

A laboratory is designing devices to create “polluted air” in a chamber, in order to expose laboratory animals to a simulation of the ambient air that is typically found in a large city. To do this the experimenters use a misting machine to create a mist with suspended particles, and then draw this mist through a glass chamber at a rate that can be adjusted. The air in the chamber goes through a set of filters, and these filters are analyzed chemically to give a measure of the amount of pollution created.

There are two different chamber designs that are being considered, and there are three types of commercial misting machines that can be used. The flow rate can be varied continuously; the relevant range is thought to be between 1 and 5 liters per minute.

The chemical analysis provides a number of measurements, but the main response of interest is the concentration of particulate matter (PM) in micrograms per meter cubed. The goal is to create a concentration as close as possible to 150 micrograms per meter cubed, but no greater than 180 micrograms per meter cubed.

- (a) The statistician they consulted said that they should use a factorial design for their experiment, with the three factors being chamber design, misting machine, and flow rate. He explained that this means that they should take measurements of PM for each of the possible combinations of chamber design, misting machine type, and flow rate. Provide advice on how they might choose which flow rates to use. The statistician also said all good experiments should be randomized. Provide some advice on why this is important, and how it might be implemented in this experiment.
- (b) Suggest a model for the analysis of the experiment, using the notation y_{ijk} for the response to chamber design i , ($i = 1, 2$), misting machine type j , $j = 1, 2, 3$ and flow rate k . Include a term in your model for the *interaction* between chamber design and misting machine type.
- (c) Chamber design 1 turned out to be completely ineffective, and the goal now is to analyse the data from chamber design 2, which is given below (the numbers are the PM measurements in $\mu g/m^3$):

Misting machine type	Flow Rate				
	1	2	3	4	5
1	130.6	135.2	144.8	152.6	148.6
2	120.0	145.1	143.7	130.2	125.4
3	150.1	148.6	149.5	150.2	190.0

Without any detailed calculation, summarize the main pattern(s) in the data, and the broad conclusions.

- (d) A second series of experiments is planned. Provide some recommendations on the factors to be used, the levels of these factors, and any other relevant considerations.

Solution

- (a) Since the range of interest for flow rate is 1 to 5, it would probably make sense to start with perhaps three settings that span this range, such as 1, 3, 5. Randomization of the various settings will help to balance any confounding factors that are not being controlled, such as time of day or time of week, ambient temperature, and so on. There might also be confounding factors associated with analysis of the filters; perhaps it is done by different operators, working at different lab stations, and so on. In this experiment it is unlikely that the choice of misting machine can be changed easily, so it is likely that each misting machine will be used for a set of glass chambers and all possible flow rates. So one might, for example, choose misting machine 1 to create a volume of mist, which is drawn through a glass chamber of type 1 or 2 (type to be chosen at random), which itself is randomly selected from the set of glass chambers available. Filter blanks would be randomly sampled from the box of filters, and flow rate might also be set randomly at one of the three levels 1, 3 or 5. The flow rate could then be changed to some other value; perhaps some 'washout' period would be needed. At the end of one 'run' one might hope to have samples from each of the three flow rates and each of the 2 chambers. Specific details would require considerable discussion with the experimenters.
- (b) A suggested model is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + \epsilon_{ijk},$$

where α_i is the effect on response of design i , β_j the effect of machine j , and γ_k the effect of flow rate k . Note that this notation does not incorporate the nesting, or split-plot, design described above, and in particular if it really is possible to use only one misting machine at a time, the interaction between misting machine and chamber doesn't make sense. However perhaps it is possible to have the three misting machines in separate rooms running at the same time. If there are replicate runs for each combination of factors then we could add a subscript ℓ to indicate this, and introduce a factor for replicates, δ_ℓ , say.

- (c) The row and column means are: 133.57 142.97 146.00 144.33 154.67 and 142.36 132.88 157.68, showing that the response seems to increase with flow rate but decrease, then increase, with machine type. Machine 3 has responses close to the desired concentration of 150, but something weird happened with flow rate 5. Machine 1 has responses close to desired at the higher flow rates. This suggests a strong interaction between machine type and flow rate.
- (d) Flow rates between 3 and 5 seem to warrant further investigation, and the outlier 190.0 should be investigated, because otherwise machine type 3 seems to be quite effective. Machine type 2 seems quite unsuitable so further experimentation should concentrate on Machine types 1 and 3. Some information on whether or not other chamber designs might be worth considering should be sought. Replication of the runs would seem important at this stage.