

From:
Sent: 3/7/2021 11:03:58 PM
Subject:
Attachments: Amended_Copy of WildfireMitigationPlans_DR_CalAdvocates_035-Q04-Atch01 (1).xlsx;
WildfireMitigationPlans_DR_CalAdvocates_049-Q01.docx



Attached are the draft response for DR_CalAdvocates_049-Q01 (Word document) along with the accompanying file required to respond Part a. (Excel File)

I created two columns, one for Equipment and one for Veg.

As a reminder, here is a summary of our discussion on Friday and the approach I took:

- Use mean as way of aggregation
- Mark with N/A circuits that do NOT have a score but are in a previously labeled HFTD
- Bring “new” circuits from 2021 models. The table below is a summary of circuits that were not in the 2018 model, but now they have a score on either model and in one instance (COAST RD 0401) a score from each model.

Circuit Name	2021 Wildfire Risk Level Equipment	2021 Wildfire Risk Level EVM
CASTRO VALLEY 1106	N/A	0.020389316
SPENCE 1122	N/A	0.002984031
COAST RD 0401	0.573212458	4.00246E-05
CARMEL 0402	N/A	8.33808E-06
CASTRO VALLEY 1101	N/A	8.13212E-06
CARMEL 0405	N/A	2.26368E-06
BIG TREES 0402	0.003149775	N/A

Lastly, here is the latest I got from the discussion with the data science team about aggregation and their preference:

“The question really is the use case. If the end user wants to just now how much risk each circuit carries, yes the risk sum is the thing to produce. If they are interested in the **risk density** in each circuit, the aggregation should return the mean. I guess the question is whether they have strong preferences, and if not, how are they planning to use the results?

Still though, without any preference from the end user, we should default to giving the sum for risk. It is the conceptually simplest thing to work with, and it certainly is the case that all else being equal, a circuit that covers more line miles will have greater risk, just because it is longer”