Concrete Compressive Strength

Regression Analysis to discover the relationship of superplasticizer to the compressive strength of the concrete

Carolyn Dunlap, Kai Ding, Nishika Abeytunge, Scott Thompson

Sources: Sample Answer, Office Hours, Course Slack, Async Modules, Live Sessions, Materials from Mark, and/or R Bridge Course.

Contents

0.1 Introduction

The compressive strength of concrete is a key concern for engineers designing structures (National Ready Mix Concrete Association, 2014). In high rise construction, higher compressive strength concrete allows for less concrete to be used (World Cement, 2014; Tecwill, 2019). This allows for thinner columns, which increases usable space in a building (World Cement, 2014; Tecwill, 2019). Additionally, using less materials can improve the sustainability of building construction (World Cement, 2014; Tecwill, 2019) and reduce overall construction costs (Portland Cement Association, 2022; World Cement, 2014; Tecwill, 2019). For concrete manufacturers, there is a key benefit of higher compressive strength concrete - it sells for a higher price (Angi.com, 2022). Generally, higher strength concrete has a lower water-to-cement ratio (I-Cheng Yeh, 1998), and concrete manufacturers use superplasticizers to make this feasible by improving workability at low water-to-cement ratios (Portland Cement Association, 2022; World Cement, 2014; Tecwill, 2019). Concrete manufacturers may vary the proportions of superplasticizer and other ingredients during production (Portland Cement Association, 2022). Our research question is as follows:

> *Beyond the simple ratio of water to cement, is the quantity of superplasticizer (X Variable) related to the compressive strength of concrete (Y Variable)?*
>
> Adapted/taken from I-Cheng Yeh (1998). Informed by Portland Cement Association (2022).

Understanding the answer to this question could be valuable to concrete manufacturers seeking to achieve a certain level of compressive strength performance for their cement, which is vital for both maximizing revenue and enhancing concrete usage in different applications.[1]

In addition to water, cement and superplasticizers, cement manufacturers include a number of other ingredients including fly ash, silica fume and blast furnace slag (Portland Cement Association, 2022; World Cement, 2014; Tecwill, 2019). As available in our data, we may include these as additional covariates to control for their effects on the compressive strength of concrete.

0.2 Data

The data set includes 1030 observations, with each observation corresponding to an individual concrete sample. For each sample of concrete, researchers measured the compressive strength, our Y variable, in megapascals (MPa) in a laboratory. Each concrete sample also has a record of the quantity of superplasticizer, measured in kg in a m3 mixture (UCI, 2007), which is our X variable of interest in addition to the quantities (also measured in kg in a m3 mixture) of other key ingredients– cement, blast furnace slag, fly ash, water, coarse aggregate and fine aggregate– which we will consider as potential additional covariates to control for. The age of each concrete sample (in days) is also recorded. Our data set is accessed from UCI Machine Learning Repository. The original owner and source is I-Cheng Yeh of Chung-Hua University (UCI, 2007).

0.3 Operationalization

Our input variable of interest, amount of superplasticizer used, is directly operationalized by the proportion of this ingredient in each concrete mixture (measured as kg in a m3 mixture), our X variable. Our key outcome of interest, our dependent (Y) variable, is the compressive strength of concrete, which is closely operationalized as the measurement in MPa (MegaPascal, a measure of pressure).

The other covariates we included are the water:cement ratio, coarse aggregate, fine aggregate, blast furnace slag, and fly ash. We considered each of these ingredients as potential X variables of interest, since they are all manipulatable by, and of potential interest to, concrete manufacturers. We decided not to use water or cement, as the relationship between the water_cement ratio and strength was well-established in the industry (first discovered in 1918) and remains a key variable that is modulated today (Abrams, 1918; Olukoun, 1994). Unsurprisingly in our dataset, this water:cement ratio had the highest correlation with compressive strength [Figure 1a]. Of the remaining other ingredients, superplasticizer showed the highest single correlation to compressive strength ($r = 0.37$ compared to 0.17, 0.16, 0.13, and 0.11 for fine aggregate, coarse aggregate, blast slag, and fly ash, respectively) [Figure 1b]. Additionally, we were interested insuperplasticizer's understood role in improving compressive strength by making lower water:cement ratios feasible (Portland Cement Association, 2022).

The last variable available in our dataset, age of the concrete samples, we decided to control for by analyzing only samples 28 days or older, as we describe in the following Model Building section.

[1] We note that our research question and findings may not be new or novel to existing industry knowledge.

Figure 1a: Strength vs Water Cement Ratio

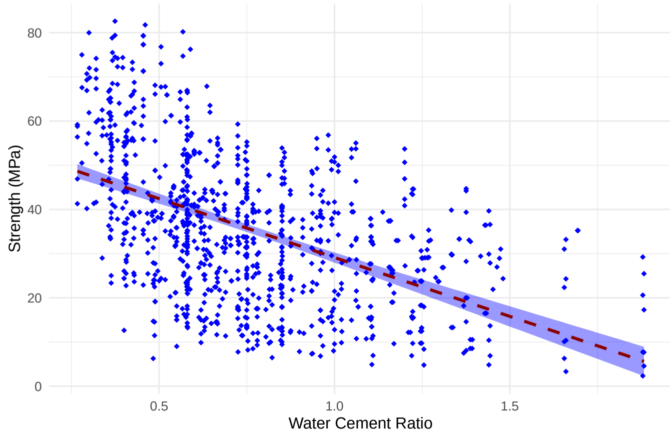
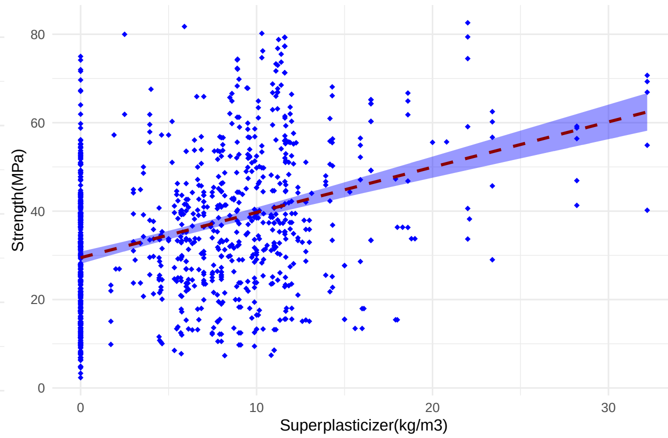


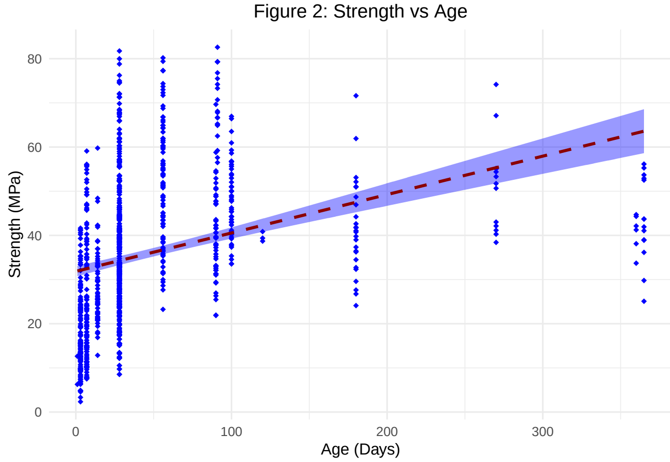
Figure 1b: Strength vs Superplasticizer

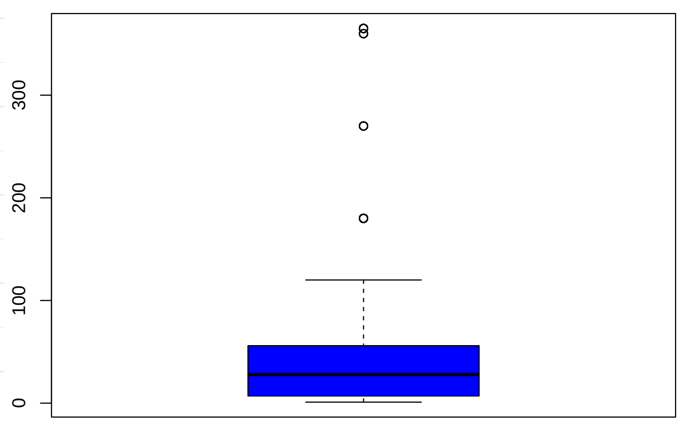


0.4 Model Building

Before building our model, we decided to remove all samples that were younger than 28 days (n=324, leaving n=706 samples for analysis). Concrete gets its compressive strength due to the chemical reaction that occurs primarily between water and cement (Herff College of Engineering, 2022). This reaction occurs over time and is well-studied, resulting in a nonlinear relationship between age and the compressive strength of concrete up to 28 days. After 28 days, the concrete has asymptotically reached its maximum compressive strength. For this reason, the industry standard is to measure the compressive strength of a concrete sample at 28 days. We observed this nonlinear relationship before 28 days in our dataset [Figure 2]. Thus we chose to only analyze samples 28 days or older, to control for these age effects and limit our model to more directly assessing the relationship between superplasticizer and compressive strength.

Figure 2: Strength vs Age





From our analysis, our resulting model was:

$$\widehat{strength} = \beta_0 + \beta_1 \cdot R \cdot (water : cement\ ratio) + \beta_2 \cdot R \cdot superplasticize + \mathbf{Z}\gamma$$

where R is an indicator for remodeling, β_1 represents the effect of water:cement ratio on concrete strength, and β_2 represents the effect of superplasticizer on strength, and $\mathbf{Z}\gamma$ is a vector of additional covariates.

We created a water:cement ratio variable from two variables water and cement because of the ratio's wide acceptance in the concrete industry. Besides the water:cement ratio and superplasticizer, we explored other model specifications, using all other variables as covariates. We analyzed models to see which would offer that would result in the highest adjusted R-squared, to both compare the effect size of superplasticizer with other concrete ingredients and identify potential omitted variable bias.

0.5 Results

Table 1 shows the results of these model explorations. Across all models, the key coefficient on superplasticizer was highly statistically significant with point estimates ranging from 0.282 to 0.546. As we can see, besides superplasticizer, slag is the most important ingredient that improves adjusted R-squared from 0.442 to 0.598.

Age and Ash strongly affect the β_1 coefficient for superplasticizer. This is not surprising, given EDA identified high correlation between Age/superplasticizer (-0.31) and Ash/superplasticizer (0.39). More specifically, not having Age in model specification (Model 2 - 3) will cause us to understate β_1 by ~40% (0.426 vs 0.601), while not having Ash in model specification (Model 3 - 4) will cause us to overstate β_1 by ~41% (0.601 vs 0.351).

To provide a sense of scale, applying Model 5, increasing superplasticizer by 1 standard deviation (5.79 units) increases concrete strength by 1.63 MPa (megapascal), which represents 0.10 standard deviation increase in strength. Applying Model 1, increasing superplasticizer by 1 standard deviation (5.79 units) increases concrete strength by 3.16 MPa (megapascal), which represents 0.20 standard deviation increase in strength. The range of superplasticizer is from 0 to 32.2 units. Therefore hypothetically, keeping all else constant, the difference between max and min superplasticizer usage results in improvement of 0.57 - 1.12 standard deviation of strength (Model 5 - Model 1).

	Dependent variable:				
	Strength				
	(1)	(2)	(3)	(4)	(5)
Water_Cement_Ratio	-30.109^{***}	-37.735^{***}	-36.374^{***}	-41.652^{***}	-43.787^{***}
	(1.688)	(1.534)	(1.479)	(1.835)	(1.973)
Plasticizer	0.546^{***}	0.426^{***}	0.601^{***}	0.351^{***}	0.282^{***}
	(0.092)	(0.079)	(0.079)	(0.094)	(0.100)
Slag		0.077^{***}	0.079^{***}	0.098^{***}	0.111^{***}
		(0.006)	(0.005)	(0.007)	(0.008)
Ash				0.047^{***}	0.058^{***}
				(0.010)	(0.011)
Coarse					0.012^{**}
					(0.006)
Fine					0.016^{***}
					(0.006)
Age			0.046^{***}	0.050^{***}	0.054^{***}
			(0.007)	(0.007)	(0.007)
Constant	60.911^{***}	61.472^{***}	56.267^{***}	57.609^{***}	33.634^{***}
	(1.576)	(1.338)	(1.483)	(1.481)	(8.404)
Observations	495	495	495	495	495
R^2	0.444	0.601	0.636	0.652	0.658
Adjusted R^2	0.442	0.598	0.633	0.648	0.653
Residual Std. Error	11.711 (df = 492)	9.939 (df = 491)	9.496 (df = 490)	9.299 (df = 489)	9.234 (df = 487)
Note:					*p<0.1; **p<0.05; ***p<0.01

0.6 Discussion and Limitations

Our biggest statistical limitation that needed to be addressed was heteroskedastic errors in the data. Because of these errors, we needed to transform compressive strength using Box-Cox transformations, which makes interpretation of our model very difficult.

Structurally, some important variables that were not included in the dataset were aspects of the concrete mixing process and curing conditions. Both of these factors heavily influence the overall compressive strength of concrete. This dataset was an accumulation of concrete samples collected across 17 different locations and so there is no way to know or control for curing or preparation conditions. Because of this, we can only assume that these factors would contribute to some of the variability in the data, which may have decreased our model's power or introduced some unknown bias.

Another important limitation of this analysis is that according to the source paper for this dataset, the superplasticizers used in each concrete sample were from multiple different manufacturers and made from different chemical compositions. Unfortunately, none of this information was kept or included in the dataset, so we have no way of knowing the different kinds of superplasticizers used, and this could very well have important implications for our model, potentially calling these results into question.

Fortunately, due to the nature of our variables and this experimental dataset, it is unlikely that there is reverse causality in the model, as compressive strength of the concrete is a physical change and would not alter the levels of superplasticizer. While we did not have any direct outcome variables on the right hand side (all ingredients are independently added to the concrete mixture), superplasticizer itself is added in order to help decrease the amount of water that can be added to the cement. Therefore, the amount of superplasticizer added to the concrete may either directly increase the compressive strength of concrete, or may allow the concrete to have a lower water:cement ratio which in turn is responsible for the increase in compressive strength.

Finally, by controlling for age, our analysis was limited to the effects of superplasticizer on the compressive strength of concrete after concrete has fully cured. Our analysis does not speak to the effects of superplasticizer on the compressive strength of concrete at earlier time points.

0.7 Conclusion

update after discussion

0.8 Sources

- UCI Machine Learning Repository. "Concrete Compressive Strength Data Set". Donated in 2007
- This Research is adapted from: I-Cheng Yeh, "Modeling of strength of high performance concrete using artificial neural networks," Cement and Concrete Research, Vol. 28, No. 12, pp. 1797-1808 (1998)
- National Ready Mix Concrete Association (2014) CIP 35 - Testing Compressive Strength of Concrete
- Angi.com "How Much Does Concrete Cost?" (2022)
- Course materials and guidance from Mark
- This research report copies from our research proposal
- World Cement. "Benefits of using high-strength concrete" (2014)
- Portland Cement Association "High-Strength Concrete" 2022
- Tecwill. "High Strength Concrete – Production and Benefits" (2019)