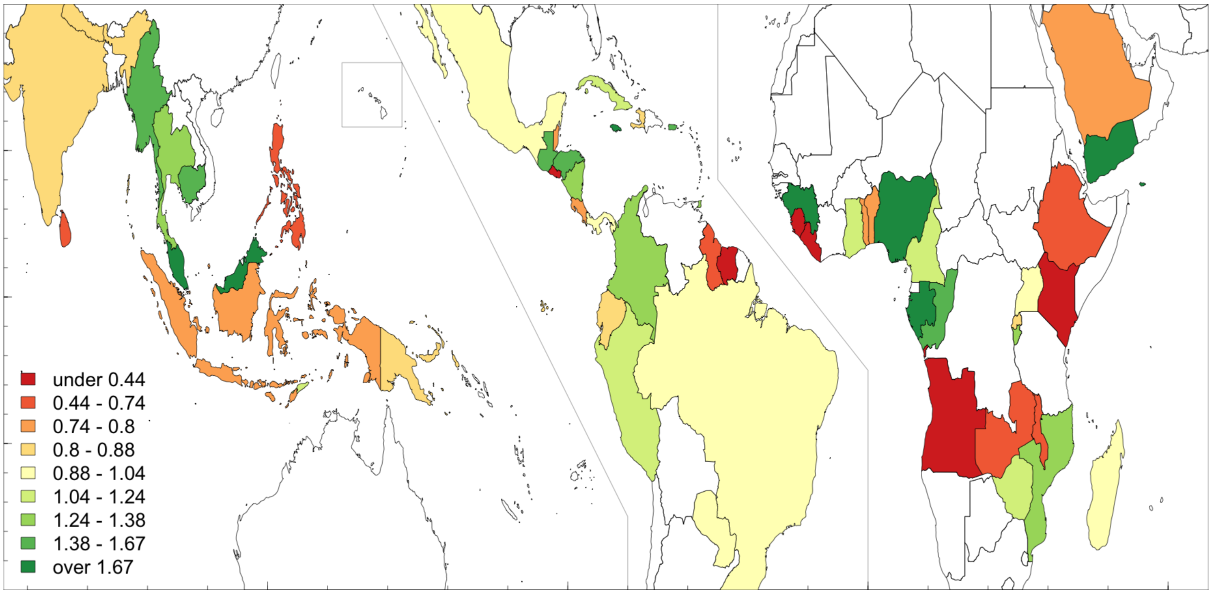


Figure 14: Changes in yield by country, for weather averaged over growing regions for each country.

As shown in the map, the impacts vary widely across countries, with some countries losing as much as 70% of their productivity, while others see increases of over 60%. Most areas in South America will experience improvements, while many countries in Central America, Southern and Eastern Africa, and Eastern Oceania will experience losses.

This result shows some general features about the variation across countries, but the actual country predictions have low confidence. In many cases, these country predictions are based on few data points, and the global distribution used to inform all of them is broad because of the uncertainty of predicting country aggregated yields. See appendix 4 for more information.

4.3 Humidity

Humidity can have varying effects on coffee. The plant needs reasonably high levels of humidity during the flowering season to avoid floral atrophy, but humidity is also crucial to the development of coffee rust. For these reasons, the timing of high humidity levels appears to be particularly important. Here we see how Arabica coffee yields respond to a one-standard deviation increase in humidity during each particular month in the year leading up to harvest. Robusta appears to be less sensitive to humidity effects than Arabica.

Humidity data is from the NCEP CFSR. The reanalysis data is available at 1/12° resolution globally, which is then aggregated to the country-month level using weights from the coffee database. The values are reported as specific humidity at 6 hour intervals, which here is averaged over each month for the year prior to harvest.

Monthly effects of humidity are shown in figure 15, and the table of coefficients is in Appendix .2. The

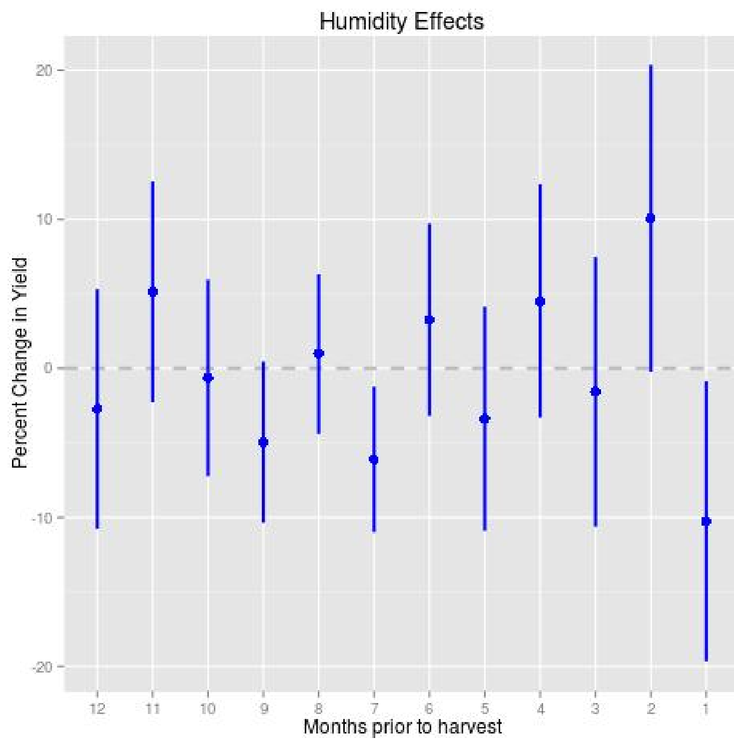


Figure 15: Arabica humidity effects. Only the humidity one and seven months before harvest are significant at 95% confidence.

coefficients result from the following model:

$$log(y) = f(T) + \sum_{m=1}^{12} \beta_m q_m + h_c(t) + \alpha_c + \gamma_t + \epsilon_{ct}$$

where $f(T)$ is a non-linear function of temperature, estimated using the number of days spent in 1-degree C temperature bins, $h_c(t)$ is a country-specific linear time trend, α_c and γ_t are country and year fixed-effects. Each β_m is the effect of specific humidity m months prior to the beginning of harvest on log yield.

4.4 Interpreting empirical model results

Climate change impacts coffee production through many different channels. Foremost, climate change reflects changes in temperature and patterns of precipitation– that is, changes in climate mean changes in weather. The models above estimate the relationship between changes in weather and changes in yields, and then extrapolate those changes to their responses under climate change.

There are important differences between unexpected weather shocks and prolonged climate changes. Coffee farming will find ways to adapt to repeated shocks of higher temperatures, and we hope our estimates provide an upper bound on the production impacts of climate change. However, the evidence for such adaptation is limited. Burke and Emerick (2012) study maize in the United States, and while there is a clear potential for adaptation to warmer temperatures, they find almost no evidence of it. The reasons for this empirical result are unclear.

The effects that we measure of temperature on yields cannot be unambiguously interpreted as the biological response to temperatures. Temperatures could be simultaneously affecting other species that then affect coffee. For example, the harmful affects of average minimum temperature could reflect a greater capacity for coffee rust or the coffee berry borer to proliferate in these warmer years. It could also reflect decreased activity on the part of farmers on hot days.

Our results should be taken as representing a holistic effect as it has occurred in the past. The extent to which it will occur in the future may be up to us.

.1 Selecting temperature limits

.2 Humidity

.3 Harvest month effects

Figure 16 shows the estimated "effect" of harvesting in a given month on yields, from 1962 to 2011, after accounting for country-specific and monthly effects. The gradual increase reflects improvements in coffee production technology, but this increase is not without large shocks. An increase in yields between 1985 and 1990 was followed by a decrease and then another period of increased yields. Countries that harvest in different months also show different fortunes, with the greatest yields to countries that harvest in January and the lowest to those that harvest in February. Since the only country that harvests in January but not February is Colombia, this probably reflects the difference between Colombia yields and yields in other February-harvesting countries.

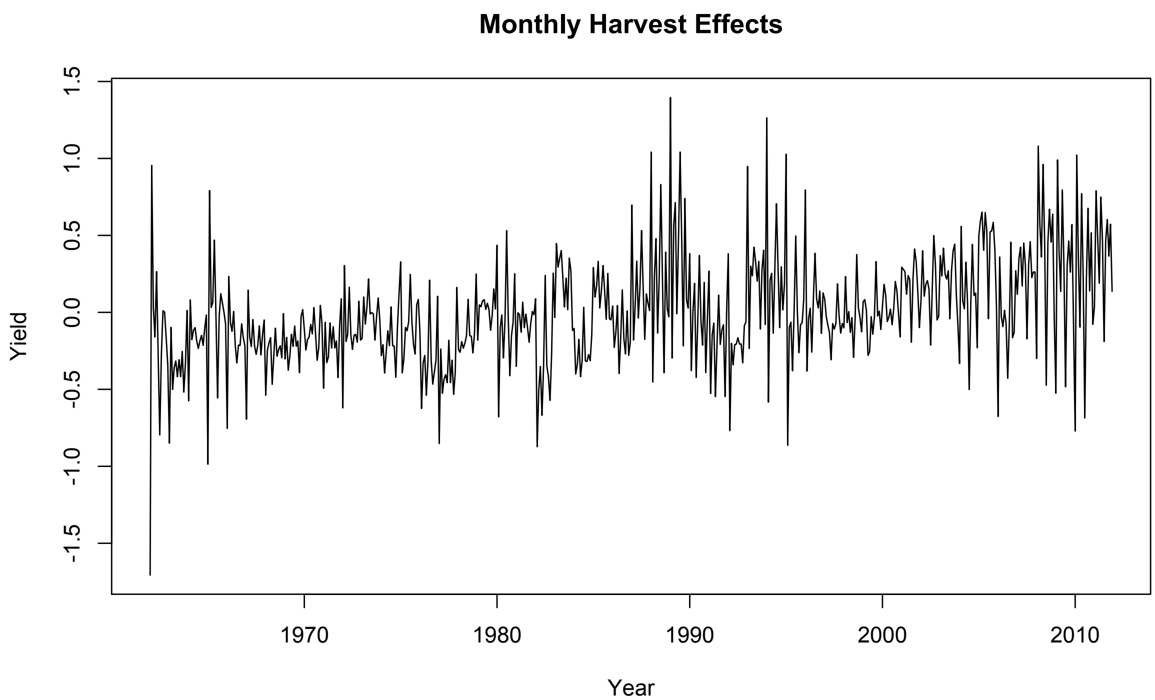


Figure 16: Monthly harvesting effects. Each point on this curve represents the difference in yields predicted by harvesting in a given month, according to coffee harvest calendars, after accounting for country-specific and month effects. Uses calendars from `https://www.sweetmarias.com/coffee.prod.timetable.php`

Low \High	28	29	30	31	32	33	34
-4					87.4211	87.4289	87.3933
-3							
-2						87.4290	
-1							
0				87.2986	87.4213	87.4290	87.3934
1						87.4290	
2					87.4212	87.4289	87.3933
3						87.4288	
4				87.2983	87.4210	87.4286	
5			87.0758	87.2979	87.4206		
6				87.2978	87.4205	87.4281	
7			87.0755	87.2978	87.4204		
8				87.2981			
9			87.0749	87.2979	87.4199		
10	86.7490		87.0737	87.2975			
11			87.0729	87.2975	87.4182		
12	86.7398		87.0700	87.2954			
13							
14							
15	86.6988		87.0369	87.2645			

Table 6: F-statistics for a growing degree-day and killing degree-day model of coffee production, across all countries. The highest F-stats use a maximum temperature of 30°C and a minimum temperature between -3°C and 1 °C.

.4 Hierarchical model coefficients

Only statistically-significant coefficients are listed below. The remaining are available online at `http://eicoffee.net/`.

	Dependent variable:	
	Countries only	
	(1)	(2)
GDDs / 1000, Liberia (Robusta)	0.515**	0.743***
	(0.213)	(0.202)
GDDs / 1000, Gabon (Robusta)	0.223	0.448**
	(0.215)	(0.204)
GDDs / 1000, Yemen (Arabica)	0.274	0.368**
	(0.189)	(0.183)
GDDs / 1000, Benin (Robusta)	0.146	0.409**
	(0.221)	(0.207)
GDDs / 1000, Cuba (Arabica)	0.222	0.322*
	(0.194)	(0.189)
GDDs / 1000, Angola (Robusta)	0.121	0.354*
	(0.217)	(0.205)
GDDs / 1000, Malaysia (Robusta)	0.266	0.495**
	(0.220)	(0.209)
GDDs / 1000, Brazil (Combined)	0.079	0.158***
	(0.208)	(0.052)

	Dependent variable:	
	Countries only	
	(1)	(2)
GDDs / 1000, Guinea (Robusta)	0.356*	0.603***
	(0.199)	(0.185)
GDDs / 1000, Nigeria (Robusta)	0.377*	0.659***
	(0.212)	(0.197)
GDDs / 1000, Suriname (Combined)	0.346*	0.484**
	(0.204)	(0.189)
GDDs / 1000, Zambia (Arabica)	0.217	0.300*
	(0.178)	(0.173)
GDDs / 1000, Paraguay (Arabica)	0.248	0.405***
	(0.165)	(0.156)
GDDs / 1000, Guyana (Robusta)	0.140	0.374*
	(0.223)	(0.211)
GDDs / 1000, Congo (Robusta)	0.145	0.382*
	(0.215)	(0.203)
KDDs / 1000, Cambodia (Combined)	−0.112	−1.798***
	(0.567)	(0.363)
KDDs / 1000, Ethiopia (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Cameroon (Combined)	−0.111	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Ghana (Robusta)	−0.180	−1.787***
	(0.568)	(0.386)
KDDs / 1000, Saudi.Arabia (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Guatemala (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Guatemala (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Dominica (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Liberia (Robusta)	−0.123	−1.732***
	(0.568)	(0.386)
KDDs / 1000, Gabon (Robusta)	−0.155	−1.764***
	(0.568)	(0.386)
KDDs / 1000, Gabon (Combined)	−0.110	−1.800***
	(0.568)	(0.364)
KDDs / 1000, Yemen (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Yemen (Arabica)	−0.078	−1.728***
	(0.581)	(0.394)
KDDs / 1000, Jamaica (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Samoa (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Kenya (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)

	Dependent variable:	
	Countries only	
	(1)	(2)
KDDs / 1000, Kenya (Combined)	−0.114	−1.804***
	(0.568)	(0.364)
KDDs / 1000, India (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Saint.Lucia (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Rwanda (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Peru (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Vanuatu (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Malawi (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Benin (Robusta)	−0.156	−1.754***
	(0.565)	(0.384)
KDDs / 1000, Benin (Combined)	−0.114	−1.773***
	(0.559)	(0.358)
KDDs / 1000, Cuba (Arabica)	−0.076	−1.725***
	(0.581)	(0.394)
KDDs / 1000, Togo (Robusta)	−0.244	−1.827***
	(0.560)	(0.380)
KDDs / 1000, Tonga (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Indonesia (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Mauritius (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Angola (Combined)	−0.109	−1.799***
	(0.568)	(0.364)
KDDs / 1000, Angola (Robusta)	−0.159	−1.768***
	(0.568)	(0.386)
KDDs / 1000, Trinidad.and.Tobago (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Nicaragua (Arabica)	−0.084	−1.733***
	(0.581)	(0.394)
KDDs / 1000, Malaysia (Robusta)	−0.159	−1.768***
	(0.568)	(0.386)
KDDs / 1000, Mozambique (Combined)	−0.111	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Uganda (Combined)	−0.111	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Brazil (Combined)	−0.110	−1.971***
	(0.568)	(0.309)
KDDs / 1000, Guinea (Robusta)	−0.101	−1.703***
	(0.566)	(0.384)

	Dependent variable:	
	Countries only	
	(1)	(2)
KDDs / 1000, Panama (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Costa.Rica (Arabica)	−0.081	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Nigeria (Robusta)	−0.085	−1.674***
	(0.562)	(0.382)
KDDs / 1000, Ecuador (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, El.Salvador (Arabica)	−0.081	−1.729***
	(0.581)	(0.393)
KDDs / 1000, Puerto.Rico (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Thailand (Combined)	−0.109	−1.796***
	(0.567)	(0.363)
KDDs / 1000, Thailand (Robusta)	−0.165	−1.773***
	(0.568)	(0.386)
KDDs / 1000, Haiti (Arabica)	−0.087	−1.731***
	(0.580)	(0.393)
KDDs / 1000, Belize (Combined)	−0.110	−1.799***
	(0.568)	(0.364)
KDDs / 1000, Sierra.Leone (Robusta)	−0.239	−1.835***
	(0.563)	(0.383)
KDDs / 1000, Philippines (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Timor.Leste (Combined)	−0.109	−1.799***
	(0.568)	(0.364)
KDDs / 1000, Colombia (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Burundi (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Burundi (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Fiji (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Madagascar (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Nepal (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Suriname (Combined)	−0.089	−1.779***
	(0.568)	(0.364)
KDDs / 1000, Zambia (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Papua.New.Guinea (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Zimbabwe (Arabica)	−0.094	−1.742***
	(0.581)	(0.393)

	Dependent variable:	
	Countries only	
	(1)	(2)
KDDs / 1000, New.Caledonia (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, New.Caledonia (Arabica)	−0.082	−1.731***
	(0.581)	(0.394)
KDDs / 1000, Paraguay (Arabica)	−0.043	−1.660***
	(0.571)	(0.387)
KDDs / 1000, Guyana (Robusta)	−0.157	−1.766***
	(0.568)	(0.386)
KDDs / 1000, Guyana (Arabica)	−0.081	−1.730***
	(0.581)	(0.394)
KDDs / 1000, Guyana (Combined)	−0.111	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Honduras (Arabica)	−0.084	−1.733***
	(0.581)	(0.394)
KDDs / 1000, Myanmar (Combined)	−0.110	−1.800***
	(0.568)	(0.364)
KDDs / 1000, Mexico (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
KDDs / 1000, Congo (Robusta)	−0.165	−1.773***
	(0.568)	(0.386)
KDDs / 1000, Congo (Combined)	−0.111	−1.800***
	(0.568)	(0.364)
KDDs / 1000, Sri.Lanka (Combined)	−0.108	−1.793***
	(0.567)	(0.363)
KDDs / 1000, Comoros (Combined)	−0.110	−1.801***
	(0.568)	(0.364)
Avg. Min., Liberia (Robusta)	−0.817***	−0.873***
	(0.141)	(0.140)
Avg. Min., Gabon (Robusta)	−0.642***	−0.714***
	(0.181)	(0.180)
Avg. Min., Yemen (Combined)	0.402*	0.369*
	(0.208)	(0.207)
Avg. Min., Jamaica (Arabica)	0.297**	0.269**
	(0.117)	(0.116)
Avg. Min., Kenya (Arabica)	−1.325***	−1.369***
	(0.288)	(0.288)
Avg. Min., Kenya (Combined)	−0.755***	−0.794***
	(0.156)	(0.154)
Avg. Min., Malawi (Arabica)	−0.255*	−0.288**
	(0.141)	(0.140)
Avg. Min., Angola (Combined)	−0.367*	−0.400**
	(0.199)	(0.198)
Avg. Min., Angola (Robusta)	0.218**	0.178*
	(0.110)	(0.108)
Avg. Min., Malaysia (Robusta)	2.766***	2.680***
	(0.164)	(0.162)

	Dependent variable:	
	Countries only	
	(1)	(2)
Avg. Min., Brazil (Combined)	−0.077	−0.091***
	(34.091)	(0.020)
Avg. Min., Guinea (Robusta)	0.329***	0.300**
	(0.125)	(0.124)
Avg. Min., El.Salvador (Arabica)	−0.525***	−0.526***
	(0.145)	(0.145)
Avg. Min., Sierra.Leone (Robusta)	−1.196***	−1.197***
	(0.147)	(0.147)
Avg. Min., Suriname (Combined)	−1.532***	−1.559***
	(0.174)	(0.172)
Avg. Min., Zambia (Arabica)	−0.204	−0.237*
	(0.125)	(0.123)
Avg. Min., Congo (Robusta)	−1.389***	−1.441***
	(0.155)	(0.154)
Avg. Min., Sri.Lanka (Combined)	−0.426***	−0.402***
	(0.140)	(0.139)
Precip., Brazil (Combined)	−4.285	0.347***
	(6.691)	(0.030)
Precip., Suriname (Combined)	−12.378*	−10.271**
	(6.552)	(4.007)
Precip.^{2}, Brazil (Combined)	5.340	0.366***
	(88.871)	(0.039)
Observations	3,011	3,016
R^2	0.902	0.903
Adjusted R^2	0.885	0.886
Residual Std. Error	0.335 (df = 2561)	0.336 (df = 2566)
F Statistic	52.575*** (df = 450; 2561)	52.962*** (df = 450; 2566)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 7: Humidity Effects

	Dependent variable:
Month prior to harvest	log(yield)
1	−8.562
	(12.703)
2	17.386
	(12.509)
3	−2.607
	(13.406)
4	−23.317*
	(12.756)
5	4.781
	(13.223)
6	−30.035**
	(12.008)
7	15.021
	(14.797)
8	−14.813
	(16.496)
9	19.024
	(16.429)
10	−6.111
	(17.747)
11	35.636*
	(18.228)
12	−33.730**
	(15.444)
Observations	738
R^2	0.895
Adjusted R^2	0.881
Residual Std. Error	0.191 (df = 653)
F Statistic	66.164*** (df = 84; 653)
Note:	*p<0.1; **p<0.05; ***p<0.01

References

Bernardes, T., Moreira, M. A., Adami, M., Giarolla, A., and Rudorff, B. F. T. (2012). Monitoring biennial bearing effect on coffee yield using modis remote sensing imagery. *Remote Sensing*, 4(9):2492–2509.

Burke, M. and Emerick, K. (2012). Adaptation to climate change: Evidence from us agriculture. *Available at SSRN 2144928*.

Dauzat, J., Griffon, S., Roupsard, O., Vaast, P., and Rodrigues, G. (2014). Building the foundations of a coffea arabica fspm. In *Embrapa Informática Agropecuária-Resumo em anais de congresso (ALICE)*. In: INTERNATIONAL CONFERENCE ON FUNCTIONAL-STRUCTURAL PLANT MODELS, 7., 2013, Saariselkä. Proceedings... Vantaa: Finnish Society of Forest Science, 2013.

Gay, C., Estrada, F., Conde, C., Eakin, H., and Villers, L. (2006). Potential impacts of climate change on agriculture: A case of study of coffee production in veracruz, mexico. *Climatic Change*, 79(3-4):259–288.

Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2014). *Bayesian data analysis*, volume 2. Taylor & Francis.

Guzmán Martínez, O., Jaramillo Robledo, A., and Baldión Rincón, J. V. (1999). Anuario meteorologico cafetero, 1998.

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L., Wilkens, P. W., Singh, U., Gijsman, A. J., and Ritchie, J. T. (2003). The dssat cropping system model. *European journal of agronomy*, 18(3):235–265.

Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., Huth, N. I., Hargreaves, J. N., Meinke, H., Hochman, Z., et al. (2003). An overview of apsim, a model designed for farming systems simulation. *European Journal of Agronomy*, 18(3):267–288.

Maro, G., Mrema, J., Msanya, B., Janssen, B., and Teri, J. (2014). Developing a coffee yield prediction and integrated soil fertility management recommendation model for northern tanzania. *International Journal of Plant & Soil Science*, 3(4):380–396.

Menke, W. (2012). *Geophysical data analysis: discrete inverse theory*. Academic press.

Rodríguez, D., Cure, J. R., Cotes, J. M., Gutierrez, A. P., and Cantor, F. (2011). A coffee agroecosystem model: I. growth and development of the coffee plant. *Ecological modelling*, 222(19):3626–3639.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., et al. (2010). The ncep climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, 91(8):1015–1057.

Schlenker, W. and Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to us crop yields under climate change. *Proceedings of the National Academy of sciences*, 106(37):15594–15598.

Acknowledgments

We would like to thank **Walter Baethgen** at the International Research Institute for Climate and Society for his thoughtful reviews of the work here and all of his comments and suggestions.

Many thanks also to:

Andrea Illy, illycaffè S.p.A., Chairman/CEO

Mario Cerutti, Lavazza S.p.A., Corporate Relations

Belay Begashaw, Columbia Global Centers, Director

Mauricio Galindo, International Coffee Organization, Head of Operations

Alexandra Tunistra, Root Capital, Advisory Services

Amir Jina, University of Chicago, Economics

Joann de Zegher, Stanford University, Environment and Resources

Marion Dumas, Columbia University, School of International and Public Affairs