Covid-19 Model for the United States

David I Gustafson, Ph.D.
St. Louis, MO

Last Daily Update: April 25, 2020

Worldometer (https://www.worldometers.info/coronavirus/country/us/) reported 1951 covid-19 deaths in the US yesterday, dropping to a weekly average of 2106, prompting a new calibration of the **CDE-k** model released on April 19. This new model had been prompted by two major developments: (1) the Wuhan data upon which the original March 23 model was based have since been disclosed to have been erroneous, with actual mortality about 50% higher than originally reported; and (2) as deaths decline in Italy, they are strongly diverging away from the symmetrical model, which was based on the faulty Wuhan data. Still preferring a simple analytical solution to this modeling challenge, I borrowed from a paper I wrote back in my twenties,[1] when I could still solve partial differential equations! Although the topic of that paper was using novel solutions to the Convective Dispersion Equation (CDE) to better describe the movement of chemicals through soil, it turns out that completely analogous physical processes are apparently involved in the "diffusion" of covid-19 through human communities. As shown in **Figure 1**, such an equation gives a compelling fit to observed mortality data for Italy. Here is the equation (simplified from Equation 11 in the 1988 paper):

$$D_i = \frac{\delta \exp[-\{(t_m/t_i) - 1\}^2/2k]}{(t_i/t_m)}$$

> where D_i is the number of deaths on day t_i (the number of days since first death); k is a fitted constant (0.25) proportional to the rate at which dispersion increases; t_m is a fitted constant (37 days) proportional to the time from first death to peak deaths: $t_m = 2\,k\,t_{peak}$/(sqrt[1+4k]-1); and δ is a fitted constant proportional to mortality.

The parameter k was initially fit to the observed Italian data using a transformation method given in the 1988 paper (see **Figure 2**). The fact that the data are linear when transformed in this manner (lower right of **Figure 2**) is strong evidence that this equation is giving a good fit to the data. I found that the data far before and far after the peak had to be removed to preserve linearity, and so when reconstructing the overall curve shown in **Figure 1**, the model was built by assuming Italy was infected by four waves offset by 5 days each. Based on the still slowing decline in Italy revealed in the weekly averages, the model was re-parameterized on April 25, resulting in a higher value for k: 0.25.

This re-parameterized model was applied to all individual cities in the US with one other major change: the imputed first death was assumed to have taken place 7 days before the reported date (see **Figure 3**). Reported dates of first death were used for the top 100 metro areas. The remaining one-third of the country was modeled using 1000 progressively smaller simulated cities, with their days to first death modeled stochastically using a regression among the largest cities, which showed a significant inverse trend between population size and date of first death.

Both the **CDE-k** model (dark blue dots) and the original Wuhan model proposed on March 23 (light blue dots) are shown in **Figure 3**. The **CDE-k** model predicts a US death toll of 98,400 through May 31, and a 6-month death toll of 133,400. IHME (https://covid19.healthdata.org/projections) currently projects a death toll of only 67,000 and much more rapid decline that the **CDE-k** model. As a resident of St. Louis, I've added the specific result for the St. Louis metro region (**Figure 4**), which suggests the peak in deaths has passed and that a slow decline is underway. Given the observation that the peak has now passed at both the US scale and within the St. Louis region, the frequency of these reports will be reduced.

---

[1] DI Gustafson (1988). Modeling root zone dispersion, *Chem. Eng. Comm.*, **73**:77-94.

© 2020, Real Whirlwind LLC (314) 409-7123 dr.dave@real-whirlwind.com

Disclaimer: This is only a model. As the saying goes: "All models are wrong - some are useful." So we know this model is wrong, but we don't know if it is useful. Nevertheless, perhaps it helps calibrate expectations for what is likely to come here in the US in the coming weeks.

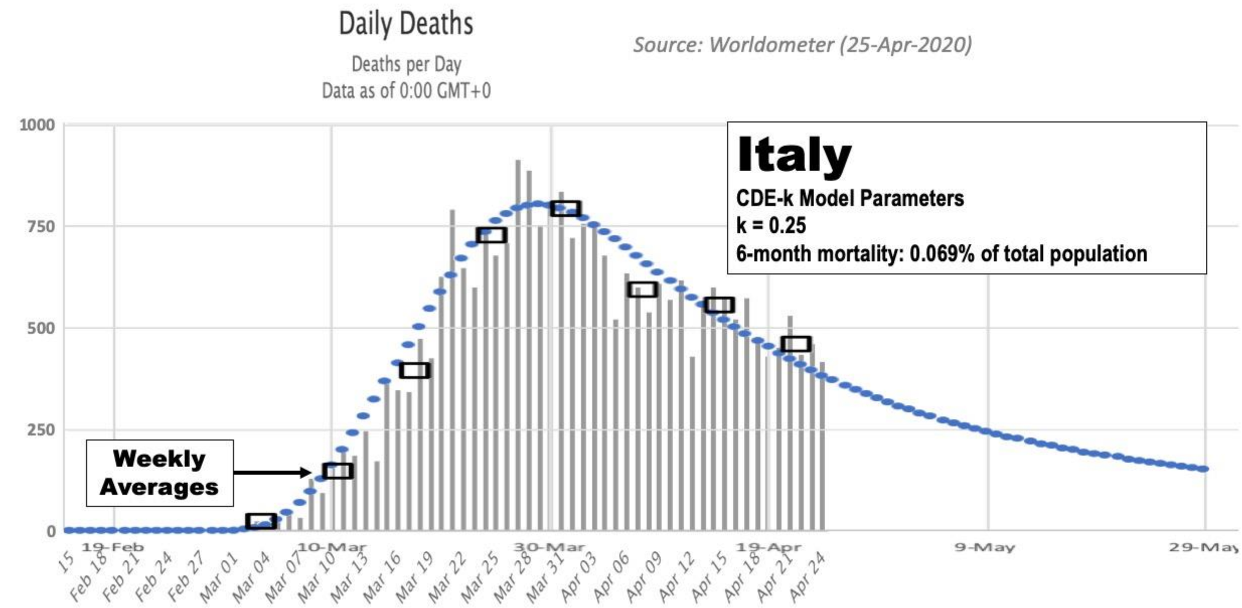
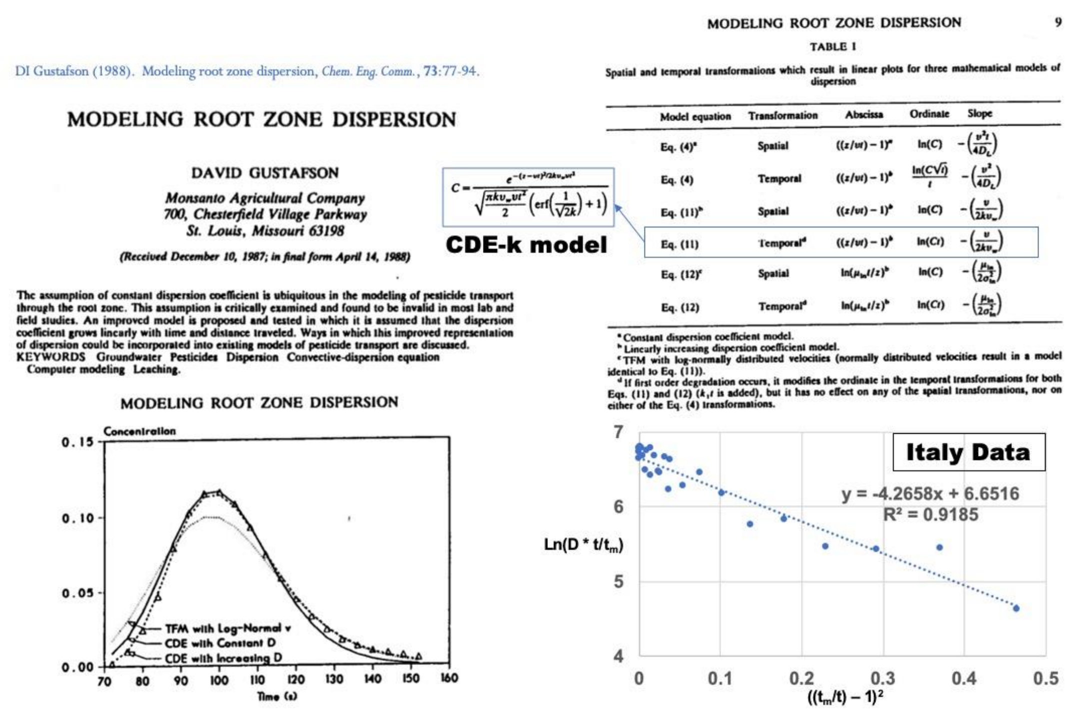


Figure 1. Application of the **CDE-k** model to observed mortality data in Italy.



DI Gustafson (1988). Modeling root zone dispersion, *Chem. Eng. Comm.*, **73**:77-94.

MODELING ROOT ZONE DISPERSION

DAVID GUSTAFSON

Monsanto Agricultural Company
700, Chesterfield Village Parkway
St. Louis, Missouri 63198

(Received December 10, 1987; in final form April 14, 1988)

The assumption of constant dispersion coefficient is ubiquitous in the modeling of pesticide transport through the root zone. This assumption is critically examined and found to be invalid in most lab and field studies. An improved model is proposed and tested in which it is assumed that the dispersion coefficient grows linearly with time and distance traveled. Ways in which this improved representation of dispersion could be incorporated into existing models of pesticide transport are discussed.
KEYWORDS Groundwater Pesticides Dispersion Convective-dispersion equation Computer modeling Leaching.

$$C = \frac{e^{-(z-vt)^2/2kv_m vt^2}}{\sqrt{\frac{\pi k v_m v t^2}{2}}\left(\mathrm{erf}\left(\frac{1}{\sqrt{2k}}\right)+1\right)}$$

CDE-k model

MODELING ROOT ZONE DISPERSION 9

TABLE I

Spatial and temporal transformations which result in linear plots for three mathematical models of dispersion

Model equation	Transformation	Abscissa	Ordinate	Slope
Eq. (4)[a]	Spatial	$((z/vt)-1)^a$	$\ln(C)$	$-\left(\frac{v^2t}{4D_L}\right)$
Eq. (4)	Temporal	$((z/vt)-1)^b$	$\frac{\ln(C\sqrt{t})}{t}$	$-\left(\frac{v^2}{4D_L}\right)$
Eq. (11)[b]	Spatial	$((z/vt)-1)^b$	$\ln(C)$	$-\left(\frac{v}{2kv_m}\right)$
Eq. (11)	Temporal[d]	$((z/vt)-1)^b$	$\ln(Ct)$	$-\left(\frac{v}{2kv_m}\right)$
Eq. (12)[c]	Spatial	$\ln(\mu_{ln}t/z)^b$	$\ln(C)$	$-\left(\frac{\mu_{ln}}{2\sigma_{ln}^2}\right)$
Eq. (12)	Temporal[d]	$\ln(\mu_{ln}t/z)^b$	$\ln(Ct)$	$-\left(\frac{\mu_{ln}}{2\sigma_{ln}^2}\right)$

[a] Constant dispersion coefficient model.
[b] Linearly increasing dispersion coefficient model.
[c] TFM with log-normally distributed velocities (normally distributed velocities result in a model identical to Eq. (11)).
[d] If first order degradation occurs, it modifies the ordinate in the temporal transformations for both Eqs. (11) and (12) (k_1t is added), but it has no effect on any of the spatial transformations, nor on either of the Eq. (4) transformations.

MODELING ROOT ZONE DISPERSION

Figure 2. Source document (Gustafson 1988) for the **CDE-k** model.

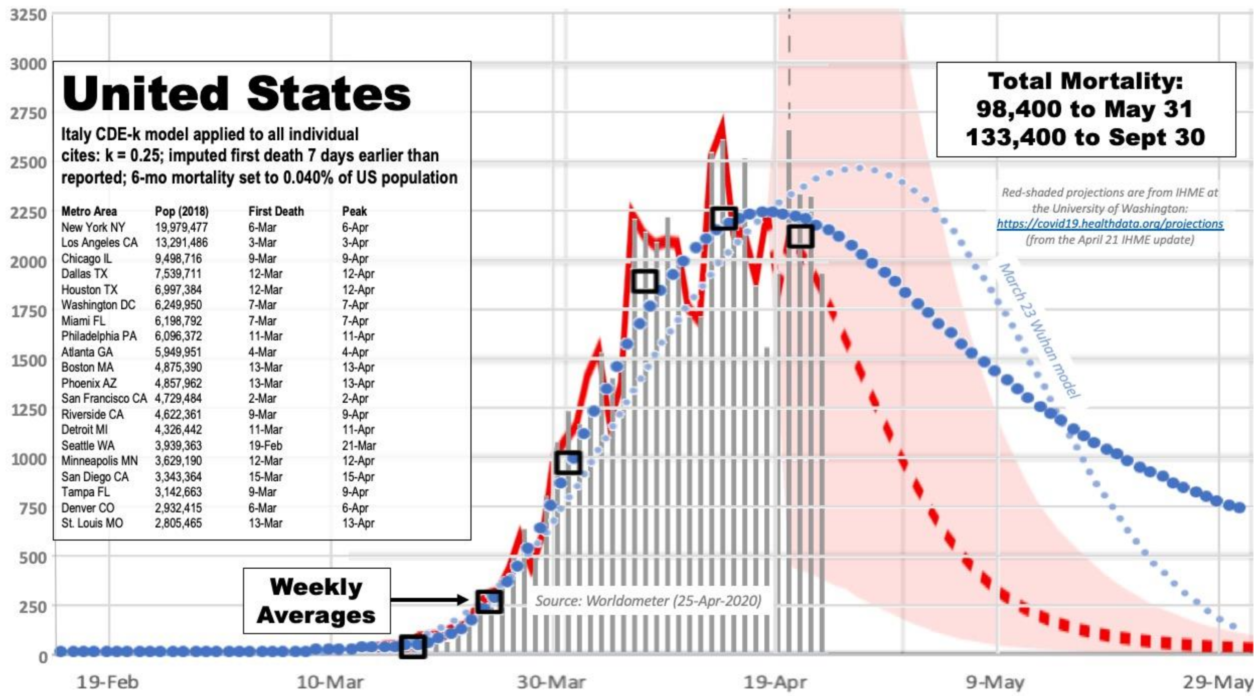


Figure 3. Application of the March 23 Wuhan model (light blue) and the **CDE-k** model (dark blue) to the US, overlaid Worldometer data (gray bars, black boxes), as well as April 21 data (solid red line) and projections (red dashed line and shading) from IHME.

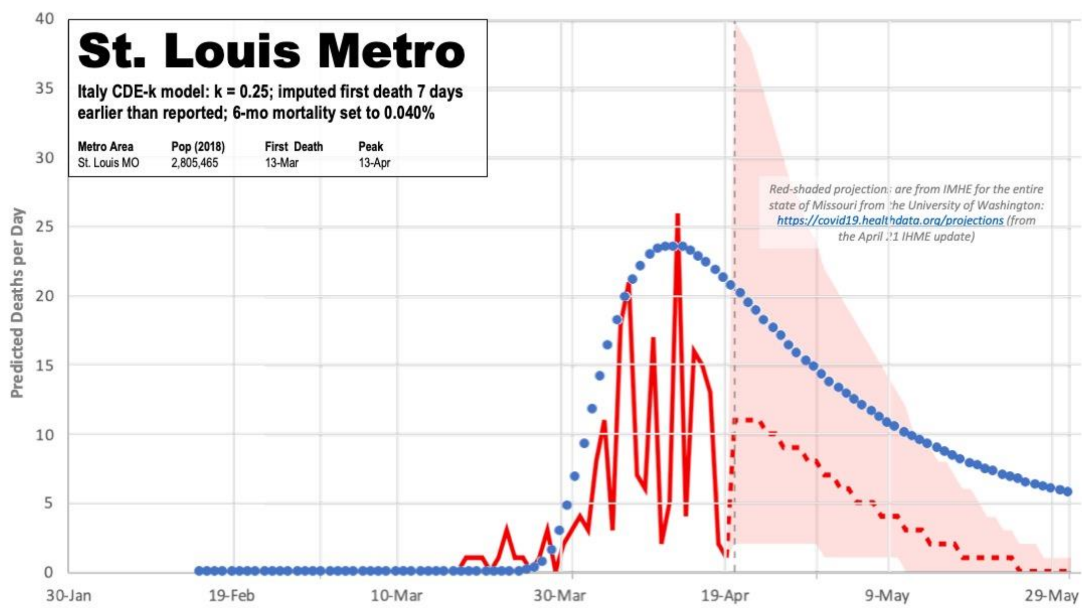


Figure 4. Application of the **CDE-k** model to St. Louis, with April 21 Missouri data (solid red line) and projections (red dashed line and shading) from IHME.