CCAO Airport Sound Analysis, 2022

Gabe Morrison, CCAO Data Intern

March, 2022

Abstract:

Airport noise might affect property values. In the past, the Cook County Assessor's Office (CCAO) captured the effects of O'Hare Airport noise on property values and property assessments using a hard boundary. Properties were either inside or outside of a significant O'Hare noise exposure region created by the Federal Aviation Administration (FAA). But there are issues with this approach: noise itself does not follow hard boundaries, the effects of airport noise on sale prices might not follow a hard boundary, and effects of Midway noise on property owners in southwest Chicago and nearby suburbs were not accounted for.

In this work, the CCAO empirically created a noise surface that accounts for noise from O'Hare and Midway airports. This kriging surface used 11 years of publicly available sound data from 53 noise monitoring stations near these airports. Instead of a hard "in or out" boundary, the surface is like a noise heat map for Cook County that captures estimated day-night average sound levels caused by O'Hare and Midway airports. The bulk of this report details the methodology to create and validate the kriging surface. We believe this airport noise surface reflects the true airport-induced sound level experienced by Cook County residents. The next output of this process is, for every residential parcel in Cook County, an estimate of that property's noise exposure. Additional work is needed to determine whether the noise surface is predictive of property values.

Introduction:

This report is structured as follows. The rest of the introduction section provides a brief introduction to some of the technical language used in the report and a description of previous data used by the CCAO to address sound quality issues. We subsequently describe data sources in more detail, introduce our methods to create the surfaces, summarize and visualize our results, and discuss some limitations of the work.

Technical Background and Language

The **Federal Aviation Administration (FAA)** is the federal government agency regulating all aspects of aviation and airports operations in the United States.

Day-Night Average Sound-Levels (DNL) is a metric used by the FAA to measure the effect of airport noise on surrounding communities. It reflects a location's cumulative sound exposure from annual aircraft operations on an average day in a given year. To determine DNL, the FAA uses a year's worth of data on flights from an airport and takes into account plane type, trajectory, and flight path, among other characteristics. DNL is measured in **decibels (dBA)**. To capture the effect of the increased disturbance of noise at night, DNL calculations add 10 dBA to flights which occur between 10:00 pm and 7:00 am. The FAA uses 65 dBA as a threshold for "significant noise exposure" ("Community Response…" 2022). The FAA produced the graphic below to provide intuition on the DNL of different areas.

We use the language "**noise surface**" to refer to a grid covering areas near Midway and O'Hare airports. We estimate DNL for each cell in the grid.

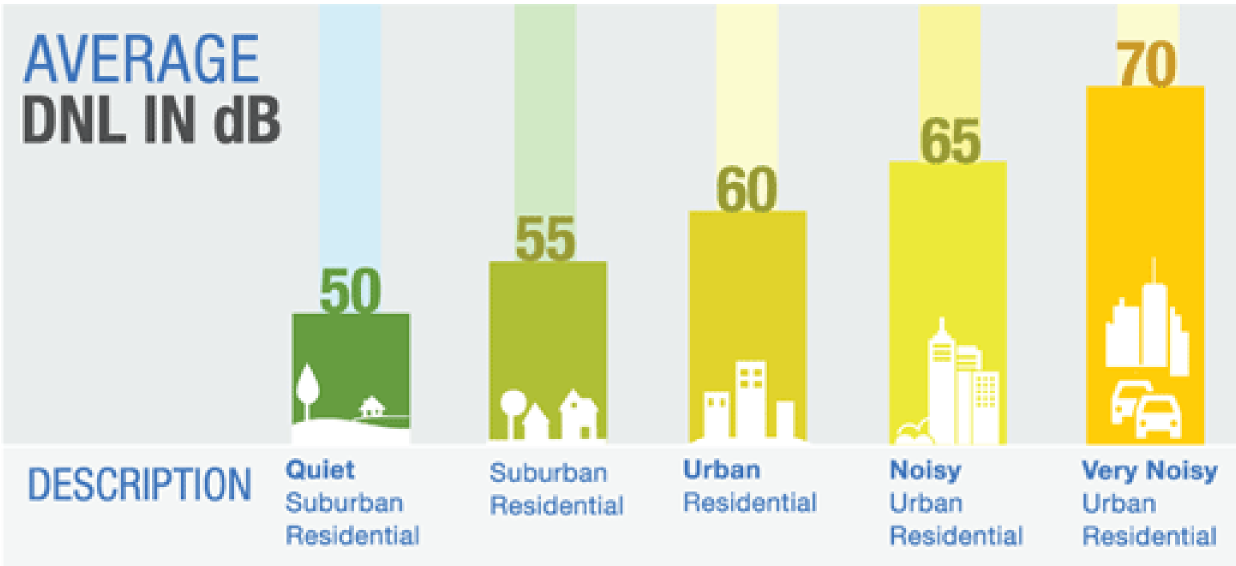
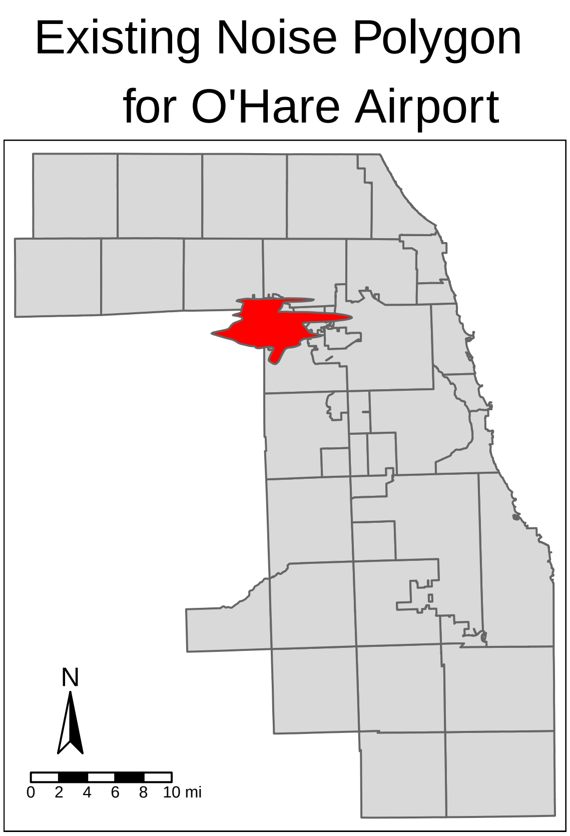


Figure 1: FAA DNL Illutration ("Community Response…" 2022)

Previous Work at the CCAO and Motivation:

Previously, the CCAO had used the polygon below. It represents the estimated areas surrounding O'Hare Airport that were exposed to 65 DNL or higher. The Data Team created a variable that indicated whether properties were within this polygon as a feature and used it in our model.

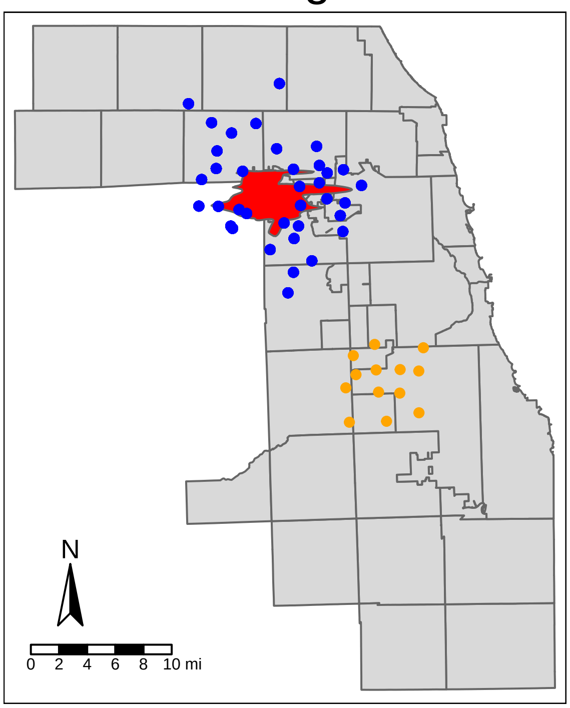


We build on our previous work for three reasons. First, preliminary analysis of model feature importance suggested that the variable created from the polygon was not useful in predictions. Second, while the FAA uses the 65 dBA threshold to determine "significant noise exposure," it seems reasonable to believe that this arbitrary threshold does not adequately capture the impact of airport sound on home buyers' preferences. That is, we would expect that, for each marginal increase in DNL, a potential home buyer would have slightly less willingness to purchase that house. Third, previous analysis did not incorporate the impact of Midway's noise on its surrounding properties.

Data Sources:

We used sound data collected from the Chicago Department of Aviation to build the noise surface. To assess the impact of sound on neighborhoods, the Chicago Department of Aviation uses a collection of noise monitors installed beginning in 1996. See Figure 2 for more details. There are 13 noise monitors surrounding Midway and 40 surrounding O'Hare ("Quarterly Reports"; "Introduction..." 2022). We used annual DNL values collected from reports about both Midway's and O'Hare's noise ("Quarterly Reports"). The locations of the monitoring stations can be seen on the map below. Orange dots reflect Midway sound monitoring stations, and blue dots show O'Hare monitoring stations.

Location of Monitoring Stations



We also used estimated DNL values based on the O'Hare Modernization Program (OMP). The OMP was an over decade-long effort to redesign O'Hare to increase capacity and minimize predictable delays (CDA Media Relations 2022). As part of these changes, the FAA modeled potential sound level calculations in DNL ("Executive Summary" 2005). We have access to these modeled DNL values at the noise monitor locations surrounding O'Hare and use them in our final surface.

Methods:

We created noise surfaces based on the DNL values of the noise monitors surrounding O'Hare and Midway airports using two sets of techniques: inverse distance weighting (IDW) and ordinary kriging.

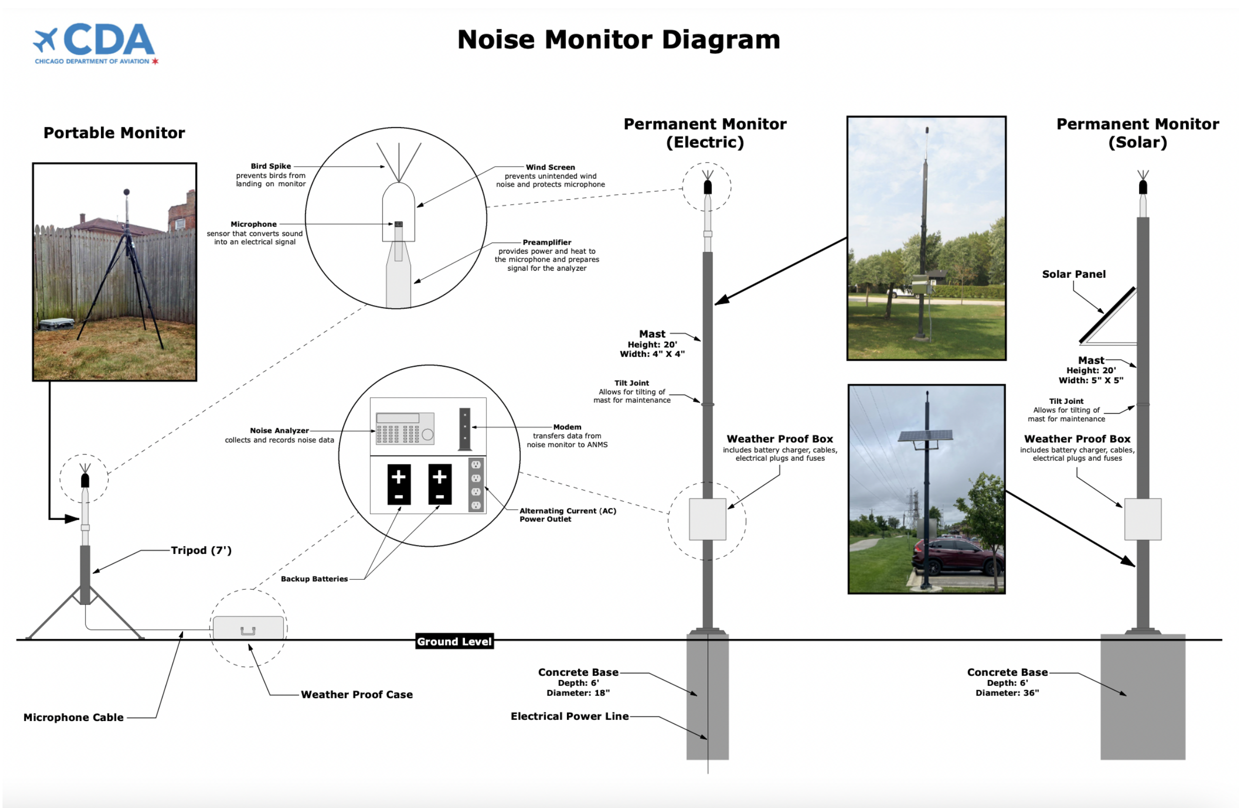


Figure 2: Chicago Department of Aviation Noise Monitor Diagrams ("Introduction to Noise Monitors… 2021")

Inverse Distance Weighting (IDW)

IDW interpolation assumes that the unknown value (in this case DNL) at any given location is the average of all values at other locations weighted by a function of the distance between the unknown location and all known locations. More formally, this is:

$$\hat{Z}(s_0) = \frac{\sum_{i=1}^{n} w(s_i) Z(s_i)}{\sum_{i=1}^{n} w(s_i)}$$

where

$$w(s_i) = \frac{1}{d(s_0, s_i)^p}$$

$\hat{Z}(s_0)$ represents the estimated value, in this case DNL, at the point s_0. $Z(s_i)$ represents DNL at the ith sound monitoring station. $w(s_i)$ is the weight of the DNL at the ith sound monitoring station on the unknown point at s_0. It is determined by computing the distance between the points s_0 and s_i to the power p. p determines how the strength of the weight changes over distance. With high p, only the nearest neighbors to an unknown point influence it. With low p, even far away points influence the estimate of an unknown point's value.

As an extreme example, imagine $p = 0$. This would mean that $w(s_i) = 1$ for all i. In this case, the equation above could be simplified to $\hat{Z}(s_0) = \frac{\sum_{i=1}^{n} Z(s_i)}{n}$. The sum of a set of values over the number of values is simply the mean. In other words, with $p = 0$, all points are weighted equally regardless of distance (Dorman 2022).

Ordinary Kriging:

The mathematics of ordinary kriging is somewhat more complicated, so we provide only high level details of the process and references to more comprehensive materials. However, the key point of ordinary kriging is that it is similar to but more complicated than IDW. It still uses values of known points weighted by distance between known points and an unknown location to estimate the value at the unknown location. However, it determines those weights flexibly based on a model built from the data (Gimond 2022).

The key steps to ordinary kriging are as follows:

1. **De-trend the data**: This could be as simple as subtracting the mean of the data from each observation or fitting a more complex model and using the residuals.
2. **Create an empirical semivariogram**: An empirical semivariogram captures the relationship between the distance between points with known values and the squared difference in their de-trended values (their semi-variance). In other words, in this step, one calculates the distance between all points and how related their de-trended values are.
3. **Fit a model to the empirical semivariogram**: This step builds a model that captures how the semi-variance changes as distance increases. Since these values are often noisy, it is common to first bin the semi-variances by distance and then fit a model to the binned average semi-variance values.
4. **Interpolate unknown locations with a variogram model**: Finally, using the model to understand the spatial relationship between unknown points and the values of the known points, interpolations of unknown points can be created (Gimond 2022; Pebesma and Bivand 2022).

For a more detailed but gentle introduction, see Gimond (2022), or for a sophisticated coverage, see Pebesma and Bivand (2022).

Leave One Out Cross-Validation:

The above description demonstrates that in interpolating these values, a modeler can choose between any variety of options. For IDW, a modeler can pick any positive value of p. For ordinary kriging, the modeler can choose to detrend the data with different functions, aggregate semi-variance values in differently sized bins, and fit those averaged values to different models. To test which parameters best fit the model, we used Leave One Out Cross-Validation (LOOCV).

LOOCV for this application involves the following steps:

1. Pick one sound monitoring location and remove it from the dataset
2. Build either an ordinary kriging or IDW surface using all points but the one point that was withheld
3. Record the predicted value at the location of the withheld monitoring surface
4. Repeat steps 1-3 for every sound-monitoring location
5. Compare the predicted values with the true, held-out values

For step 5, we evaluated the surfaces with Root Mean-Squared Error. Formally, this is:

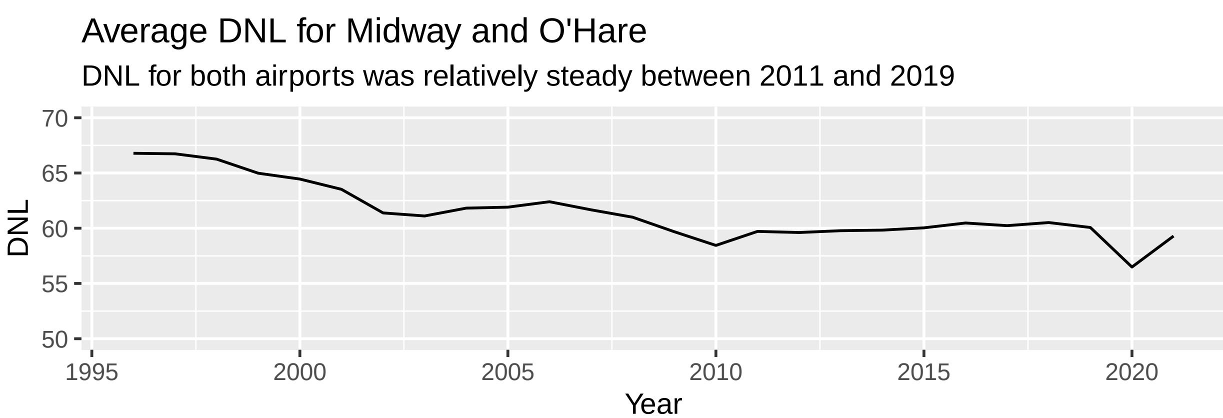
$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

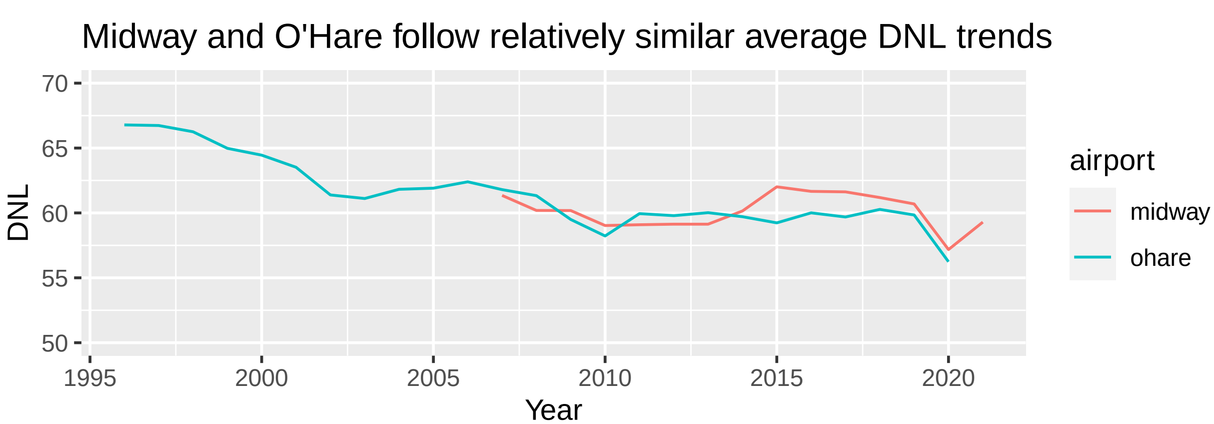
In the equation above, $\hat{y}_i$ is the estimated DNL at a point, and y_i is the true value at the point. We seek a low RMSE because that reflects that most estimated values were similar to their true values.

This method was inspired by, and builds off code from, Gimmond (2022).

Full Method Process:

We first created a dataset of all observed DNL levels for all noise monitoring stations around both Midway and O'Hare airports as well as the modeled OMP values for O'Hare. We subsequently plotted average DNL over time as illustrated in the chart below.





The charts illustrate that DNL was relatively consistent between 2011 and 2019. We believe that the drop in 2020 likely came as a consequence of dramatically reduced flights due to the COVID-19 pandemic. Regardless, based on the relative consistency between 2011 and 2019, we averaged the yearly DNL values at those locations.

With this data, we then tested IDW models with the p feature ranging from 1 to 20 and ordinary kriging surfaces with a variety of specifications. More specifically, we tested different de-trending functions: simple average, linear function of latitude and longitude coordinates, and function of latitude and longitude squared and interaction, following Gimond (2022). We also tested numerous bin width sizes to aggregate the sample semi-variance estimates. These ranged from 250 to 10,000 feet. Finally, we tested four different functions to fit to the experimental semi-variogram estimates. For the ordinary kriging, we performed "Grid Search" whereby we tested each parameter with each other parameter. For all these tests, we used a 1000 ft by 1000 ft grid for all areas within two miles of any O'Hare monitoring station and 3.5 miles of any Midway monitoring station.

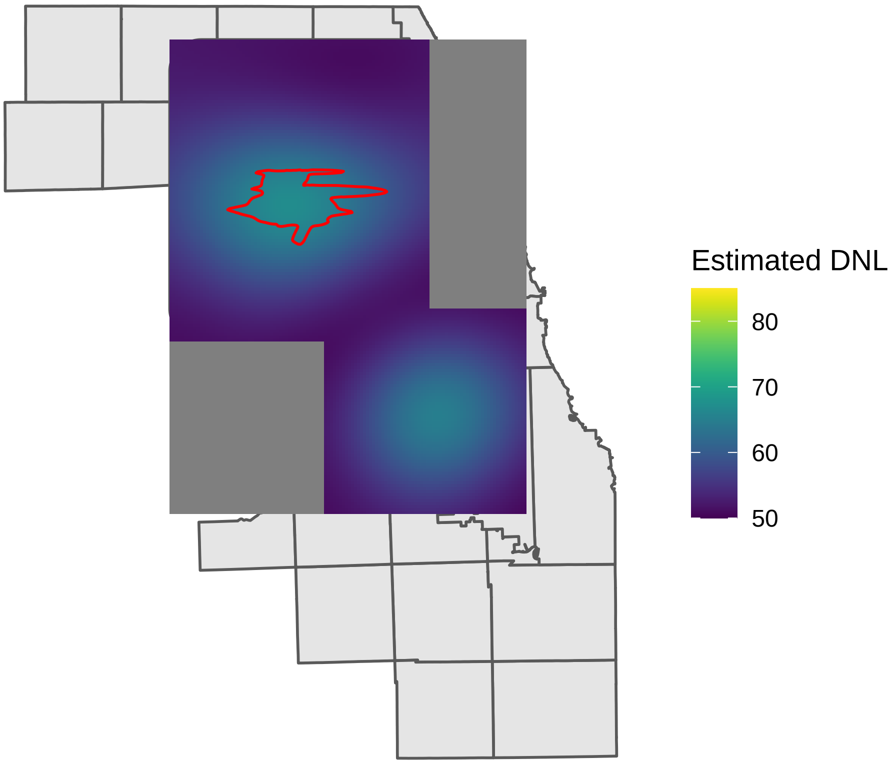
One challenge with this process was that noise monitors appear to be placed in noisy locations. Thus, we also tried adding points to the dataset at the locations of Midway and O'Hare airports with high DNL values and also points at the edges of the bounding boxes where there is effectively no airport noise (50 DNL). We experimented with different airport DNL values (80, 90, 100, 110) and compared the performance of the best performing model with and without these added points. We had the best model performance with the added points and a 90 DNL level.

To create a feature to be used in the property evaluation model, we re-ran the best-performing model using the data that best reflects future sound levels surrounding the airports. For Midway, this was the 2019 sound data (the last data before COVID). For O'Hare, these are the modeled OMP values. We interpolated this surface to a more fine grid with dimensions of 100 ft by 100ft. Finally, we determined the cell of the noise surface in which each parcel was located. Using that information, we joined the DNL of the cell to existing parcel data. We assigned parcels not covered by the kriging surface to be 55 if covered by the grid and 52.5 otherwise. These values are in line with the FAA's description of residential neighborhoods ("Community Responses…" 2022).

Results

Model Type	Airport and Neighborhood Points Added?	RMSE
IDW	No	5.13
IDW	Yes	4.94
Ordinary Kriging	No	4.97
Ordinary Kriging	Yes	3.86

We report the performance of our best surfaces above. Each row lists the RMSE of the model whose hyper-parameters correspond to the surface with the lowest RMSE of models of that type. The best performing overall model was an ordinary kriging model using the added airport and quiet neighborhood points. It used a Gaussian model to fit to the semi-variogram, 500 foot bins, and a simple average as the original de-trending function. This surface, built from average 2011-2019 values, is shown below.



Based on our best surface (using 2021 Midway data and OMP data for O'Hare), we determined the number of parcels exposed to different airport-related noise levels. We show both the percent and total number of residential parcels within each DNL range in the table below. Our results indicate that only about 4 percent of parcels are exposed to DNLs of 65 or higher.

DNL	Percent	Count
50 - 55	76.6	1,083,567
55 - 60	19.5	275,849
60 - 65	2.4	33,584
65 - 70	0.8	11,143
70 - 75	0.6	8,323
75 - 80	0.2	2,932
80 - 85	0.0	12

Limitations and Conclusion:

This work was not without limitations. To begin, access to data was limited. The number of sound monitoring stations around both airports, and particularly Midway, was limited, making the task inherently challenging. Because the stations are all relatively close to airports, estimating DNL farther away from the airports was challenging. We partially addressed this issue by adding some points at the edge of the bounding boxes. The general impact of this lack of observations can be seen in the surface's smoothness. A "true" noise surface is likely more angular and higher at areas aligned with airport runways, but it was difficult to capture this level of granularity. With respect to building the ordinary kriging surfaces, while we tuned the most important features, we could have assessed others and tried more variables to find an even better surface. Finally, we only used sound monitoring data. Similar exercises performed by the FAA use far more data including plane take-off and landings as well as private software. We used the publicly available yearly DNL averages of the sound monitoring and R, a powerful statistical computing program but that was not built specifically for this task.

Despite these limitations, we were able to build an ordinary kriging surface with an RMSE of less than 4. We believe that the final surface created reflects the true airport-induced sound level far closer than the previously-used 65 DNL threshold.

Citations:

CDA Media Relations. "Mayor Lightfoot and Chicago Department of Aviation (CDA) Celebrate the O'Hare Modernization Program (OMP)." *The Chicago Department of Aviation.* September 9, 2021. https://www.chicago.gov/city/en/depts/doa/provdrs/dbata/news/2021/september/992021.html. Accessed February 2022.

Dorman, Michael. "Chapter 12 Spatial interpolation of point data," *Introduction to Spatial Data Programming with R.* February 21, 2022. http://132.72.155.230:3838/r/index.html. Accessed February 2022.

"Executive Summary Final EIS Chicago O'Hare International Airport." *The Federal Aviation Administration.* July 2005. https://www.faa.gov/airports/airport_development/omp/eis/feis/Media/ExecutiveSummary.pdf. Accessed February 2022.

"Community Response to Noise." *The Federal Aviation Administration.* February 25, 2022. https://www.faa.gov/regulations_policies/policy_guidance/noise/community/. Accessed February 2022.

Gimmond, Manuel. "Chapter 14 Spatial Interpolation," *Intro to GIS and Spatial Analysis.* January 9, 2022. https://mgimond.github.io/Spatial/index.html. Accessed February 2022.

"Introduction to Noise Monitors at Chicago O'Hare International Airport." *The Chicago Department of Aviation.* December 2021. https://www.flychicago.com/SiteCollectionDocuments/Community/Noise/OHare/ANMS/ORD_Monitors_Introduction_FactSheet.pdf. Accessed February 2022.

Pebesma, Edzer and Roger Bivand. "Chapter 12 Spatial Interpolation," *Spatial Data Science." February 11, 2022. https://keen-swartz-3146c4.netlify.app/interpolation.html. Accessed February 2022.

"Quarterly Reports." *The Chicago Department of Aviation.* https://www.flychicago.com/community/MDWnoise/ANMS/Pages/ANMSreports.aspx. Accessed February 2022.