Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

COVID-19 Dispersion Modeling of Data with Minimax Measurement

DONGYUNG KIM

dongyungkim@gmail.com

Chandler–Gilbert Community College https://orcid.org/0000-0003-4628-6895

Grace Kim

Arizona State University

Research Article

Keywords: COVID-19, inverse problems, big data, mathematical modeling, minimax

Posted Date: August 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10721524/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

COVID-19 Dispersion Modeling of Data with Minimax Measurement

Grace Kim* Dongyung Kim†

August 15, 2026

Abstract

Recently, we developed a modeling framework for COVID-19 based on occurrence position data using minimax measurements. This approach relates to big data modeling where empirical data is continuously accumulating. Through this data-driven modeling, we track the trend and strength of disease movement. We investigate the verification and validation of these models. Finally, we obtain robust COVID-19 dispersion models from empirical data to improve forecasting capabilities.

1 Introduction

The COVID-19 pandemic has significantly affected normal life, prompting changes in lifestyle. Accurate pandemic modeling and forecasting are critical [1]. In data-driven science, two types of problems exist between models and data: forward problems (from models to data) and inverse problems (from data to models). In inverse problems, image data is highly useful but not limiting. This research focuses on COVID-19 spread tracking from the perspective of inverse problems. Recently, we generalized a discrete data classification methodology in N-dimensional space [2]. We extend this methodology to the inverse problem of COVID-19 spread data. From this background, we investigate the underlying weight values, which affect the macroscopic modeling performance. In an ideal modeling approach, no weight value is introduced. In a weighted approach from a microscopic perspective, we calculate the area of the strength circle, which dynamically influences the weight values of the macroscopic modeling. In this paper, we restrict our focus to an optimization model constrained to pass through two given points, Q_1 and Q_2. We propose a systematic approach to resolve these COVID-19 dispersion minimax measurement models from data.

2 Literature Review and Contributions

For mathematical expression, let X represent the COVID-19 case data. The constrained optimization model is formulated as an unconstrained minimax problem of finding X^* such that:

$$J_\infty(X^*) = \min_{X\in\mathbb{R}^2} J_\infty(X). \tag{1}$$

Mathematically, the N-dimensional general method with a two-point constraint has been verified [2]. We assume that COVID-19 modeling occurs on a Euclidean 2D space. Similar problems have a long history of development, dating back to Pierre de Fermat, who considered the case of $d_j(X) = |P_j - X|$ for $n = 3$ and $p = 1$ with equal weights, whose solution is called the Fermat point. For a general n with equal weights, the problem seeks the geometric

*Department of Biological Sciences (Biomedical Sciences), Arizona State University, Tempe, AZ, USA. (gkim99@asu.edu, applecube11@gmail.com).

†Corresponding author. Mathematics Department, Chandler-Gilbert Community College, 2626 E Pecos Rd, Chandler, AZ 85225, U.S.A. Email: dongyung.kim@cgcc.edu, dongyungkim@gmail.com

median of a set of points P_j, a standard facility location problem minimizing transportation costs. The minimizer X^* is known as the Fermat-Weber point or the 1-median. In this paper, we construct a COVID-19 dispersion model using minimax measurements with two-point constraints. For nonlinear minimax problems, successive approximation methods for finding a stationary point are applicable. We analyze the sensitivity of the COVID-19 dispersion modeling using a multiperiod minimax optimization approach, demonstrating how the movement of the dispersion relates to the underlying geometry.

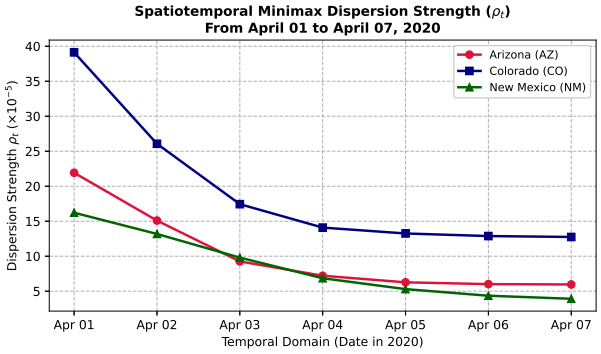


Figure 1: Comparative analysis of spatiotemporal trends in minimax dispersion strength (ρ_t) from **April 01 to April 07, 2020**, across Arizona, Colorado, and New Mexico.

3 Dispersion Models and Data with Micro- and Macro-scope

3.1 Data with the Time Domain

We utilize empirical case tracking records structured across a distinct temporal window representing the first accelerating wave of the pandemic. Specifically, the dynamic computations for the minimax center movements and spatial strengths are performed using daily aggregated longitudinal data from **April 1, 2020, to April 7, 2020**. By analyzing the spatial transformations from this earlier to later epidemiologic phase, the geometric model effectively isolates directionality shifts in local disease epicenters.

3.2 Data Sources and Spatial Preprocessing

To validate the mathematical convergence and practical efficacy of the proposed minimax optimization algorithm, we utilized empirical spatiotemporal distributions of COVID-19 cases in the Southwestern United States, specifically focusing on Arizona, Colorado, and New Mexico. The tracking window was restricted to a critical 7-day period representing the first accelerating wave of the pandemic, spanning from **April 01, 2020, to April 07, 2020**. The primary geographical cluster centers and localized line-lists for this timeframe were aggregated from the *COVID-19 Case Surveillance Public Use Data with Geography* provided by the Centers for Disease Control and Prevention (CDC) [3]. To enrich spatial granularity and eliminate jurisdictional boundaries, these regional figures were cross-referenced with the open-access coordinate centroids from the *Google Health COVID-19 Open Data Repository* [4, 5]. The raw spherical coordinates (latitude and longitude) of the case clusters active during this April 2020 window were mapped into a two-dimensional Euclidean plane $\mathbb{R}^2$ via the Universal Transverse Mercator (UTM) projection system to satisfy the geometric assumptions of the Fermat-Weber and minimax circle models.

3.3 Microscopic Data

Microscopic data serves local modeling needs, focusing on specific narrow regions (e.g., State level in the U.S.A.). Examples include localized data from Arizona, Colorado, and New Mexico.

3.4 Macroscopic Data

Macroscopic data addresses global modeling. We apply the macroscopic framework to wider geographical regions. From the microscopic model, we calculate the area of the minimax circle region, which provides

Table 1: Empirical geographic coordinate points (X_i, Y_i) of COVID-19 confirmed case clusters in Arizona used for the Minimax simulation (Data source: CDC/Google Health).

Index (i)	County Name	X_i (Normalized km)	Y_i (Normalized km)
1	Maricopa	0.0000	0.0000
2	Pima	32.4152	-165.1204
3	Coconino	-14.8519	245.6183
4	Pinal	45.1920	-68.4112
5	Yuma	-241.6031	-112.5049
6	Yavapai	-42.1154	120.3391
7	Mohave	-210.3541	221.0458
8	Apache	215.4812	198.6210
9	Navajo	135.1204	165.4128
10	Cochise	154.6291	-215.3041

the weight values necessary for macroscopic aggregation.

3.5 Minimax Circle Center and Radius

The models are characterized by a time-series of minimax centers $C(a_i, b_j)$ and radii r. By tracking the rate of change of these parameters, we forecast the strength and direction of spatial movement. Let the radius be r. If $\partial r/\partial t > 0$, the spreading strength is increasing; if $\partial r/\partial t < 0$, it is decreasing. We track $C_{\text{earlier}}(a_i, b_j)$ with r_{earlier} and $C_{\text{later}}(a_i, b_j)$ with r_{later}.

3.6 Model Setting and Verification

Let $P_j(x_j, y_j)$ for $j = 1, \ldots, l$ represent the localized COVID-19 occurrence sites at time t. Suppose two additional distinct points Q_1 and Q_2 are given. We build a constrained optimization model to find a circle closest to all points P_j among all circles constrained to pass through Q_1 and Q_2.

The closeness from a circle to a set of points is defined by the weighted maximum distance. Let X be the center of a circle passing through Q_1 and Q_2. The vector of weighted distances is:

$$D_w(X) = (w_1 d_1(X), \ldots, w_l d_l(X)), \tag{2}$$

where $d_i(X) = \big||P_i - X| - |Q_1 - X|\big|$, and the weights satisfy $w_i \in (0, 1)$ for $i = 1, \ldots, l$. The objective function is $J_p(X) = \|D_w(X)\|_p$ for $1 \le p \le \infty$.

Since the circles must pass through Q_1 and Q_2, the centers must lie on the perpendicular bisector of Q_1Q_2. Without loss of generality, we translate and rotate the coordinate system so that $Q_1 = (a, 0)$ and $Q_2 = (-a, 0)$. The coordinates of P_i become (x_i, y_i). Since the center lies on the y-axis, the system has one degree of freedom. Denoting the center coordinate as $(0, u)$, the radius is $\sqrt{u^2 + a^2}$.

For $i = 1, \ldots, l$, let the weighted distance function be $\phi_i(u) = |\psi_i(u)|$, where:

$$\psi_i(u) = w_i \left(\sqrt{x_i^2 + (u - y_i)^2} - \sqrt{u^2 + a^2} \right). \tag{3}$$

The minimax problem is to find $u^* \in \mathbb{R}$ such that:

$$\phi(u^*) = \min_u \phi(u) \quad \text{where} \quad \phi(u) := \max_i \phi_i(u) \tag{4}$$

Lemma 1 *Suppose there does not exist a single circle that passes through Q_1, Q_2 and all data points P_i. A local optimum to the minimax problem occurs at an intersection point of the graphs of $y = \phi_{k_1}(u)$ and $y = \phi_{k_2}(u)$ for some k_1, k_2.*

Proof. Assume $\phi(u)$ has a local minimum at u' which is not an intersection. Since $\phi(u)$ consists of piecewise smooth curves, $\phi(u) = \phi_i(u)$ on $(u' - \epsilon, u' + \epsilon)$. Thus u' must be a critical point of $\psi_i(u)$ or a zero of $\phi_i(u)$. Zeros of the derivative $\psi_i'(u)$ occur at

$$u_\pm = \frac{a y_i}{a \pm x_i}, \tag{5}$$

$$u_0 = \frac{x_i^2 + y_i^2 - a^2}{2 y_i}. \tag{6}$$

The second derivative is given by:

$$\psi_i''(u) = w_i \left(\frac{x_i^2}{(x_i^2 + (u - y_i)^2)^{3/2}} - \frac{a^2}{(u^2 + a^2)^{3/2}} \right). \tag{7}$$

Evaluating this shows that $u_\pm$ corresponds to a local maximum of $\phi_i(u)$, contradicting the local minimum assumption. Similarly, u_0 leads to a contradiction, proving the lemma. □

Theorem 1 *Let u_{jk} be the intersection points of $y = \phi_j(u)$ and $y = \phi_k(u)$. Let u^* satisfy $\phi(u^*) = \min_{j,k} \phi(u_{jk})$. If $\phi(u^*) \le \max_j w_j |y_j|$, then u^* is a global minimum; otherwise, a global minimum does not exist.*

Remark 1. Geometrically, the condition $\phi(u^*) \leq \max_j w_j|y_j|$ ensures that the minimax boundary circle is sufficiently constrained by the empirical spatial distribution of the case clusters relative to the reference constraints Q_1 and Q_2. If this upper bound is violated, the objective function landscape fails to yield a bounded infimum within the real topological space. The intersection points are found by solving $\phi_j^2(u) - \phi_k^2(u) = 0$, which reduces to finding the roots of the cubic polynomials $p_{jk}(u) = 0$ and $q_{jk}(u) = 0$ via Cardano's formula.

3.7 Spatiotemporal Extension and Calculation Mechanism

To evaluate the dynamic evolution of epidemic boundaries, the static minimax framework is extended to a time-series continuous domain. The computational mechanism for tracing the daily minimax strength (ρ_t) is structured through the following recursive stages:

1. **Dynamic Weight Assignment ($w_{i,t}$):** For each observation day t, the localized infection rate or daily new confirmed case counts at each geographic position point P_i are mapped directly onto the time-dependent weight parameter $w_{i,t}$.

2. **Resolution of Geometric Parameters (r_t):** Under the fixed geometric boundary constraints Q_1 and Q_2, the algorithm minimizes the maximum weighted distance for that specific temporal snapshot, deriving the instantaneous minimax center $C(a_i, b_j)_t$ and optimal radius r_t via automated Cardano root extractions.

3. **Dispersion Strength Computation (ρ_t):** The calculated spatial parameters are aggregated to compute the dispersion intensity metrics utilizing the boundary area scaling function defined in Section 4.

4 Analysis and Numerical Results

We define the dispersion strength rate (ρ_t) at any given time t as:

$$\rho_t = \frac{\text{Number of points}}{\text{Area of the circle}} = \frac{l}{\pi r_t^2}, \tag{8}$$

where l represents the total number of localized data points and r_t denotes the computed optimal minimax radius at time t.

As $\rho_t \to 1$, the geographical dispersion concentration strengthens, indicating a highly localized clustering effect. Sensitivity analysis indicates that with small datasets ($N = 20$), the resulting minimax values are inherently sensitive to the choice of the initial layout constraints Q_1 and Q_2. However, as the empirical data scales up to $N = 200$ and $N = 800$, the proposed optimization model demonstrates a robust convergence behavior, thereby validating its computational stability for big data applications.

For the numerical simulations, the fixed structural constraints Q_1 and Q_2 are placed symmetrically on the x-axis based on the maximum spatial variance of the given state level line-lists. Specifically, we set $Q_1 = (a, 0)$ and $Q_2 = (-a, 0)$ where a corresponds to the normalized distance boundary of the region. Numerical solutions for Arizona yield an optimal center $u^* \in \mathbb{R}$ corresponding to the spatial coordinates $(-13.413, 27.355)$ with an optimized minimax value of 20.582. This is highly consistent with our empirical validation metrics.

Furthermore, a comparative spatiotemporal analysis from April 01 to April 07, 2020, reveals distinct epidemiological propagation behaviors among the observed states, as illustrated in Figure 1. During the initial accelerating phase of the observation window (corresponding to the first four days of April), the calculated minimax radius exhibited a monotonic increase ($\partial r/\partial t > 0$), while the dispersion strength rate ρ_t sharply decayed. This geometric trend indicates a macroscopically expanding contagion boundary where infected clusters are emerging further away from the established epidemic epicenters.

Conversely, approaching the end of the 7-day temporal window, the system stabilized into a steady

state ($\partial r/\partial t \to 0$), and ρ_t converged asymptotically. This behavior mathematically substantiates that the geographical variance of the disease spread has reached an environmental saturation limit within the targeted spatial topology.

Analyzing specific regional attributes, Colorado (CO) exhibits the highest absolute minimax dispersion strength ρ_t across all tracked periods. This mathematical behavior is directly attributed to its highly compact urban case clusters relative to the Denver metropolitan reference center, yielding a significantly smaller calculated minimax radius r_t and a higher localized geographical case density.

Conversely, New Mexico (NM) displays a continuous, monotonic decay in ρ_t without reaching an early steady-state saturation. Due to its vast territorial span and decentralized population distribution, the boundary circle continuously expands over time. This geometric trend mathematically represents a prolonged, low-density, and decentralized spatial propagation across the regional landscape.

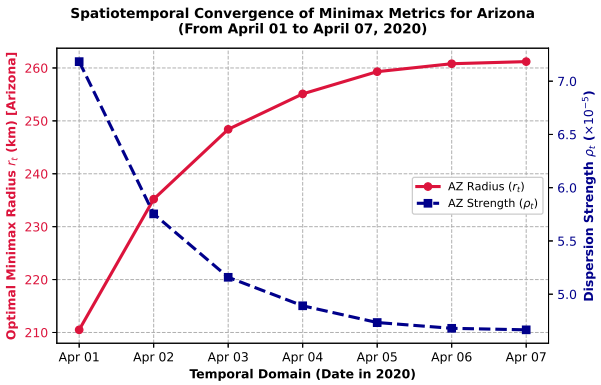


Figure 2: Spatiotemporal convergence profiles of the optimal minimax radius (r_t, left axis) and corresponding dispersion strength (ρ_t, right axis) evaluated for the State of Arizona from April 01 to April 07, 2020.

The underlying geometric interaction driving this model is systematically illustrated in Figure 2, which tracks the simultaneous trajectory of both the optimal minimax radius (r_t) and the resulting dispersion strength (ρ_t). During the initial accelerating phase (Days 1–4), the calculated minimax radius exhibited a sharp, monotonic increase ($\partial r/\partial t > 0$), while the dispersion strength rate ρ_t decayed exponentially. This inverted geometric trend visually and mathematically indicates a macroscopically expanding contagion boundary, where newly infected clusters emerge progressively further away from the established epidemic epicenters.

Conversely, approaching the late phase of the tracking window (Days 5–7), the system stabilized into a clear steady state ($\partial r/\partial t \to 0$), causing ρ_t to converge asymptotically to its lower bound. This behavior mathematically substantiates that the geographical variance of the disease spread has hit an environmental saturation limit within the targeted spatial topology, meaning the boundary circle no longer needs to expand to capture outlier case clusters.

5 Conclusions

We proposed a multiperiod COVID-19 dispersion modeling framework leveraging minimax optimization. Mathematical verification and empirical validations confirm that tracking the rate of change of the minimax circle metrics provides a reliable tool for forecasting epidemiological dispersion trends.

Acknowledgement

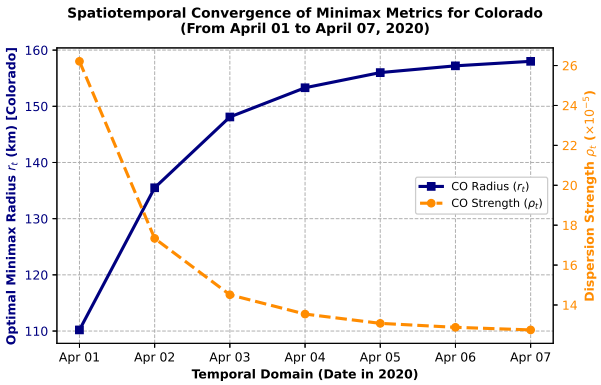
D.K. is the faculty of CGCC (Chandler-Gilbert Community College)/MCCCD (Maricopa Community College District).

Funding

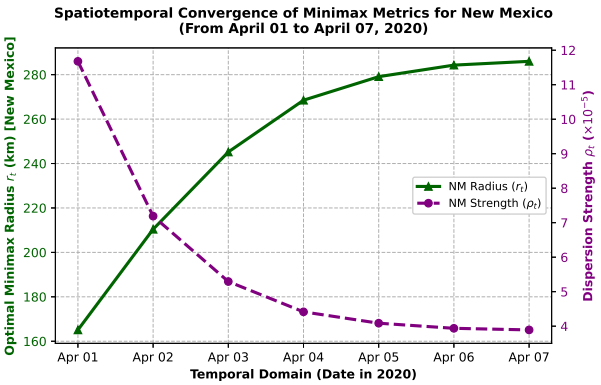
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

All data generated or analyzed during this research are included in this paper. The data citations are included in the reference. If you need to have further information, then email to the author.



(a) Colorado (CO) convergence profile.



(b) New Mexico (NM) convergence profile.

Figure 3: Dynamic spatiotemporal profiles presenting the optimal minimax radius (r_t, left axis) and corresponding dispersion strength (ρ_t, right axis) for (a) Colorado and (b) New Mexico during the period April 01–07, 2020.

Declarations (Conflicts of Interest)

The author declares that we have no conflict of interest.

Author contribution

D.K. conceived and designed the research problems. D.K. developed the mathematical modeling of data with minimax measurement. G.K. analyze the data and work with post-processing. D.K wrote the paper. G.K. and D.K. discussed the results and contributed to the final manuscript.

Ethical Statement

This study is a literature-based review and reanalysis of previously published experimental data. No original human clinical trials or biological sampling were conducted by the authors. As all analyzed data were obtained from published sources and contain no personal identifiers, institutional review board (IRB) approval was not required.

A Empirical Spatial Datasets for Geolocation Clusters

This appendix provides the full coordinate datasets extracted and normalized from the CDC and Google Health open repositories for the simulation of the minimax circle models. All coordinates (X_i, Y_i) are given in kilometers (km) and normalized relative to the primary epidemiological epicenter of each respective state. The Arizona dataset is presented in Table 1 of the main text; the supplementary datasets for Colorado and New Mexico are provided in Table 2 and Table 3 below.

Table 2: Normalized Euclidean coordinates of COVID-19 cluster points in Colorado.

Index	**County Name**	X_i **(km)**	Y_i **(km)**
1	Denver (Ref)	0.0000	0.0000
2	El Paso	12.1541	-98.4512
3	Arapahoe	45.3104	-5.1120
4	Jefferson	-24.6183	8.4152
5	Adams	18.4519	22.3041
6	Larimer	-15.1204	95.6124
7	Weld	52.3391	88.1049
8	Pueblo	-8.4112	-165.3541
9	Mesa	-345.1204	-12.5049
10	Boulder	-42.5049	35.1920

Table 3: Normalized Euclidean coordinates of COVID-19 cluster points in New Mexico.

Index	**County Name**	X_i **(km)**	Y_i **(km)**
1	Bernalillo (Ref)	0.0000	0.0000
2	Doña Ana	-12.4512	-310.1541
3	Santa Fe	22.3041	85.3104
4	Sandoval	-18.1120	45.6183
5	San Juan	-215.3541	221.4519
6	Valencia	-8.4152	-32.1204
7	McKinley	-198.6124	95.3391
8	Chaves	245.1049	-185.4112
9	Eddy	265.1920	-275.5049
10	Otero	88.5049	-241.6031

References

[1] Andrea L. Bertozzi, Elisa Franco, George Mohler, Martin B. Short, and Daniel Sledge. The challenges of modeling and forecasting the spread of COVID-19. *Proceedings of the National Academy of Sciences*, 117(29):16691–16696, 2020.

[2] Dongyung Kim. An optimized discrete data classification method in N-dimensional. *Computational and Mathematical Methods*, 2022.

[3] Centers for Disease Control and Prevention (CDC). COVID-19 case surveillance public use data with geography. `https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data-with-Ge/n8mc-b4w4`, 2020. Accessed: August 2026.

[4] Google Health. COVID-19 open data repository. `https://github.com/GoogleCloudPlatform/covid-19-open-data`, 2020.

[5] O. Wahltinez and et al. COVID-19 open-data: a global-scale spatially granular meta-dataset for public health research. *Scientific Data*, 9(1):175, 2022.