Review of: "A Multi-factor Model of COVID-19 Epidemic in California"

Wei Sun[1]

1 Beijing University of Chemical Technology

Potential competing interests: No potential competing interests to declare.

In this paper, a statistic model is established to estimate the cumulative cases and duration of the COVID-19, and multiple area characteristics are considered in this model. However, the method details and application value of the proposed model are not clear. My major concerns are listed as follows:

1. It seems inappropriate to choose these five factors to estimate the cumulative cases and duration of the epidemic, because the linear dependences of the population density, family income, Gini coefficient, and land area with the cumulative cases are not significant as shown in Figure 2.

2. As the author said “Many models have been developed for estimating and predicting the spread of contagions via contact”, what is the advantages of the proposed multi-factor model compared with other statistic models or machine learning models?

3. The method should be addressed with more details. As the author said “The Fama-French equation uses a similar approach to estimate …”, what is the specific similarities between the economic method and the proposed epidemic estimation model? In addition, how to obtain the values of correlation coefficients shown in Table 2, it seems different with Figure 2. And the definition of the duration is also needed to be cleared.

4. The calculation equation of I(x) and D(x) are different with the definition of P(x), the coefficient in front of x5 and the denominator of D(x) need more explanations. According to the calculation equation of P(x), the value of I(x) is supposed to be less than 1, but the x-coordinate in Figure 3 is greater than 1.