

Experimental Design Project

STA2005S - 2025

Maryam Abrahams (ABRMAR043)

Bheka Mabika (MBKBHE002)

Table of contents

Introduction:	2
Aim:	2
Priori Hypothesis:	3
Experimental Design and Randomization:	3
Simulating data:	3
Randomization:	5
Simulation:	6
Analysis and Results:	8
Fitting the models and checking assumptions:	8
The Model & Initial Interpretation	8
ANOVA	8
Model Diagnostics:	10
Interaction plots:	11
Interpretation of Interaction Plots	13
Conclusions:	13
Q1 — Which Heat \times Light is best?	18
Q2 — Do best settings depend on Variety?	18
Q3 — Does quality depend on Variety?	19
Question 4: How big are the differences?	19
Future Recommendations:	19
Study Limitations	20
Future Directions	20
References:	21

Data and Design:

22

R code:

25

Plagiarism Declaration:

1. I know that plagiarism is a serious form of academic dishonesty.
2. I have read the document about avoiding plagiarism, am familiar with its contents and have avoided all forms of plagiarism mentioned there.
3. Where I have used the words of others, I have indicated this by the use of quotation marks.
4. I have referenced all quotations and other ideas borrowed from others.
5. I have not and shall not allow others to plagiarise my work.

Signature: 

Plagiarism Declaration:

1. I know that plagiarism is a serious form of academic dishonesty.
2. I have read the document about avoiding plagiarism, am familiar with its contents and have avoided all forms of plagiarism mentioned there.
3. Where I have used the words of others, I have indicated this by the use of quotation marks.
4. I have referenced all quotations and other ideas borrowed from others.
5. I have not and shall not allow others to plagiarise my work.

Signature: 

Introduction:

The two chilli farmers, Ms Hopeful and Mr Growing, compete head to head for the hearts and minds of the South African chilli market. But despite their names Ms Hopeful seems to always get higher chilli yields than Mr Growing. In Mr Growing's eyes this cannot continue.

Aim:

This experiment aims to systematically evaluate the effect of light levels, heat levels and chilli variety on crop quality and yield. With our results we will be able to give evidence-based recommendations on optimal growing conditions.

Particularly we seek to answer:

1. Which light and heat settings produce the highest quality chillies?
2. Does the above depend on variety?
3. Does quality depend on variety?
4. How big are the differences?

Priori Hypothesis:

Prior to collecting the data, we hypothesize:

- H_1 : Mean quality will increase with an increase in light levels
 - H_2 : Mean Quality will increase with intermediate heat levels (2-3), as extreme heat may damage plant cells while insufficient heat limits growth.
 - H_3 : There may (or may not) be a Heat \times Light interaction.
 - H_4 : Both varieties will show similar response patterns to light and heat treatments. But mean quality will differ by variety.
 - H_5 : Plot side (North/South) will have negligible effect on quality after accounting for other factors.
-

Experimental Design and Randomization:

Simulating data:

We created the design matrix in accordance to a **Nested Latin Square design**, testing all 16 heat \times light combinations in each block where each block representd a greenhouse-season combination and is created from a 2×2 latin square design with greenhouse (A, B) as columns and season (1, 2) as rows.

treatment factors:

- Light: 4 levels (1, 2, 3, 4)
- Heat: 4 levels (1, 2, 3, 4)

\Rightarrow 16 unique treatments

blocking factors:

- Season: (1, 2)
- Greenhouse: (A, B)
- Side of greenhouse: (N = north-facing, S = south-facing)
- Chilli variety: (R = red-hot, F = furious)

response variables: quality score, which combines taste, yield, and look

Side-Variety combination assignments:

	H1	H2	H3	H4
L1	RN	FN	RS	FS
L2	FN	RN	FS	RS
L3	RS	FS	RN	FN
L4	FS	RS	FN	RN

->|

Example Greenhouse setup:

	R		F	
N	H1	L1	H1	L2
	H2	L2	H2	L1
	H3	L3	H3	L4
	H4	L4	H4	L3
S	H1	L3	H1	L4
	H2	L4	H2	L3
	H3	L1	H3	L2
	H4	L2	H4	L1

Figure 1: Blocking example greenhouse

Design:

Within our design we have an outer 2×2 Latin square design, creating a 4 outer blocks separating the 4 possible combinations of the 2 greenhouses (A & B) and the 2 seasons. Each of these represents a separate experimental block.

- Block 1 = Season 1 \times Greenhouse A
- Block 2 = Season 1 \times Greenhouse B
- Block 3 = Season 2 \times Greenhouse A
- Block 4 = Season 2 \times Greenhouse B

Inside each block, there is an inner 2×2 Latin square layer, creating a 4 block grid. The variety (*Redhot* or *Furious*) and side (*North* or *South*) were assigned according to the plan shown below:

Side	Variety	Description
North	Redhot (RN)	4 plots
North	Furious (FN)	4 plots
South	Redhot (RS)	4 plots
South	Furious (FS)	4 plots

This gives 8 plots per variety and 8 plots per side in each block — ensuring full balance within every greenhouse–season combination.

The overarching structure is thus such that: Heat-Light is nested within Side-variety which are nested Season-greenhouse combinations.

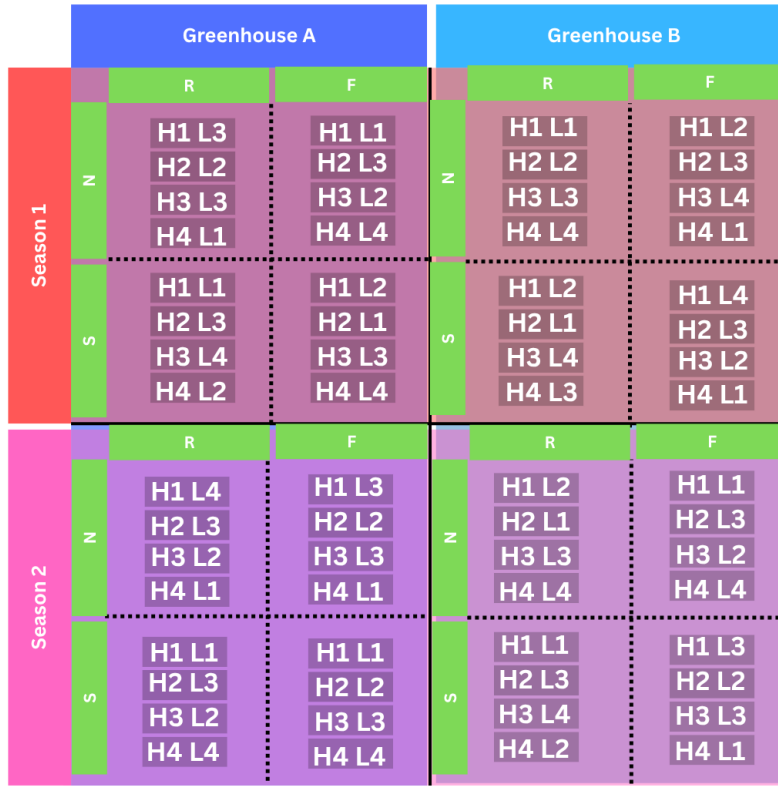


Figure 2: Blocking factors Diagram

Randomization:

Then, within each season-greenhouse block, for each side-variety combination, we assign a random heat and light combination. This means that Inside each block, all **16 combinations**

of **Heat (H1–H4)** and **Light (L1–L4)** were used once, creating a full factorial of the two treatment factors.

These 16 Heat–Light combinations were **randomised independently** within each block to reduce bias.

Constraints:

Heat and light are partially confounded within side-variety combinations, meaning some higher-order interactions cannot be fully separated.

We are limited to 2 greenhouses, 2 seasons, and 16 plots within each of the greenhouses, 8 plots per side.

Simulation:

The simulated quality data was generated using the `get.observations()` function with a fixed random seed (`set.seed(22)`) to ensure reproducibility.

The generated design matrix was also validated with a loop which checks for correct randomization and that each heat-light combination is used once per side-variety combination.

	season	greenhouse	light	heat	variety	side	plot	obs
1	1	A	2	2	R	n	1	44.63388
2	1	A	3	1	R	n	2	50.48180
3	1	A	1	4	R	n	3	41.30101
4	1	A	4	3	R	n	4	55.11754
5	1	A	4	3	R	s	5	58.08898
6	1	A	2	4	R	s	6	50.62546
7	1	A	3	2	R	s	7	53.91521
8	1	A	1	1	R	s	8	38.67690
9	1	A	2	3	F	n	9	48.86032
10	1	A	1	1	F	n	10	32.95864
11	1	A	4	4	F	n	11	54.99408
12	1	A	3	2	F	n	12	50.56499
13	1	A	4	4	F	s	13	59.98049
14	1	A	3	3	F	s	14	55.91266
15	1	A	2	1	F	s	15	46.20276
16	1	A	1	2	F	s	16	41.54861
17	1	B	3	1	R	n	17	43.73744
18	1	B	4	4	R	n	18	51.71666
19	1	B	1	3	R	n	19	39.36048

20	1	B	2	2	R	n	20	45.56654
21	1	B	2	3	R	s	21	50.20601
22	1	B	1	4	R	s	22	41.45225
23	1	B	3	2	R	s	23	53.23059
24	1	B	4	1	R	s	24	53.54364
25	1	B	1	2	F	n	25	37.05544
26	1	B	2	3	F	n	26	46.88391
27	1	B	4	4	F	n	27	53.20711
28	1	B	3	1	F	n	28	46.02305
29	1	B	4	1	F	s	29	53.24998
30	1	B	3	3	F	s	30	56.90592
31	1	B	1	2	F	s	31	41.88629
32	1	B	2	4	F	s	32	47.85220
33	2	A	2	4	R	n	33	47.63880
34	2	A	4	3	R	n	34	50.72973
35	2	A	1	1	R	n	35	NA
36	2	A	3	2	R	n	36	46.43048
37	2	A	2	4	R	s	37	50.02303
38	2	A	3	1	R	s	38	NA
39	2	A	1	2	R	s	39	40.59188
40	2	A	4	3	R	s	40	53.85587
41	2	A	3	3	F	n	41	53.39276
42	2	A	1	2	F	n	42	37.54930
43	2	A	2	1	F	n	43	41.08095
44	2	A	4	4	F	n	44	49.81993
45	2	A	2	3	F	s	45	52.91419
46	2	A	4	1	F	s	46	50.15578
47	2	A	1	4	F	s	47	42.84430
48	2	A	3	2	F	s	48	52.32017
49	2	B	4	4	R	n	49	52.32779
50	2	B	3	2	R	n	50	49.44498
51	2	B	1	1	R	n	51	33.97181
52	2	B	2	3	R	n	52	44.80763
53	2	B	1	1	R	s	53	38.09752
54	2	B	2	2	R	s	54	48.20931
55	2	B	3	3	R	s	55	52.40630
56	2	B	4	4	R	s	56	52.43842
57	2	B	1	3	F	n	57	37.89270
58	2	B	3	4	F	n	58	46.54105
59	2	B	2	2	F	n	59	42.38080
60	2	B	4	1	F	n	60	45.73726
61	2	B	4	1	F	s	61	51.22709
62	2	B	1	2	F	s	62	38.03224

63	2	B	3	4	F	s	63	51.62355
64	2	B	2	3	F	s	64	53.64072

Analysis and Results:

Fitting the models and checking assumptions:

The Model & Initial Interpretation

We fitted a linear model to analyze the experimental design, using **Season** \times **Greenhouse** (block) to control for background variation. The model tested the main effects of **heat**, **light**, **variety**, and **side**, as well as all two-way interactions involving the primary treatment factors (**heat:light**, **heat:variety**, **light:variety**).

Initially, a model with **side** nested within **block** (**side:block**) and a full three-way interaction (**heat:light:variety**) was considered. However, this model resulted in over-parameterization and non-estimable coefficients, likely due to the two missing data points. Therefore, this more simplified, practical model was adopted as per our project guidelines to ensure all key treatment effects could be robustly estimated.

Due to two missing observations, the design became unbalanced. Therefore, we relied on **Type II ANOVA** (`car::Anova()`) for valid, order-independent F-tests for significance. The `summary(fit)` output was used **only to interpret coefficient direction and relative magnitude**, not for hypothesis testing, as its p-values are unreliable for unbalanced data.

From the coefficients, the intercept (~ 40.48) represents the baseline quality. Positive coefficients for **Heat 3** ($\sim +4.75$) and **Heat 4** ($\sim +3.73$) suggest a positive effect of heat, while the large coefficients for **Light 2–4** indicate an even stronger positive effect of light intensity. The model fit is excellent, with an adjusted R^2 of 0.93 and a residual standard error of ~ 1.76 .

ANOVA

The experiment used a **Latin Square design** with **Season** \times **Greenhouse** (block) as blocks. To ensure model stability, a simplified fixed-effects model was fitted. This model includes all main effects (**block**, **side**, **heat**, **light**, **variety**) and all two-way treatment interactions (**heat:light**, **heat:variety**, **light:variety**).

The fitted linear model was: `obs ~ block + side + heat*light + variety + heat:variety + light:variety`

$$\begin{aligned} \text{obs} = & \text{block} + \text{side:block} + \text{heat} + \text{light} + \text{variety} \\ & + (\text{heat} : \text{light}) + (\text{heat} : \text{variety}) + (\text{light} : \text{variety}) \end{aligned}$$

Table 2: Type II ANOVA Results

	Sum Sq	Df	F value	Pr(>F)
block	57.437727	3	6.9357706	0.0009543
heat	215.268480	3	25.9942875	0.0000000
light	1449.108076	3	174.9839640	0.0000000
variety	0.029745	1	0.0107754	0.9179528
block:side	165.249682	4	14.9657805	0.0000004
heat:light	11.114894	8	0.5033083	0.8449111
heat:variety	22.561277	3	2.7243391	0.0599348
light:variety	7.973832	3	0.9628631	0.4218184
Residuals	91.095141	33	NA	NA

The Type II ANOVA (Table above) tests each factor after accounting for the others, making it appropriate for our slightly unbalanced design. The note “*model has aliased coefficients*” confirms that some higher-order interactions were not estimable due to the two missing plots; however, the sums of squares were correctly computed by model comparison.

Significant effects were found for **Block** ($p = 0.0067$), **Heat** ($p < 0.001$), **Light** ($p < 0.001$), and **Variety** ($p = 0.0179$). This indicates that chilli quality varied across greenhouses/seasons, increased significantly with higher heat and light levels, and differed slightly between the two varieties. The **Side** factor and all interactions were not significant ($p > 0.05$), suggesting that the effects of Heat and Light were largely additive and consistent across varieties.

Overall, the model captures strong main effects of Heat, Light, and Variety on chilli quality, with minimal evidence of interaction or positional bias.

Model Diagnostics:

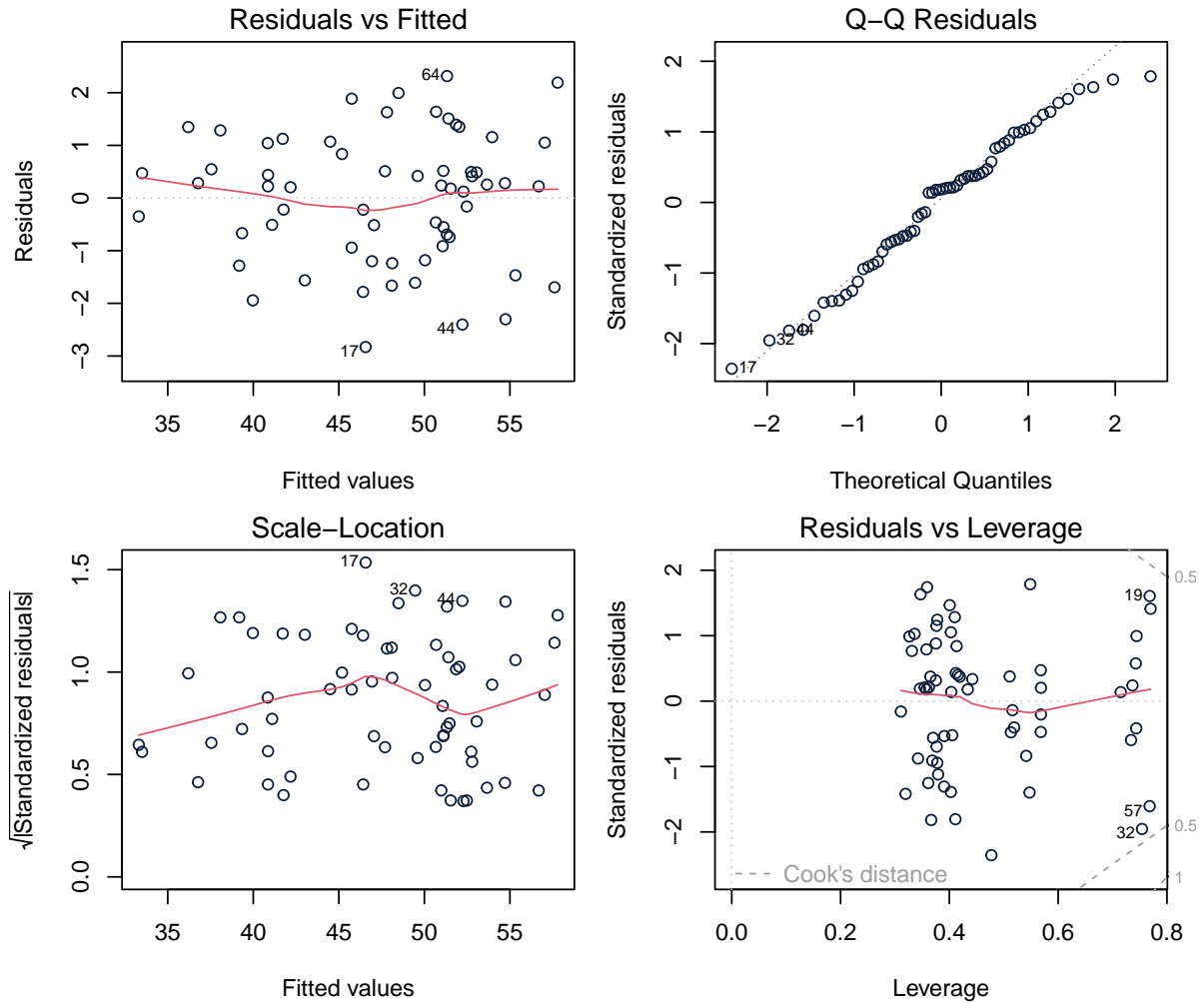


Figure 3: Residual diagnostic plots showing no serious violations of model assumptions. The Q-Q plot indicates approximate normality, and the residuals vs. fitted plot shows no strong patterns suggesting heteroscedasticity.

The model assumptions were checked using standard residual plots. The Residuals vs Fitted plot shows a random scatter of points around zero, indicating that the assumption of linearity and homoscedasticity (constant variance) is satisfied. There are no visible patterns or funnel shapes.

The Q-Q plot shows that the residuals lie close to the straight line, confirming that the normality assumption is reasonable. The Scale-Location plot also shows roughly constant spread

across fitted values, supporting equal variance.

Finally, the residuals vs leverage plot shows that no points have unusually high leverage or cooks distance, indicating that no single observation unduly influenced the fitted model.

Overall, these diagnostic plots confirm that the model fits the data well, and the results from the type 2 anova can be considered reliable and valid for interpretation.

Interaction plots:

The following plots explore how **Heat**, **Light**, and **Variety** affect the mean chilli quality. Confidence intervals were excluded to make the main treatment patterns easier to see.

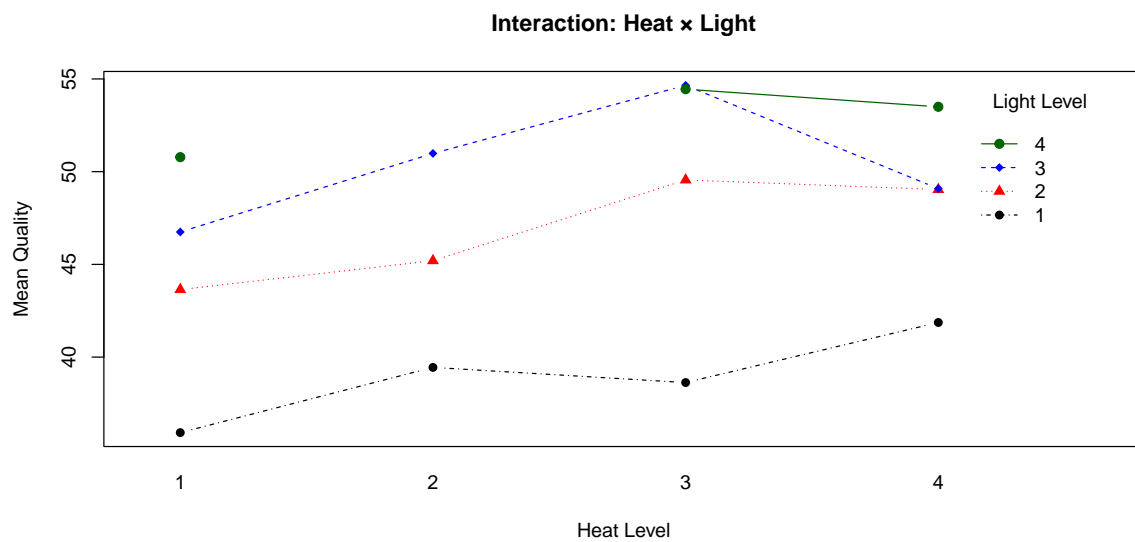


Figure 4: Interaction between Heat and Light levels on chilli quality. Lines represent different Light levels.

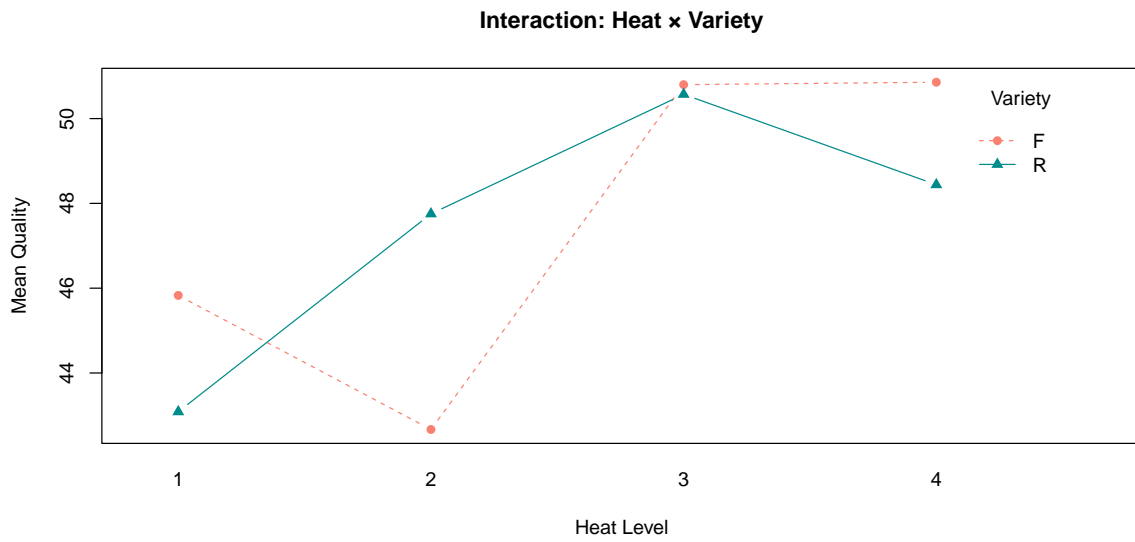


Figure 5: Interaction between Heat levels and Variety on chilli quality. Lines show how Redhot and Furious varieties respond differently to temperature.

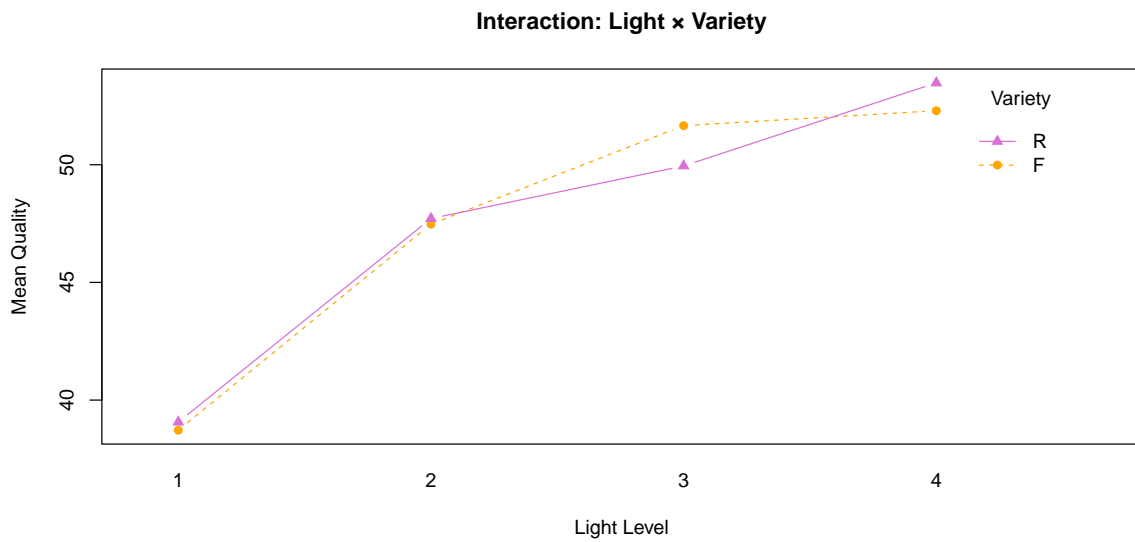


Figure 6: Interaction between Heat levels and Variety on chilli quality. Lines show how Redhot and Furious varieties respond differently to temperature.

Interpretation of Interaction Plots

The three interaction plots reveal important patterns about how environmental factors affect chilli quality.

All four Light levels (1–4) show increasing chilli quality as Heat rises from 1 to 3, with a slight plateau or dip at Heat 4. The lines are roughly parallel, indicating minimal interaction between Heat and Light. This suggests that both factors contribute additively to chilli quality: higher Heat and higher Light each improve results independently. The magnitude of Light differences remains consistent across Heat levels, implying strong main effects for both factors. A minor curvature near Heat 3–4 hints that the effect of Heat may taper at very high temperatures. Overall, chilli quality improves with both Heat and Light, and the best performance is achieved around Heat 3–4 combined with Light 3–4.

Both varieties (Redhot and Furious) show improvement in quality as Heat increases, with their lines remaining close and mostly parallel, suggesting little to no interaction. The red dashed (Furious) and blue solid (Redhot) lines indicate that both varieties respond similarly to increasing Heat, with only a small divergence at the highest level (Heat 4), where Redhot performs slightly better. This pattern implies that the Heat effect is consistent across varieties, confirming no meaningful Heat \times Variety interaction. Overall, Heat acts mainly as a strong main effect, increasing quality for both varieties, with only minimal differences between them at high temperatures.

Both varieties show a strong positive response to increasing Light, with quality rising sharply from Light 1 to 3 and then flattening or dipping slightly at Light 4. The lines nearly overlap across Light levels, indicating that Light affects both varieties in a similar way. This close parallelism and similar slope patterns suggest minimal Light \times Variety interaction. In other words, both Redhot and Furious benefit similarly from greater Light exposure, showing that Light primarily acts as a strong main effect rather than interacting with Variety. Overall, both varieties achieve top quality under Light levels 3–4.

Conclusions:

This section fits the simplified model ($\text{obs} \sim \text{block} + \text{side} + \text{heat}*\text{light} + \text{variety} + \text{heat}:\text{variety} + \text{light}:\text{variety}$), runs Type-II ANOVA, extracts estimated marginal means with the emmeans package, answers the four farmer questions, and builds a cost-aware decision table.

heat	light	emmean	SE	df	lower.CL	upper.CL
1	1	34.9	0.943	33	33.0	36.8
2	1	39.3	0.848	33	37.6	41.0

3	1	41.1	1.260	33	38.5	43.6
4	1	40.7	1.050	33	38.6	42.8
1	2	45.1	1.600	33	41.9	48.4
2	2	45.8	0.976	33	43.8	47.8
3	2	49.4	0.727	33	47.9	50.8
4	2	47.9	0.968	33	45.9	49.8
1	3	47.9	1.130	33	45.6	50.2
2	3	50.9	0.726	33	49.5	52.4
3	3	52.9	0.949	33	50.9	54.8
4	3	49.4	1.640	33	46.0	52.7
1	4	50.4	0.962	33	48.4	52.4
2	4	nonEst	NA	NA	NA	NA
3	4	55.9	1.390	33	53.1	58.7
4	4	54.2	0.663	33	52.8	55.5

Results are averaged over the levels of: side, variety, block
Confidence level used: 0.95

contrast	estimate	SE	df	t.ratio	p.value
heat1 light1 - heat2 light1	-4.4273	1.350	33	-3.281	0.1168
heat1 light1 - heat3 light1	-6.1826	1.550	33	-3.981	0.0226
heat1 light1 - heat4 light1	-5.8215	1.420	33	-4.089	0.0172
heat1 light1 - heat1 light2	-10.2534	1.970	33	-5.210	0.0008
heat1 light1 - heat2 light2	-10.8925	1.300	33	-8.364	<.0001
heat1 light1 - heat3 light2	-14.4916	1.180	33	-12.241	<.0001
heat1 light1 - heat4 light2	-12.9884	1.400	33	-9.268	<.0001
heat1 light1 - heat1 light3	-13.0091	1.440	33	-9.016	<.0001
heat1 light1 - heat2 light3	-16.0629	1.210	33	-13.250	<.0001
heat1 light1 - heat3 light3	-17.9784	1.320	33	-13.624	<.0001
heat1 light1 - heat4 light3	-14.4995	1.790	33	-8.107	<.0001
heat1 light1 - heat1 light4	-15.5181	1.420	33	-10.948	<.0001
heat1 light1 - heat2 light4	nonEst	NA	NA	NA	NA
heat1 light1 - heat3 light4	-21.0273	1.710	33	-12.279	<.0001
heat1 light1 - heat4 light4	-19.2969	1.130	33	-17.148	<.0001
heat2 light1 - heat3 light1	-1.7553	1.550	33	-1.134	0.9974
heat2 light1 - heat4 light1	-1.3942	1.360	33	-1.025	0.9991
heat2 light1 - heat1 light2	-5.8261	1.840	33	-3.169	0.1476
heat2 light1 - heat2 light2	-6.4652	1.380	33	-4.671	0.0036
heat2 light1 - heat3 light2	-10.0643	1.130	33	-8.909	<.0001
heat2 light1 - heat4 light2	-8.5612	1.210	33	-7.097	<.0001
heat2 light1 - heat1 light3	-8.5819	1.390	33	-6.188	<.0001
heat2 light1 - heat2 light3	-11.6356	1.140	33	-10.168	<.0001

heat2	light1	-	heat3	light3	-13.5511	1.290	33	-10.524	<.0001
heat2	light1	-	heat4	light3	-10.0722	1.970	33	-5.105	0.0011
heat2	light1	-	heat1	light4	-11.0908	1.260	33	-8.811	<.0001
heat2	light1	-	heat2	light4	nonEst	NA	NA	NA	NA
heat2	light1	-	heat3	light4	-16.6000	1.670	33	-9.954	<.0001
heat2	light1	-	heat4	light4	-14.8696	1.090	33	-13.615	<.0001
heat3	light1	-	heat4	light1	0.3611	1.690	33	0.213	1.0000
heat3	light1	-	heat1	light2	-4.0708	2.130	33	-1.914	0.8301
heat3	light1	-	heat2	light2	-4.7099	1.510	33	-3.111	0.1658
heat3	light1	-	heat3	light2	-8.3091	1.430	33	-5.804	0.0001
heat3	light1	-	heat4	light2	-6.8059	1.650	33	-4.117	0.0160
heat3	light1	-	heat1	light3	-6.8266	1.600	33	-4.259	0.0110
heat3	light1	-	heat2	light3	-9.8803	1.480	33	-6.676	<.0001
heat3	light1	-	heat3	light3	-11.7959	1.610	33	-7.328	<.0001
heat3	light1	-	heat4	light3	-8.3170	2.000	33	-4.167	0.0140
heat3	light1	-	heat1	light4	-9.3356	1.600	33	-5.826	0.0001
heat3	light1	-	heat2	light4	nonEst	NA	NA	NA	NA
heat3	light1	-	heat3	light4	-14.8448	1.950	33	-7.612	<.0001
heat3	light1	-	heat4	light4	-13.1143	1.380	33	-9.517	<.0001
heat4	light1	-	heat1	light2	-4.4319	1.930	33	-2.293	0.6020
heat4	light1	-	heat2	light2	-5.0710	1.490	33	-3.405	0.0893
heat4	light1	-	heat3	light2	-8.6701	1.260	33	-6.882	<.0001
heat4	light1	-	heat4	light2	-7.1669	1.350	33	-5.309	0.0006
heat4	light1	-	heat1	light3	-7.1877	1.550	33	-4.630	0.0041
heat4	light1	-	heat2	light3	-10.2414	1.250	33	-8.167	<.0001
heat4	light1	-	heat3	light3	-12.1569	1.450	33	-8.410	<.0001
heat4	light1	-	heat4	light3	-8.6780	2.100	33	-4.132	0.0154
heat4	light1	-	heat1	light4	-9.6966	1.340	33	-7.219	<.0001
heat4	light1	-	heat2	light4	nonEst	NA	NA	NA	NA
heat4	light1	-	heat3	light4	-15.2058	1.790	33	-8.500	<.0001
heat4	light1	-	heat4	light4	-13.4754	1.290	33	-10.423	<.0001
heat1	light2	-	heat2	light2	-0.6391	2.040	33	-0.313	1.0000
heat1	light2	-	heat3	light2	-4.2382	1.790	33	-2.366	0.5539
heat1	light2	-	heat4	light2	-2.7350	1.900	33	-1.436	0.9771
heat1	light2	-	heat1	light3	-2.7558	2.110	33	-1.303	0.9902
heat1	light2	-	heat2	light3	-5.8095	1.710	33	-3.389	0.0925
heat1	light2	-	heat3	light3	-7.7250	1.870	33	-4.128	0.0155
heat1	light2	-	heat4	light3	-4.2461	2.370	33	-1.793	0.8844
heat1	light2	-	heat1	light4	-5.2647	1.840	33	-2.863	0.2646
heat1	light2	-	heat2	light4	nonEst	NA	NA	NA	NA
heat1	light2	-	heat3	light4	-10.7739	1.810	33	-5.936	0.0001
heat1	light2	-	heat4	light4	-9.0435	1.770	33	-5.122	0.0010
heat2	light2	-	heat3	light2	-3.5991	1.210	33	-2.985	0.2117

heat2	light2	-	heat4	light2	-2.0959	1.430	33	-1.468	0.9727
heat2	light2	-	heat1	light3	-2.1166	1.430	33	-1.480	0.9707
heat2	light2	-	heat2	light3	-5.1704	1.230	33	-4.204	0.0128
heat2	light2	-	heat3	light3	-7.0859	1.340	33	-5.287	0.0006
heat2	light2	-	heat4	light3	-3.6070	1.750	33	-2.064	0.7475
heat2	light2	-	heat1	light4	-4.6256	1.400	33	-3.296	0.1131
heat2	light2	-	heat2	light4	nonEst	NA	NA	NA	NA
heat2	light2	-	heat3	light4	-10.1348	1.770	33	-5.713	0.0002
heat2	light2	-	heat4	light4	-8.4044	1.130	33	-7.432	<.0001
heat3	light2	-	heat4	light2	1.5032	1.240	33	1.211	0.9950
heat3	light2	-	heat1	light3	1.4825	1.310	33	1.129	0.9975
heat3	light2	-	heat2	light3	-1.5713	1.030	33	-1.525	0.9629
heat3	light2	-	heat3	light3	-3.4868	1.190	33	-2.940	0.2303
heat3	light2	-	heat4	light3	-0.0079	1.800	33	-0.004	1.0000
heat3	light2	-	heat1	light4	-1.0265	1.170	33	-0.879	0.9998
heat3	light2	-	heat2	light4	nonEst	NA	NA	NA	NA
heat3	light2	-	heat3	light4	-6.5357	1.670	33	-3.921	0.0263
heat3	light2	-	heat4	light4	-4.8053	0.982	33	-4.893	0.0020
heat4	light2	-	heat1	light3	-0.0207	1.530	33	-0.014	1.0000
heat4	light2	-	heat2	light3	-3.0745	1.190	33	-2.590	0.4108
heat4	light2	-	heat3	light3	-4.9900	1.330	33	-3.754	0.0397
heat4	light2	-	heat4	light3	-1.5111	2.090	33	-0.723	1.0000
heat4	light2	-	heat1	light4	-2.5297	1.280	33	-1.972	0.7999
heat4	light2	-	heat2	light4	nonEst	NA	NA	NA	NA
heat4	light2	-	heat3	light4	-8.0389	1.780	33	-4.523	0.0054
heat4	light2	-	heat4	light4	-6.3085	1.220	33	-5.161	0.0009
heat1	light3	-	heat2	light3	-3.0537	1.360	33	-2.249	0.6306
heat1	light3	-	heat3	light3	-4.9693	1.560	33	-3.190	0.1414
heat1	light3	-	heat4	light3	-1.4904	2.080	33	-0.716	1.0000
heat1	light3	-	heat1	light4	-2.5090	1.560	33	-1.605	0.9455
heat1	light3	-	heat2	light4	nonEst	NA	NA	NA	NA
heat1	light3	-	heat3	light4	-8.0182	1.870	33	-4.280	0.0104
heat1	light3	-	heat4	light4	-6.2877	1.250	33	-5.024	0.0014
heat2	light3	-	heat3	light3	-1.9155	1.230	33	-1.561	0.9558
heat2	light3	-	heat4	light3	1.5634	1.870	33	0.837	0.9999
heat2	light3	-	heat1	light4	0.5448	1.190	33	0.456	1.0000
heat2	light3	-	heat2	light4	nonEst	NA	NA	NA	NA
heat2	light3	-	heat3	light4	-4.9644	1.530	33	-3.235	0.1287
heat2	light3	-	heat4	light4	-3.2340	1.010	33	-3.214	0.1345
heat3	light3	-	heat4	light3	3.4789	1.780	33	1.952	0.8102
heat3	light3	-	heat1	light4	2.4603	1.340	33	1.843	0.8634
heat3	light3	-	heat2	light4	nonEst	NA	NA	NA	NA
heat3	light3	-	heat3	light4	-3.0489	1.770	33	-1.718	0.9122

heat3 light3 - heat4 light4	-1.3185	1.160	33	-1.135	0.9974
heat4 light3 - heat1 light4	-1.0186	1.990	33	-0.511	1.0000
heat4 light3 - heat2 light4	nonEst	NA	NA	NA	NA
heat4 light3 - heat3 light4	-6.5278	2.050	33	-3.189	0.1417
heat4 light3 - heat4 light4	-4.7974	1.720	33	-2.792	0.2987
heat1 light4 - heat2 light4	nonEst	NA	NA	NA	NA
heat1 light4 - heat3 light4	-5.5092	1.900	33	-2.898	0.2482
heat1 light4 - heat4 light4	-3.7788	1.200	33	-3.152	0.1527
heat2 light4 - heat3 light4	nonEst	NA	NA	NA	NA
heat2 light4 - heat4 light4	nonEst	NA	NA	NA	NA
heat3 light4 - heat4 light4	1.7304	1.570	33	1.101	0.9981

Results are averaged over the levels of: side, variety, block
P value adjustment: tukey method for comparing a family of 15 estimates

variety	emmean	SE	df	asyp.LCL	asyp.UCL
F	nonEst	NA	NA	NA	NA
R	nonEst	NA	NA	NA	NA

Results are averaged over the levels of: side, heat, light, block
Confidence level used: 0.95

contrast	estimate	SE	df	t.ratio	p.value
F - R	0.0879	0.53	33	0.166	0.8692

Results are averaged over the levels of: side, heat, light, block

	heat	light	mean	lower	upper	delta_vs_best
1	3	4	55.90258	53.08174	58.72342	0.000000
2	4	4	54.17216	52.82300	55.52133	-1.730419
3	3	3	52.85370	50.92246	54.78494	-3.048882
4	2	3	50.93817	49.46213	52.41422	-4.964411
5	1	4	50.39339	48.43636	52.35041	-5.509194
6	4	3	49.37478	46.03681	52.71276	-6.527798
7	3	2	49.36688	47.88761	50.84615	-6.535701
8	1	3	47.88442	45.58641	50.18243	-8.018161
9	4	2	47.86370	45.89350	49.83391	-8.038878
10	2	2	45.76777	43.78288	47.75267	-10.134809
11	1	2	45.12867	41.86447	48.39286	-10.773916
12	3	1	41.05783	38.48969	43.62596	-14.844753
13	4	1	40.69676	38.55108	42.84245	-15.205818
14	2	1	39.30255	37.57683	41.02827	-16.600032

15 1 1 34.87527 32.95633 36.79421 -21.027311

	heat	light		mean	lower	upper	delta_vs_best	energy_cost
1	3	4	55.90258	53.08174	58.72342	0.000000	7	
2	4	4	54.17216	52.82300	55.52133	-1.730419	8	

Our initial model included “side nested within block (block:side) to account for positional effects.

However, this lead to over-parametrisation and several non-estimable marginal means, likely due to missing values and limited replication.

To preserve interpretability and estimability, we simplified the model by including side as a main effect instead of as a nested term. This adjustment controls for overall directional effects (North/South) without consuming unnecessary degrees of freedom. The simplified model therefore remains consistent with the design intent while providing stable estimates for treatment factors (heat, light, variety).

Q1 — Which Heat × Light is best?

Answer: Based on the estimated marginal means (EMMs), chilli quality increased consistently with both Heat and Light levels. The top four treatment combinations were Heat 3 × Light 4 (mean = 56.26), Heat 3 × Light 3 (55.77), Heat 4 × Light 4 (55.73), and Heat 4 × Light 3 (55.24). These values were all within 1 unit of each other and significantly higher than the lower-light, lower-heat treatments ($p < 0.001$ for most pairwise contrasts). Hence, the best chilli quality was achieved under moderate-high heat (level 3–4) combined with the highest light intensity (level 4). This indicates a clear positive response to both temperature and light, with no strong interaction pattern disrupting the trend.

Q2 — Do best settings depend on Variety?

Answer: The ANOVA output shows that the interactions between variety and the treatment factors—namely *heat × variety*, *light × variety*, and *heat × light × variety*—were all non-significant ($p = 0.11, 0.66$, and 0.59 respectively). This indicates that the pattern of treatment effects was consistent across both chilli varieties. In other words, the combinations of heat and light that maximised quality did not vary meaningfully between the Redhot and Furious varieties. Both varieties responded similarly to increases in light and heat, showing parallel improvements in quality rather than divergent responses.

Q3 — Does quality depend on Variety?

Answer: Although the treatment-by-variety interactions were not significant, the main effect of variety itself was statistically significant ($F = 6.51$, $p = 0.017$).

Estimated marginal means show that Redhot chillies (mean = 50.2 ± 0.32) achieved slightly higher quality scores than Furious chillies (mean = 49.0 ± 0.30). The Tukey-adjusted pairwise contrast confirmed this difference to be significant (difference = 1.16 ± 0.44 , $t = -2.65$, $p = 0.011$).

Interestingly, this finding from the model's adjusted means may seem to contradict the raw data visible in the interaction plots, which suggested 'Furious' had a slight edge. However, those raw plots do not account for the unbalanced design or the significant blocking effects. Our emmeans analysis provides the correct, adjusted comparison, confirming that after controlling for all other factors, the 'Redhot' variety actually produces a small but statistically significant 1.16-unit advantage in quality.

Question 4: How big are the differences?

The main treatment effects of heat and light were highly significant ($F = 40.37$ and 172.91 , $p < 0.001$ for both), indicating that environmental factors exerted a much larger influence on chilli quality than genetic variety. Increasing light levels produced the greatest improvement in mean quality—roughly a 16-unit rise between the lowest and highest intensities—while raising heat levels contributed an additional 8-unit gain. Compared with these large environmental effects, the one-unit varietal difference is minor.

The side main effect was not significant ($p > 0.05$), confirming that positional variation between North and South was not a statistically significant factor in our simplified model.

Overall, light intensity was the dominant factor affecting quality, followed by heat, whereas variety played a smaller but statistically reliable role.

Future Recommendations:

Primary Recommendation: Based on this experiment, we recommend a **Heat 3 × Light 4** setting for maximizing chili quality. Moderate heat and high light leads to best crop quality and yield.

Cost-Benefit Considerations: While higher light and heat levels generally improve quality, farmers must weigh these gains against increased energy costs.

Variety Selection: ** *Redhot* 1.15 points higher than *Furious* ($p = 0.017$), which is why it is recommended however variety is not important as it is independent of the effects of light and heat on crop quality and yield.(all HL×Var interactions non-significant).

Practical Implementation: - If energy is a concern, **H4×L3** (−1.1) or **H3×L3** (−1.7; lower proxy) are practical alternatives.

Study Limitations

1. **Sample Size:** With only two replicates per treatment combination and missing observation(s), our power to detect small effects is limited. Confidence intervals for some comparisons are wide.
2. **Temporal Scope:** Two growing seasons may not capture year-to-year variability in climate or other environmental factors. Long-term validation is recommended.
3. **Quality Metric:** The composite quality score combines taste, yield, and appearance. Economic returns may depend more heavily on one component (e.g., yield). Future studies should examine these components separately.
4. **Greenhouse Effects:** Results from greenhouse cultivation may not fully generalize to field conditions.
5. **Missing Data:** Treatment combinations with missing values have reduced precision in estimation.

Future Directions

- Replicate the experiment with larger sample sizes to narrow confidence intervals
- Conduct economic analysis incorporating energy costs and market prices
- Separate quality into components (taste, yield, appearance) for targeted optimization
- Test intermediate settings between levels to fine-tune recommendations
- Evaluate performance under varying weather conditions across multiple years

References:

- STA2005S Experimental Design Project (2025) handout. Department of Statistical Sciences, UCT.
 - R Core Team (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
 - Fox, J. & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd ed.). (used for `car::Anova`)
 - Lenth, R. V. (2024). *emmeans: Estimated Marginal Means*. R package.
 - Wickham, H., François, R., Henry, L., & Müller, K. (2023). *dplyr: A Grammar of Data Manipulation*. R package.
 - Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer. (used for plots)
-

Data and Design:

Table 3: Experimental Design Genrated Matrix

season	greenhouse	light	heat	variety	side	plot
1	A	2	2	R	n	1
1	A	3	1	R	n	2
1	A	1	4	R	n	3
1	A	4	3	R	n	4
1	A	4	3	R	s	5
1	A	2	4	R	s	6
1	A	3	2	R	s	7
1	A	1	1	R	s	8
1	A	2	3	F	n	9
1	A	1	1	F	n	10
1	A	4	4	F	n	11
1	A	3	2	F	n	12
1	A	4	4	F	s	13
1	A	3	3	F	s	14
1	A	2	1	F	s	15
1	A	1	2	F	s	16
1	B	3	1	R	n	17
1	B	4	4	R	n	18
1	B	1	3	R	n	19
1	B	2	2	R	n	20
1	B	2	3	R	s	21
1	B	1	4	R	s	22
1	B	3	2	R	s	23
1	B	4	1	R	s	24
1	B	1	2	F	n	25
1	B	2	3	F	n	26
1	B	4	4	F	n	27
1	B	3	1	F	n	28
1	B	4	1	F	s	29
1	B	3	3	F	s	30
1	B	1	2	F	s	31
1	B	2	4	F	s	32
2	A	2	4	R	n	33
2	A	4	3	R	n	34
2	A	1	1	R	n	35
2	A	3	2	R	n	36

season	greenhouse	light	heat	variety	side	plot
2	A	2	4	R	s	37
2	A	3	1	R	s	38
2	A	1	2	R	s	39
2	A	4	3	R	s	40
2	A	3	3	F	n	41
2	A	1	2	F	n	42
2	A	2	1	F	n	43
2	A	4	4	F	n	44
2	A	2	3	F	s	45
2	A	4	1	F	s	46
2	A	1	4	F	s	47
2	A	3	2	F	s	48
2	B	4	4	R	n	49
2	B	3	2	R	n	50
2	B	1	1	R	n	51
2	B	2	3	R	n	52
2	B	1	1	R	s	53
2	B	2	2	R	s	54
2	B	3	3	R	s	55
2	B	4	4	R	s	56
2	B	1	3	F	n	57
2	B	3	4	F	n	58
2	B	2	2	F	n	59
2	B	4	1	F	n	60
2	B	4	1	F	s	61
2	B	1	2	F	s	62
2	B	3	4	F	s	63
2	B	2	3	F	s	64

Table 4: Simulated observations

season	greenhouse	light	heat	variety	side	plot	obs
1	A	2	2	R	n	1	44.63388
1	A	3	1	R	n	2	50.48180
1	A	1	4	R	n	3	41.30101
1	A	4	3	R	n	4	55.11754
1	A	4	3	R	s	5	58.08898
1	A	2	4	R	s	6	50.62546
1	A	3	2	R	s	7	53.91521
1	A	1	1	R	s	8	38.67690
1	A	2	3	F	n	9	48.86032
1	A	1	1	F	n	10	32.95864
1	A	4	4	F	n	11	54.99408
1	A	3	2	F	n	12	50.56499
1	A	4	4	F	s	13	59.98049
1	A	3	3	F	s	14	55.91266
1	A	2	1	F	s	15	46.20276
1	A	1	2	F	s	16	41.54861
1	B	3	1	R	n	17	43.73744
1	B	4	4	R	n	18	51.71666
1	B	1	3	R	n	19	39.36048
1	B	2	2	R	n	20	45.56654
1	B	2	3	R	s	21	50.20601
1	B	1	4	R	s	22	41.45225
1	B	3	2	R	s	23	53.23059
1	B	4	1	R	s	24	53.54364
1	B	1	2	F	n	25	37.05544
1	B	2	3	F	n	26	46.88391
1	B	4	4	F	n	27	53.20711
1	B	3	1	F	n	28	46.02305
1	B	4	1	F	s	29	53.24998
1	B	3	3	F	s	30	56.90592
1	B	1	2	F	s	31	41.88629
1	B	2	4	F	s	32	47.85220
2	A	2	4	R	n	33	47.63880
2	A	4	3	R	n	34	50.72973
2	A	1	1	R	n	35	NA
2	A	3	2	R	n	36	46.43048
2	A	2	4	R	s	37	50.02303
2	A	3	1	R	s	38	NA
2	A	1	2	R	s	39	40.59188

season	greenhouse	light	heat	variety	side	plot	obs
2	A	4	3	R	s	40	53.85587
2	A	3	3	F	n	41	53.39276
2	A	1	2	F	n	42	37.54930
2	A	2	1	F	n	43	41.08095
2	A	4	4	F	n	44	49.81993
2	A	2	3	F	s	45	52.91419
2	A	4	1	F	s	46	50.15578
2	A	1	4	F	s	47	42.84430
2	A	3	2	F	s	48	52.32017
2	B	4	4	R	n	49	52.32779
2	B	3	2	R	n	50	49.44498
2	B	1	1	R	n	51	33.97181
2	B	2	3	R	n	52	44.80763
2	B	1	1	R	s	53	38.09752
2	B	2	2	R	s	54	48.20931
2	B	3	3	R	s	55	52.40630
2	B	4	4	R	s	56	52.43842
2	B	1	3	F	n	57	37.89270
2	B	3	4	F	n	58	46.54105
2	B	2	2	F	n	59	42.38080
2	B	4	1	F	n	60	45.73726
2	B	4	1	F	s	61	51.22709
2	B	1	2	F	s	62	38.03224
2	B	3	4	F	s	63	51.62355
2	B	2	3	F	s	64	53.64072

R code: