Mining, Metallurgy & Exploration
https://doi.org/10.1007/s42461-025-01241-0

An Algorithm for the Efficient Placement of Air Quality Sensors in Underground Mines

Kate Brown Requist[1] · **Moe Momayez**[1]

Received: 17 June 2024 / Accepted: 24 March 2025

Abstract
Current regulations for the placement of air quality sensors in underground mines in the USA have been subject to varied debate, especially regarding sensor locations and spacings. Ideally, sensor placement in underground mines should initially focus on areas where dangerous air mixtures are likely to be, such as along beltways or at the working face. Beyond this, however, lies a concern about the lack of guidelines for the placement of non-regulatory, fixed air quality monitors in underground mines. As more operations turn to real-time air quality monitoring, any configuration of sensors should consider a near-even distribution of sensors throughout these areas to support mine-scale estimation and monitoring of airborne contamination distributions. We present an algorithm for the efficient placement of air quality sensors (EPAQS) in underground mines that provides an optimized spatial distribution of sensors when considering the effects of the ventilation network's airflow. This algorithm provides suggested placements of air quality sensors in underground mining environments to provide an even statistical coverage of the ventilation system for future work in mine-scale contamination distribution estimation using spatial statistics.

Keywords Atmospheric monitoring systems · Graph theory · Combinatorics · Air quality · Underground mining

1 Introduction

The presence of carbon monoxide in underground coal mines has been documented as a health hazard since the sixteenth century [1]. With the development of improved real-time monitoring capabilities, the detection of carbon monoxide in large portions of underground US coal mines has been delegated to atmospheric monitoring systems (AMSs). These AMSs can consist of tens or hundreds of sensors used to detect carbon monoxide as an indicator of spontaneous combustion and human health risk [2]. The placement of sensors within underground coal mines is well-defined for critical areas, such as along beltways and escapeways, but there exists little guidance for the installation locations of sensors in other areas of the mine where carbon monoxide may still be present. Currently, the placement of sensors in these other regions is left to the discretion of ventilation engineers or industrial hygienists as per US 30 CFR §75.351(e)(1). This complicates the efficacy of additional monitoring, especially with the goal to move toward site-wide airborne contamination monitoring beyond current AMS capabilities. Currently, operations must rely on judgement and experience to guide non-random or worst-case sampling in non-regulatory areas. The regulations, as currently written for beltways and escapeways, highlight the need for consistent spacing of sensors to return a realistic measurement of mine atmosphere in the area monitored, to provide a mine-scale understanding of airborne contamination behavior.

Naturally, carbon monoxide is not the only airborne contaminant that constitutes a significant health hazard to underground workers, nor is it the only airborne contaminant that provides insight into the risk of mine fires and explosions. This is especially pertinent, given that AMSs are only employed in underground coal mines. Without a suitable analogue for other contaminants across underground mining method and commodity types, AMSs are a helpful starting point to discuss the merit of changing current approaches to airborne contamination monitoring in underground mines, especially at the mine-scale. Because AMSs for carbon monoxide monitoring are the only systems that currently have prescribed placement standards, real-time car-

✉ Kate Brown Requist
katebrown@arizona.edu

[1] Department of Mining and Geological Engineering, University of Arizona, Tucson, AZ, USA

Published online: 02 April 2025

bon monoxide monitoring provides a helpful starting point to considerations of the non-regulatory placement of air quality monitors in underground mines. The placement of AMS sensors is relegated to areas where fire risk is high. AMSs are purpose-built, and that purpose is to provide early warnings to operators of the risk of spontaneous combustion of coal or a mine fire [2]. As such, the regulatory expectations for AMS sensor placement are confined to the narrow use case of AMSs. Ideally, there should exist some method whereby the placement of sensors can be informed by the movement of air (the main driver of the movement of airborne contamination) to provide an even coverage of the mine's ventilation system. An ideal approach could focus on providing recommendations to operators based on the number of monitors available to the operation, which would provide a means for monitoring with even coverage without concerns of having some "proper" number of non-regulatory monitors. The development of an air quality sensor placement algorithm would alleviate operators' need to consider sensor placement when developing a mine-scale air quality monitoring system that extends well beyond modern AMSs.

Sensor spacings as per US 30 CFR §75.351(e)(1) for the detection of carbon monoxide along beltways is shown in Table 1. Similarly, US 30 CFR §75.351(f) requires the monitoring of carbon monoxide within 500 ft (approximately 150 ms) of any working area and within 500 ft inby the beginning of an active panel. These sensors provide vital health and safety information to surface AMS operators during the work cycle, but any additional monitoring does not currently have a widely accepted paradigm for the placement and spacing of carbon monoxide sensors within underground coal mines. Critically, the spacings as outlined by these regulations are intended to provide a snapshot of air quality in critical areas, despite a lack of consistent spacing.

The regulation of sensor spacing in US underground coal mines has gone through multiple iterations, largely starting in 2004. By August of 2004, underground coal mines in the USA using AMSs were expected to comply with the spacings made official in US 30 CFR §75.351(e)(1) [3]. After the Sago Mine methane explosion in Tallmansville, West Virginia in 2006 and the subsequent United States MINER Act, a technical study panel was formed to investigate the set of sensor spacing regulations and to define the role of an AMS operator more clearly [4, 5]. During discussions with the technical study panel, the placement of sensors was widely debated. Consensus was poor among operations already employing real-time carbon monoxide sensing, and some operations shared their doubt in the then-required spacings, arguing that the regulation already required too many sensors to be placed too close together, leading to regions of long beltways remaining comparatively data-poor [5]. In 2008, the technical study panel concluded their open comment period and compiled a new set of sensor spacing regulations that went into effect at the end of 2009. These changes were limited in scope; only increasing the number of sensors in some areas and clarifying the air velocities that would mandate changes in spacing. Since this time, no further regulatory changes have been made to sensor spacing requirements for underground coal mines using AMSs. Additionally, there exists no regulatory framework for air quality sensing in underground metal/nonmetal mines. Arguably, the lack of industry consensus and no regulatory framework for underground metal/nonmetal mines has posed a barrier to widespread real-time air quality monitoring in underground mining, especially for use in exposure assessment. This has left a need for guidance in placing sensors in other parts of mines, and there exists no common regulatory or best-practices framework for the placement of sensors in the US mining industry.

Spatial statistical methods are strong candidates for mine-scale airborne contamination monitoring, but paradigms need to be developed for numerically robust estimation of contaminant behavior between fixed air quality sensors [6]. Mine-scale contamination monitoring would facilitate holistic assessment of contamination transport at a scale that is not currently feasible under typical industrial hygiene personal and area monitoring approaches, especially non-random or worst-case sampling. With the intention of using fixed sensors for the estimation of airborne contamination distributions in future research, any statistical or machine

Table 1 AMS sensor spacing requirements pursuant to US 30 CFR 75.351(e)(1)

Location	Maximum separation between location and sensor
Upwind of tail pulley	150 ft
Branching or mixing of belt air	50 ft
Sensor-to-sensor with air velocity ≥100 ft per minute	1000 ft
Sensor-to-sensor with air velocity between 50 and 100 ft per minute	500 ft
Sensor-to-sensor with air velocity ≤50 ft per minute	350 ft
Downwind of any pulley or takeup	100 ft

learning method will require a relatively even-spaced sensor placement to avoid over- or under-representation of portions of the mine's ventilation system. In spatial statistics, this spacing is vital. Spatial statistical methods, like kriging, include the locations of known data points relative to all other known data points as a basis for estimation [7, 8]. The over-representation of data in certain locations can harm a spatial statistical model's accuracy and estimation variance and can further complicate actions like the construction of a variogram in the case of kriging. In the case of geostatistics, geological data is normally much less scarce and better-spaced when compared to data availability for air quality in underground mines. A typical geostatistical workflow includes declustering of data, which seeks to minimize over-representation of known data in the estimation. Declustering methods are numerous, with the most well-known method arguably being cell declustering [9–11].

Declustering methods seek to condition input data prior to statistical estimation within geostatistical workflows. Input data is typically abundant but may be spread unevenly, causing clusters of data. These clusters lead to statistical over-representation of parts of the estimation domain. This can negatively impact the quality of kriged estimates as well as have major implications on the stability and validity of the kriging equations for estimation. To remedy this, declustering weights can be assigned to input data. The application of declustering weights attempts to provide a density-weighted average of input data at a specific location rather than multiple varying values that contribute identically to the input dataset. This helps to avoid singular matrices while solving the kriging system of equations and can help to ensure the covariance matrix remains positive definite. By providing an even statistical coverage, areas of the estimation domain are less likely to be over- or under-represented during estimation [9–11].

When considering the current state of regulatory and non-regulatory air quality monitoring in underground mines, data is much scarcer and spaced significantly farther apart when considering the relative scales of a drilling program to an active mine's ventilation system. This means that declustering methods like cell declustering are not feasible to ensure even spatial representation in estimation. With the steep cost of real-time monitoring, it is not realistic to use declustering methods on monitoring data. It would otherwise be best to develop a placement paradigm that removes a need for declustering of data prior to spatial statistical estimation of contamination distributions. The volume of data is already low, and expansion of the monitoring framework is costly. Instead, priority should be given to methods that focus on the even spacing of sensors to promote an even statistical representation of input data. In turn, this would reduce the number of sensors required to estimate a contamination distribution when compared to sensors that have been naively placed, requiring declustering nonetheless. The construction of such a method will improve current capabilities to model contamination distributions without reliance on the typical mine ventilation network approaches that operations are currently beholden to.

Methods for placing sensors in underground mines have been developed previously, especially in the past two decades. In 2022, Liu et al. developed an algorithm based on graph representations of mine ventilation systems designed to place air velocity sensors in a mine using the fewest sensors possible and spacing the sensors in the network such that the flow distribution of the network could be estimated based on the monitoring data [12]. Similarly, in 2023, Yan et al. developed a method for air velocity sensors again using a graph representation of the ventilation network, but this time allowing for a user-selected number of sensors [13]. These methods rely on a traditional graph representation of the mine ventilation network and were explicitly designed for the estimation of air quantity using ventilation network software. With a desire to use spatial statistics to estimate contamination distributions, these algorithms are unfit for adoption. In 2017, Muduli et al. developed a method to suggest the placement of wireless sensors for the monitoring of air quality in longwall mines [14]. Critically, the goal in Muduli et al. (2017) was to consider wireless connectivity over the area. The method presented is not intended to consider the effects of air movement and its impact on the transport of airborne contamination. This angle of the problem has been considered across multiple works and across a wide array of wireless network topologies [15–19]. The spacing of wireless air quality sensors should still consider the limitations and spatial requirements of wireless sensor networks.

With the interest of developing a mine-scale system for airborne contamination distribution monitoring based on spatial statistics, there are no existing methods to guide the placement of sensors in a manner that promotes a statistically robust estimation of contaminant behavior. For monitoring air quality using spatial statistics, the input data must be spaced such that there is an even statistical coverage across the entire ventilation system. This would provide a supplementary means of area monitoring of air quality across an entire mine by way of a contamination distribution as can be simulated using computational fluid dynamics or ventilation network solvers. A spatial statistics-based monitoring method stands to provide mine operators and industrial hygienists with another layer of real-time data at a resolution has been previously unfeasible.

A lack of current methods for sensor placement and the time-intensive nature of calibration of sensors poses a large barrier to operational efficiency as well as health and safety capabilities when faced with the large costs (e.g., monetary/acquisition, time, labor) associated with real-time atmospheric monitoring [6]. To address these concerns, we

consider the even spacing of a known number of sensors in a space. We propose an algorithm for the efficient placement of air quality sensors (EPAQS) that is based on ventilation network airflow and contaminant transport modeling to ensure the spatial partitioning accurately represents the flow of air through the space to create the most realistic picture of real-time mine atmospheric conditions. Formally stated, the EPAQS algorithm is an algorithm that seeks to optimize the partition of vertices of a complete, bidirected graph into a set of vertices of complete bipartite graphs of form $K_{1,k}$ (otherwise referred to as star graphs) where the union of each star graph contains every vertex in the original complete, bidirected graph.

The EPAQS algorithm aims to improve the decision-making process for the placement of sensors where a widely accepted paradigm does not currently exist. Non-regulatory air quality monitoring has expanded as a critical priority for worker health in the US mining industry, especially on the heels of changes to regulatory exposure limits for respirable crystalline silica (RCS) and diesel particulate matter (DPM) [20–22]. Regulatory monitoring of RCS and DPM is expensive, and most operations do not have access to the equipment needed to analyze samples on-site. While other work has primarily focused on the development of non-regulatory, real-time personal monitors, consideration should likewise be given to the use of area monitoring as a helpful tool for the exposure assessment of underground miners. Unlike carbon monoxide monitoring under the current AMS framework outlined in 30 CFR §75.351, there does not exist any regulatory or best-practices framework under which mine-scale area monitoring of RCS or DPM should be conducted, especially with regard to the spacing of real-time monitors. Even sensor spacing paired with spatial statistical methods for site-wide contamination monitoring would provide a pathway to a robust understanding of contamination transport across a mine's ventilation system.

Further, current AMSs have an average of 38 monitors in underground coal mines in the USA [2]. These sensors are concentrated in critical areas, but there exist no formalized guidelines for the inclusion of additional sensors with regard to placement. Real-time air quality monitors remain costly, especially when considering the time required to calibrate and maintain the sensors. This poses a significant barrier to collecting sufficient volumes of quality data across the mine for deeper analysis because the addition of sensors to any monitoring platform means an increase in cost to install, operate, and maintain the monitoring system. With the relative scarcity of area monitoring data, especially well-integrated and longitudinal data, we aim to provide a method to aid in the placement of fixed air quality sensors. The goal of the algorithm presented here is to improve the current state of non-regulatory area monitoring with a special focus on eventual applications for miner health.

Even spacing of air quality sensors will allow for the use of spatial statistical methods to estimate airborne contamination distributions. Further, this will prevent the installation of unnecessary sensors whose data must otherwise be declustered prior to statistical estimation of a contamination distribution. This spacing requirement is vital for quality estimation of contamination distributions using statistical means, but no available methods currently exist for such an application. This is especially complicated when considering the complex dynamics of contamination and airflow in underground mines. With the use of the EPAQS algorithm, we attempt to approximate a solution to the capacity-constrained centroidal Voronoi diagram problem outside of methods that have been developed for Euclidean spaces with or without additive or multiplicative weights. This even spacing will likewise provide a more clear picture of underground mine contamination behavior without the need to install sensors through trial-and-error to determine the best placement. By providing this method, we hope to reduce the time needed for operators to determine locations to install non-regulatory monitors with a special focus on providing an even statistical coverage of the mine's ventilation system.

2 Methods

The EPAQS algorithm is designed to optimize the spatial distribution of sensors in an underground mine when considering the effects of ventilation network airflow. Because air moves through the mine with varying speed and direction, the most representative spatial distribution to cover all regions of the underground mine equally is not necessarily the same as one that would be constructed using a traditional distance metric, such as a Euclidean or Manhattan distance. As such, the EPAQS algorithm considers the effective distance between a sensor and a neighboring point to be a more appropriate measure. This effective distance value is derived via a pathfinding algorithm that is an extension of the well-known A* pathfinding algorithm. This algorithm is described in detail in Brown Requist and Momayez (2024) [23]. This method considers the effects of pressure losses and surface roughness in conjunction with airflow directions as a means of parameterizing an effective distance between any two points in an underground ventilation system.

To use the modified pathfinding algorithm, a mine ventilation model can be used to inform airflow directions. Figure 1 displays a portion of an underground room-and-pillar coal mine in Utah, USA, where the network ventilation model was exported to CAD so all airways could be annotated with their respective direction of airflow. Effective distances were calculated using the modified A* algorithm presented in Brown Requist and Momayez (2024) with a pressure loss of $\Delta P = 0.5$ and a surface roughness of $e/D = 0.025$.

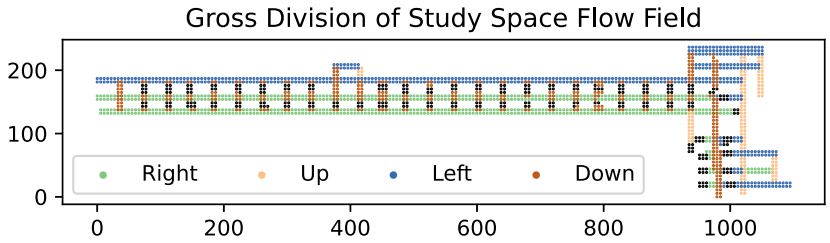

Fig. 1 Mine ventilation system exported for use with the modified pathfinding algorithm. First published in Brown Requist and Momayez (2024) [23]

2.1 Capacity-Constrained Centroidal Voronoi Diagrams

To evenly space points (i.e., sensors), we propose the use of a capacity-constrained centroidal Voronoi diagram. This is a classic problem in combinatorics that additionally has applications in graph theory. The capacity-constrained centroidal Voronoi diagram problem can be grouped under the larger umbrella of optimal set partition problems, for which exact computational solutions are famously NP-complete [24]. Typically, the capacity-constrained centroidal Voronoi diagram is considered in Euclidean space or in Euclidean spaces with the consideration of additive or multiplicative weight fields.

A centroidal Voronoi diagram is a set of locations that act as the centroid of some number of cells in a continuous or discrete space [24]. These cells are effectively sets of locations in any discrete, Euclidean formulation of the problem. Voronoi diagrams have wide-ranging applications, but a simple example might be the decision between visiting a business with multiple locations. Individuals, generally, would more preferentially select the closest location over another based on convenience. The various business locations constitute the centroid of the Voronoi cell, and the surrounding areas for which one location is the most convenient can be represented as the area of the Voronoi cell.

The size of a cell of a Voronoi diagram can therefore be realized as the cardinality of the set of locations (i.e., the area) that form the distinct cell within the partition. For applications of this paper, the cardinality of a cell V can be indicated as $|V|$ and is essentially equivalent to the number of locations that are closest to a given sensor at the cell's centroid and no other sensor. The construction of a Voronoi diagram can be generalized as a set S of m locations in some subset Ω of $\mathbb{R}^n$ that partitions the subset of $\mathbb{R}^n$ into m cells. Each cell, V_i, is defined by a location s_i such that $s_i \in S$. V_i is a set of points $x \in V_i \subset \Omega \subseteq \mathbb{R}^n$, such that any point x in V_i is closer to the location s_i compared to any other location $s_j \in S, i \neq j$. The resulting partitions defined by the general Voronoi diagram are referred to collectively as the Voronoi tessellation, $\mathcal{V}(S)$ of set S in a subset of $\mathbb{R}^n$ [24]. There is an important stipulation to Voronoi diagrams, however, that any point x in cell V_i may not be in any other cell. Stated otherwise, $x \in V_i$ requires $x \notin V_j$.

In the case of an underground mine's ventilation system, the space cannot be realized as strictly Euclidean. Simply stated, the centroid of a cell is not necessarily the geometric center of the cell. The movement of air plays a major role in the transport of airborne contamination, and this must be considered while constructing the Voronoi tessellation with the use of the modified A* algorithm. In every case, centroids must first be considered prior to the construction of a centroidal Voronoi tessellation. To create a centroidal Voronoi tessellation, we must define a bounded space $\Omega \subseteq \mathbb{R}^n; S \in \Omega$ such that some location $s_i \in S$ is the centroid of the cell V_i [25].

From this definition, the centroid of the cell, p_i, can be calculated in a continuous case as

$$p_i = \frac{\int_{V_i} x\rho(x)\,dx}{\int_{V_i} \rho(x)\,dx} \tag{1}$$

where $\rho(x)$ describes some non-negative density at location x, where $\rho(x)$ is defined, at minimum, for all $x \in \Omega$. This equation can be generalized as the density-weighted average location within the cell V_i [26]. The inclusion of a density function $\rho(x)$ implies the presence of an additive or multiplicative weight field, where the centroid of a Voronoi cell is not necessarily the geometric center of the cell. This is a classical problem in the capacity-constrained centroidal Voronoi diagram problem. This formulation stems from the power diagram problem itself, which is highly applicable to algebraic and computational geometry and numerous practical applications like computer graphics [25]. The determination of centroids is predicated on a comprehensive understanding of the density function, $\rho(x)$, for each Voronoi cell, V_i, in the bounded space Ω containing the Voronoi diagram of $\mathcal{V}(S)$. The density function $\rho(x)$ must be continuous, integrable, and known, which is not feasible for a mine ventilation system.

2.2 Design of the EPAQS Algorithm

The application presented in this work presumes that distances between locations in the mine's ventilation system cannot be recognized in a Euclidean manner. This is obvious when considering that air, and by extension airborne contamination, tends to move from upwind (higher pressure)

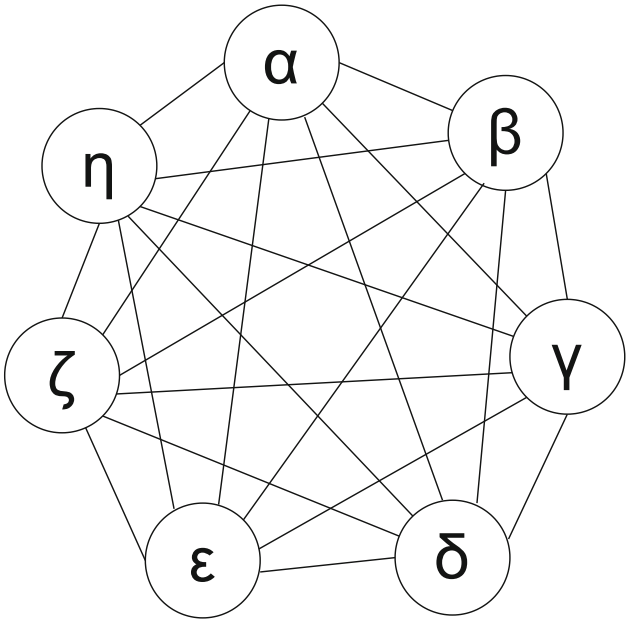

Fig. 2 A complete graph. Each vertex is connected to every other vertex within the graph. Graph edges may be weighted or directed

locations to downwind (lower pressure) locations. Consider two points, A and B. A is upwind of B, so air tends to preferentially move from location A to location B. This, of course, neglects the possibility of turbulent flow or reversal in the turbulent boundary layer near the walls of an excavation, but this statement holds, by and large. If we wish to represent this space as a mathematical construct, we must give priority to the fact that contamination is more likely to flow from A to B than B to A. To accomplish this, we can realize the mine's ventilation system as a complete, bidirected graph. This is managed by using the modified A* algorithm presented in Brown Requist and Momayez (2024) to assign up- and downwind effective distances to all pairs of points in the ventilation system [23].

A graph is a mathematical structure that concerns a set of objects or points (vertices) and any existing pairwise relationships between two items within that set (edges). Figure 2 provides an example of a complete graph. Complete graphs require that all vertices are connected to every other vertex. Graphs may be directed, meaning that pairwise relationships may only occur in on direction. Additionally, these pairwise relationships can be quantified as weights within the graph. This assigns a numerical value to the pairwise relationship. These weights can also observe direction.

Consider a mine's ventilation network that can be represented as a directed and weighted graph, indicating the direction of airflow and the quantity through each branch. Here, the intersection of branches forms a vertex of the graph, and an airway within the network forms the edge of the graph. Figure 3 provides an example of a simple ventilation network, with arrows indicating airflow direction (directed graph) and numbers along the graph's edges representing airflow quantity (weighted graph). Note that this graph is not complete, as not every vertex is connected to all other vertices.

The EPAQS algorithm requires a departure from the directed, weighted graphs that are used in ventilation network analysis. Rather than representing mine airways as edges of a graph, the EPAQS algorithm uses a complete, bidirected graph. In this complete, bidirected graph, discrete locations within the mine's ventilation system (as outlined in Fig. 1) are the vertices of the graph, and the up- and downwind effective distances constitute the weights along the edges connecting the vertices.

Figure 4 shows an example of a complete, bidirected graph using six vertices. Complete, bidirected graphs used by EPAQS have been tested up to 2500 vertices. The goal of the EPAQS algorithm is to select some subset of vertices contained in the complete, bidirected graph and create distinct subgraphs. These subgraphs, as shown in Fig. 5, are complete and bipartite subgraphs of the original complete, bidirected graph. Complete bipartite graphs of this form are specifically referred to as star graphs. A bipartite graph is a graph whose vertices can be separated into two distinct sets, in which a vertex is connected to all other vertices outside of its set, but none within its set. This is a helpful framing of the problem to adopt for sensor placement: the sensor location is the sampling location, and that sensor has a range of influence over all locations closest to that sensors. In another framing, the sensor and sensor location form one unique set with one vertex, and all locations within the range of influence of the sensor form the other set that the sensor must have a connection to via the edge weights.

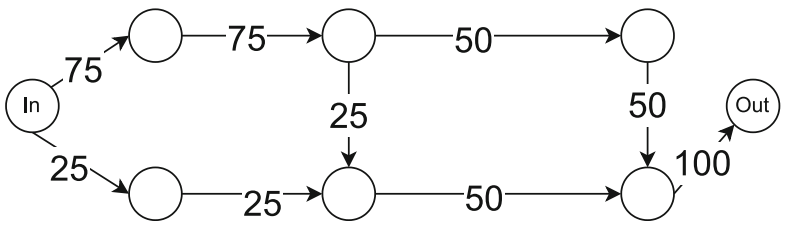

Fig. 3 Simple ventilation network using a graph representation of the network. Air quantity is indicated along the edges of the graph, with arrow direction indicating the direction of the edge

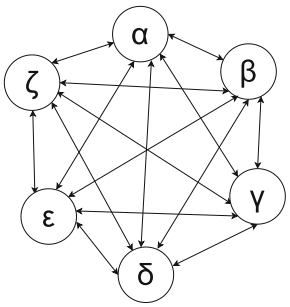

Fig. 4 Example complete, bidirected graph

Using Fig. 5 as an example, the complete bidirected graph in Fig. 4 can be partitioned into two subgraphs of form $K_{1,2}$, or stars with two branches. From this partition, we could state that the sensors are located at points α and δ, and α and δ additionally provide statistical coverage for locations β and ζ, and γ and ϵ, respectively. This is the general function of the EPAQS algorithm, but the weighted nature of the original complete, bidirected graph causes additional logistical challenge for arriving at a perfect partition.

Let the mine's ventilation system be realized as a complete, bidirected graph $\mathcal{G}$. Graph $\mathcal{G}$ contains vertices that represent a set of discrete locations in the mine's ventilation system and bidirected edges that represent the effective distance between locations as given by the modified A* pathfinding algorithm that considers airflow. We seek to partition $\mathcal{G}$ into a set $\mathcal{V}$ of star graphs V of form $K_{1,k}$ where the number of vertices in each subgraph is equal; $|V_i| = |V_j| \, \forall i, j \in \mathcal{V}$. Additionally, to satisfy the construction of a proper capacity-constrained Voronoi tessellation, no vertices may be repeated between two star subgraphs of $\mathcal{G}$, and all vertices in graph $\mathcal{G}$ must be present in the union of subgraphs $V \in \mathcal{V}$. This is complicated with a discrete representation of a space, such as that presented using a graph. A perfect partition may not be possible, but care should be taken to approximate the perfect partition.

Lemma 1 *A perfect capacity-constrained partition may not be feasible, as this requires all star subgraphs to only contain vertices that are unique to the star graph* (Fig. 6).

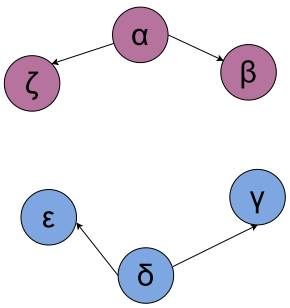

Fig. 5 Partition of the graph into star subgraphs

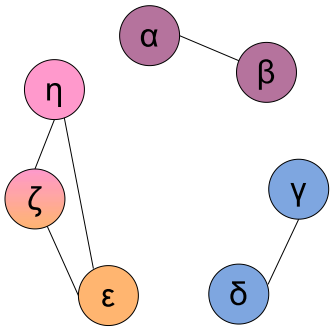

Fig. 6 Graph representation of the condition presented under Lemma 1

Proof By way of contradiction, assume that all complete graphs can be partitioned into star graphs and that these star graphs all contain unique vertices.

Consider a complete graph $\mathcal{G}$ with vertices in the set $\Omega = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta\}$. We wish to partition this graph into four star graphs. Disregarding a desire to group vertices by spatial proximity or lowest weight in the bidirected complete graph, we could create four sets of these vertices, A, B, C, and D. Let $A = \{\alpha, \beta\}$, $B = \{\gamma, \delta\}$, $C = \{\epsilon, \zeta\}$, and $D = \{\eta, \zeta\}$. Vertex ζ (or some other vertex under this arbitrary partition) must be used twice.

To meet the capacity-constraint requirement to partition graph $\mathcal{G}$ into subgraphs of form $K_{1,1}$, at least one subgraph must contain a vertex that is already contained in another subgraph. This shows that a perfect capacity-constrained partition may not be feasible, because some partitions may lead to the over-representation of one or more vertices found in graph $\mathcal{G}$. □

Lemma 2 *A perfect capacity-constrained partition may not be feasible, as this requires the union of all star graphs to contain all vertices that are in the complete bidirected graph* (Fig. 7).

Proof By way of contradiction, assume that all complete graphs can be partitioned into star graphs and that the union of the sets of vertices in these star graphs is equivalent to the set of vertices in the complete bidirected graph.

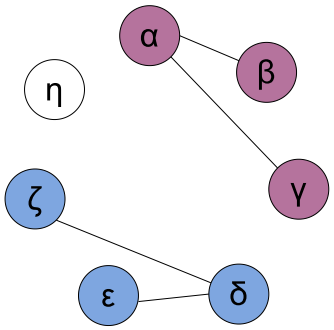

Fig. 7 Graph representation of the condition presented under Lemma 2

Once again, consider a complete graph $\mathcal{G}$ with vertices in the set $\Omega = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta\}$. We now wish to partition this graph into two star graphs. Disregarding a desire to group vertices by spatial proximity or lowest weight in the bidirected complete graph, we could create two sets of these vertices, A and B. Let $A = \{\alpha, \beta, \gamma\}$ and $B = \{\delta, \epsilon, \zeta\}$. Vertex η (or some other vertex under this arbitrary partition) is not present in $(A \cup B) = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta\}$.

To meet the capacity-constraint requirement of the partition graph $\mathcal{G}$ into star subgraphs, the union of subgraphs containing vertices in sets A and B must omit at least one vertex. This shows that a perfect capacity-constrained partition may not be feasible, because some partitions may lead to the omission of one or more vertices found in graph $\mathcal{G}$. □

Lemmas 1 and 2 highlight the two requirements for a perfect partition of the complete, bidirected graph $\mathcal{G}$ into subgraphs in $\mathcal{V}$. A perfect partition would have subgraphs of the form $K_{1,k}$, where $k = \frac{|\Omega \setminus S|}{|S|}$. Here, Ω is the set of all vertices that form the basis for the complete, bidirected graph $\mathcal{G}$, and S is the set of locations that form the bases for the subgraphs V that form the tessellation $\mathcal{V}(S)$. This value k must be equal to the number of total vertices in $\mathcal{G}$, less the number of locations in S, divided by the number of locations in S. As outlined through Lemmas 1 and 2, a perfect partition of $\mathcal{G}$ is not guaranteed. This is additionally complicated when considering the bidirected nature of $\mathcal{G}$.

Lemma 3 *A perfect capacity-constrained partition may not be feasible when considering a bidirected graph, because the relative rank of vertices in k for multiple graphs of form $K_{1,k}$ may lead to over- or under-representation of vertices.*

Proof By way of contradiction, assume that all complete, bidirected graphs can be partitioned into star graphs and that no vertex present in the bidirected graph is duplicated or omitted in the star subgraphs.

Now, consider a complete graph $\mathcal{G}$ that is bidirected with vertices in the set $\Omega = \{\alpha, \beta, \gamma, \delta, \epsilon, \zeta\}$. We now wish to partition this graph into two star graphs. Assume we select the set S of our locations as $S = \{\alpha, \beta\}$. Let the respective weights from α to γ, δ, ϵ, and ζ be notated as $W_\alpha = \{1, 2, 3, 4\}$. Similarly, let the weights from β to the other vertices be notated as $W_\beta = \{4, 2, 3, 1\}$. Although $\mathcal{G}$ is a bidirected graph, for the purpose of constructing the centroidal realization of $\mathcal{V}(S)$ from $\mathcal{G}$, only weights in the direction from a location in S to another vertex are relevant.

Because of the centroidal constraint, locations α and β must select the k vertices from their respective sets of weights that ensure that the sum of weights is the smallest. Note that vertex δ has an identical weight from location α to δ as from location β to δ.

To meet the capacity-constraint requirement to partition graph $\mathcal{G}$ into subgraphs of form $K_{1,2}$, the union of subgraphs containing vertices in sets A and B must omit at least one vertex. This shows that a perfect capacity-constrained partition may not be feasible, because some partitions may lead to the omission of one or more vertices found in graph $\mathcal{G}$. This results in the set A of vertices in the subgraph with location α being $A = \{\alpha, \gamma, \delta\}$. Similarly, for the subgraph with location β, the set of vertices in the subgraph is $B = \{\beta, \zeta, \delta\}$.

The vertex δ is shared between the subgraphs that form the partition $\mathcal{V}(S)$ of $\mathcal{G}$. While the capacity constraint holds, vertex δ is over-represented in the subgraphs, and vertex ϵ is omitted entirely. As a consequence of the bidirected nature of the graph $\mathcal{G}$ and the centroidal and capacity requirements for the partition $\mathcal{V}(S)$, a perfect partition of a complete, bidirected graph into subgraphs of form $K_{1,k}$ is not always feasible even when the number of vertices in $\mathcal{G}$ less than the number of locations in S is divisible by the number of locations in S. □

Together, Lemmas 1, 2, and 3 highlight the need to construct a means to gauge the performance of a partition of $\mathcal{G}$. A perfect partition of $\mathcal{G}$ is possible, but is by no means guaranteed. As such, a measurement of overlap between cells or subgraphs V in $\mathcal{V}(S)$ can provide an easy basis from which to optimize the partition using the EPAQS algorithm. Figures 6 and 8, corresponding to conditions under Lemmas 1 and 3, each display overlap that is measurable. This overlap serves as a good metric for partition performance, as the goal of the EPAQS algorithm is to provide even spacing or statistical coverage through a space. With fewer vertices duplicated between subgraphs and the capacity constraint, a perfect partition has no overlap whatsoever. To measure this overlap, the number of vertices shared between two cells V in $\mathcal{V}(S)$ can be used to provide an indicator of partition performance as in Eq. 2.

$$\chi_{ij} = \frac{|V_i \cap V_j|}{k} \tag{2}$$

where χ_{ij} is the number of vertices shared between subgraphs V_i and V_j in $\mathcal{V}(S)$ as a fraction of the number of vertices not in S, k, present in each subgraph of $\mathcal{V}(S)$. The average value for χ_{ij} for all pairs of subgraphs, $\bar{\chi}_{ij}$, can likewise be

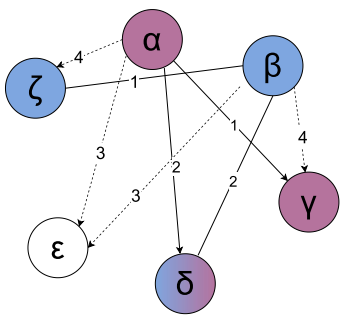

Fig. 8 Graph representation of the condition presented under Lemma 3

constructed as a general measure for the performance of a partition when considering all subgraphs V in $\mathcal{V}(S)$.

In the case of the partition presented under Lemma 3, we can use Eq. 2 to assess the performance of the partition. Let V_i represent the subgraph containing vertices $\{\alpha, \gamma, \delta\}$, and let V_j represent the subgraph containing vertices $\{\beta, \delta, \zeta\}$. The intersect of V_i and V_j is the set $\{\delta\}$, which has a cardinality of 1. There are two vertices not in S in each subgraph, so $k = 2$. Following the equation, $\chi_{ij} = 0.5$ for the pair of subgraphs presented in Lemma 3. When assessing a full-scale graph under the EPAQS algorithm, each pair of subgraphs created in the partition of $\mathcal{G}$ is compared, and the overlap between pairs is measured using Eq. 2 for optimization of the partition.

To optimize the partition of graph $\mathcal{G}$, χ_{ij} should be minimized. This will work to promote a partition that does not have large overlaps of vertices between any two subgraphs V_i and V_j in $\mathcal{V}(S)$. By constraining the maximum value of χ_{ij}, installation locations can be re-selected from a group of potential pairs of installation locations such that a partition that satisfies $\chi_{ij} < \chi_{max}$ will return an overlap between all pairs of cells V_i and V_j in $\mathcal{V}(S)$ that is necessarily less than or at worst equal to some maximum value χ_{max}.

Prior work from Lloyd (1982) suggests that the capacity-constrained Voronoi tessellation should first assume that the tessellation occurs in a bounded, non-infinite space Ω that is a proper subset of $\mathbb{R}^n$ [27]. Some initial tessellation $\mathcal{V}_0(S)$ should be created such that there is an arbitrary partition of Ω. All locations $s_i \in S$ are moved to the centroid ρ_i of $V \in \mathcal{V}(S)$ via least-squares quantization. This allows for the measurement of ρ to describe the distribution of points $x \in V \subset \Omega$. Cells V in $\mathcal{V}(S)$ are checked against the capacity constraint, and if the constraint is not met (i.e., the tessellation does not converge), Ω is once again arbitrarily partitioned until some tessellation $\mathcal{V}_n(S)$ converges. This method clearly has a risk of running into incredibly long times required for convergence, and further work by Balzer, Schlömer, and Deussen (2009) saw acceleration with some changes to data preprocessing and the use of heap data structures [26].

The methods presented by Lloyd (1982) and Balzer, Schlömer, and Deussen (2009) require that the space $\Omega \subset \mathbb{R}^n$ is realizable as a Euclidean space [26, 27]. Because $\Omega \subseteq \mathbb{R}^n$ is instead realized as a set of vertices that inform the complete, bidirected graph $\mathcal{G}$, the density function obtained via least-squares quantization in both methods is unattainable in a mine ventilation application. This gives rise to the need for the overlap measurement, χ_{ij}, because locations $s \in S$ are initially assigned as the centroid of cells $V \in \mathcal{V}(S)$.

To select the locations that should form set S as the basis of $\mathcal{V}(S)$ as a partition of $\mathcal{G}$, aspects of the approaches presented by Lloyd (1982) and Balzer, Schlömer, and Deussen (2009) are the most logical, considering the NP-hard nature of optimal set partitioning problems, to which the construction of capacity-constrained Voronoi tessellations typically belong [26–28]. In truth, the partitioning of a complete, bidirected graph into star graphs is not NP-hard, but some of the considerations must change. Rather than the random selection of locations and repeated arbitrary partitioning of $\Omega \subset \mathbb{R}^n$ suggested by Lloyd (1982), tracking unvisited pairs of locations V_i and V_j with unknown overlap χ_{ij} should provide an improvement in time-to-convergence. For two subgraphs V_i and V_j realized as cells within the tessellation $\mathcal{V}(S)$ of the complete, bidirected graph $\mathcal{G}$, if $\chi_{ij} \geq \chi_{max}$, $\mathcal{V}(S)$ of $\mathcal{G}$ does not converge under the (relaxed) perfect partition condition. Rather than selecting entirely new locations S from vertices Ω that form graph $\mathcal{G}$, locations $s_i, s_j \in S$ that form the bases for subgraphs V_i and V_j such that $\chi_{ij} > \chi_{max}$ should be discarded from the set of pairs of candidate locations.

The overall time complexity of the EPAQS algorithm is $O(bn^3)$, where n is the number of points in the set of vertices Ω of the graph $\mathcal{G}$, and b is the number of sensors to be installed. Algorithm 1 shows the general structure of the EPAQS algorithm in pseudocode, where offending pairs of locations are removed from consideration. This presents a reduction in time needed to optimize the partition when compared to a truly random selection of locations, as unfit pairs are removed. In doing so, EPAQS has no way to suggest the inclusion of locations that are already known to not meet the χ_{max} requirement.

2.3 Validation of the EPAQS Algorithm

Because there exist no comparable methods for sensor placement in the same manner, we have selected to benchmark the performance of the EPAQS algorithm on random selections of locations across the space. This benchmark testing was conducted on the same portion of the underground room-and-pillar coal mine in Utah, USA, for the placement of 5, 10, 20, and 50 sensors. To generate a suitable, random baseline, 10,000 sets of sensor locations were generated, and the maximum pairwise overlap and average pairwise overlap values were averaged for each of the random sets. This allows for the qualitative assessment of the EPAQS algorithm's performance relative to the baseline, but is not suitable for tests of significance because the large volume of samples can impact the meaningfulness of significance tests by obscuring small effect sizes [29, 30].

To test for statistical significance with a sample size that provides an appropriate level of statistical power while minimizing the chances of obscuring effect size, 100 random sets of sensor locations were mathematically tested for 5, 10, 20, and 50 sensors. The average performance statistics for the randomly selected sensor locations and the EPAQS-derived sensor locations were then compared in aggregate to assess if the EPAQS algorithm could be capable of providing a statistically significant improvement over random selection of locations.

Algorithm 1 EPAQS Algorithm Pseudocode.

Import graph $\mathcal{G}$ as a DataFrame containing each pair of vertices and their corresponding directed edge weights
Initialize a hash table with all pairs of vertices in $\mathcal{G}$
Set the number of installation locations
Calculate the capacity of each cell
Initialize a list $generators$ with an arbitrary selection of installation locations from the vertices in $\mathcal{G}$
Initialize maximum allowable overlap χ_{max} as the capacity of each cell minus 1
while the number of keys in the hash table $\neq 0$ **do**
    Initialize a flag for $\chi_{ij} \geq \chi_{max}$
    while the flag is true and the number of keys in the hash table $\neq 0$ **do**
        Initialize an empty list for $\chi_{ij} \forall i, j$
        for each location in $generators$ **do**
            if flag is true **then**
                Break the loop
            end if
            Collect all vertices within the capacity of the cell defined by the installation location
            for each other location in $generators$ **do**
                if the two points are not the same **then**
                    Collect all vertices within the capacity of the cell defined by the installation location
                    Calculate χ_{ij} for the two points
                    Add χ_{ij} to $\chi_{ij} \forall i, j$
                    if $\chi_{ij} \geq \chi_{max}$ **then**
                        Update the first flag for $\chi_{ij} \geq \chi_{max}$
                        Delete the pair of points from the hash table
                        Select a new pair of points from the hash table
                        Update the corresponding points in $generators$" with the new pair
                        Set flag to true and break the loop
                    end if
                end if
            end for
        end for
    end while
    Update flag to false
    Record $\bar{\chi}_{ij}$, χ_{max}, and the current $generators$
    Update χ_{max} as $max(\chi_{ij} \forall i, j)$
    Print final results (optional)
end while
Report optimized list of $generators$, $\bar{\chi}_{max}$, and $max(\chi_{ij} \forall i, j)$

3 Results

The EPAQS algorithm seeks to provide an optimized partition of a space where distances between a pair of points are not necessarily identical when measuring from point A to point B and from point B to point A. This can arise in the case of minimum cost paths, like the modified A* pathfinding algorithm when applied to mine ventilation systems as presented in Brown Requist and Momayez (2024) [23]. The shortest path from one position to another is not necessarily the same length when the direction of travel is reversed. As a genetic algorithm, EPAQS does not require these distances to be the same regardless of measurement direction, unlike previous work focusing on capacity-constrained centroidal Voronoi tessellations as presented in Lloyd (1982) and Balzer, Schlömer, and Deussen (2009) [26, 27].

In a case where the space to partition is a square grid and distances between points are Euclidean, EPAQS can still be used to provide an optimized tessellation. Importantly, this comes with the caveat that the resulting tessellation meets the capacity constraint, but is not necessarily centroidal. The centroidal feature within the EPAQS algorithm is merely approximate in nature, and EPAQS struggles with this constraint in Euclidean space. Further work should focus on modifications to this problem formulation such that the centroidal constraint is better met, especially when considering Euclidean spaces. When considering a Euclidean space with or without a quantizable additive or multiplicative weight field, the methods presented by Lloyd (1982) or Balzer, Schlömer, and Deussen (2009) are better approaches, as up- and downwind relationships are irrelevant to those formulations of the problem. This case is perhaps more intuitive, as the results are more visually clear. The space is strictly Euclidean because any distance measured between a set of two points is identical regardless of the direction of measurement. In partitioning a space of 2025 points, the initial (naive) partition for 5 sensors as shown in Fig. 9 has an average overlap of $\bar{\chi} = 0.100$ and a maximum pairwise overlap of $\chi_{ij} = 0.985$. After 15 iterations, the optimized partition (Fig. 10) returned has an average overlap of $\bar{\chi} = 0.034$ and a maximum pairwise overlap of $\chi_{ij} = 0.077$. Regardless of the method of distance measurement, the EPAQS algorithm provides consistent improvement in the placement of air quality sensors in a space as evidenced by its ability to minimize pairwise overlap. This in turn reduces the average pairwise overlap for all cell pairs in the tessellation. For this Euclidean space, the EPAQS algorithm yields a reduction of average overlap of 66.4% and a reduction of maximum pairwise overlap of 92.2%.

Outside of a strictly Euclidean space, the EPAQS algorithm has been applied to a portion of an underground room and pillar coal mine in Utah, USA, measuring approximately 1200 by 300 ms. A graphical representation of the space is displayed in Fig. 11, where the space has been separated into 5.5 by 5.5-meter regions to reduce the time needed to converge on an optimized partition with higher resolution. The general direction of airflow is indicated via the arrows in the graphic, including intake quantities.

With a 5.5 by 5.5-meter resolution, the study space contains 1577 distinct locations that must be partitioned for a set number of available sensors. Considering the placement of 10 sensors, the initial tessellation yields an average overlap of $\bar{\chi}$ =0.307 and a maximum pairwise overlap of χ_{ij} =0.987. The initial partition is displayed in Fig. 12. The iterative process as outlined in the algorithm pseudocode results in an optimized partition with $\bar{\chi}$ =0.030 and a maximum pairwise overlap of χ_{ij} =0.250 after 18 iterations. This method

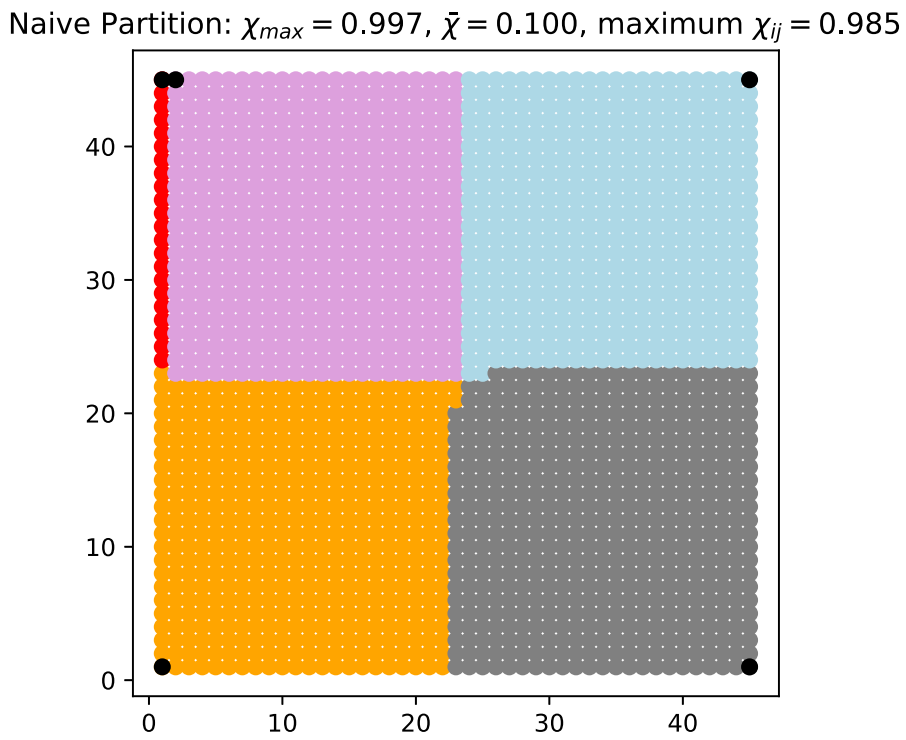

Fig. 9 Naive partition of a Euclidean space with 5 installation locations

of optimization achieves a 90.0% reduction of average pairwise overlap and a 74.7% reduction of maximum pairwise overlap from the initial partition to the optimized partition. For partitions optimized using the EPAQS algorithm within the portion of the mine, the centroidal nature is much more enforced. Sensors are more likely to be placed such that there is more representation of locations downwind of the sensor than upwind (Fig. 13).

When instead considering the placement of five sensors in the space, the initial partition has an average overlap of

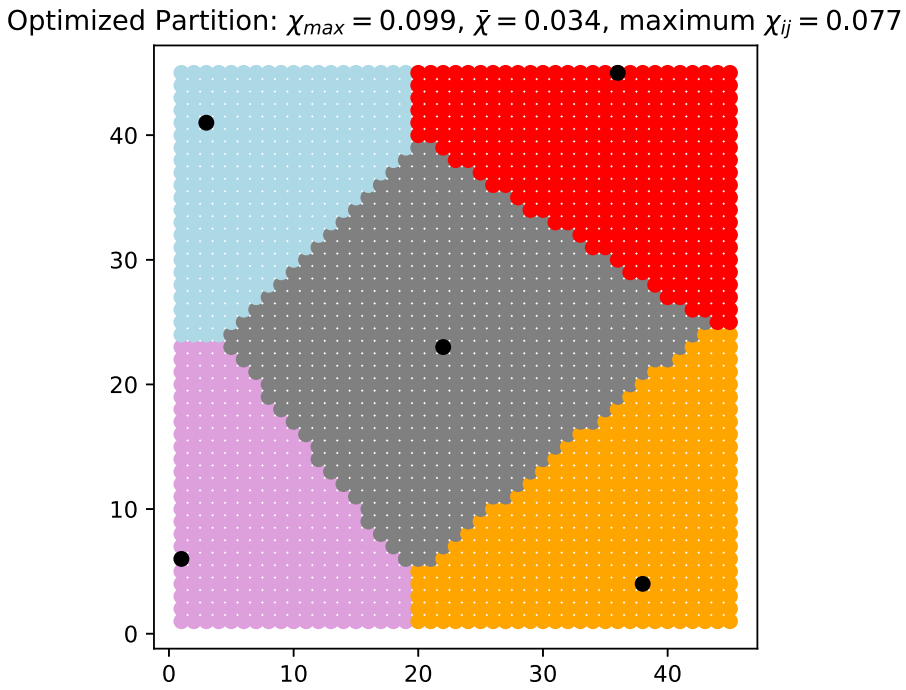

Fig. 10 Optimized partition of the Euclidean space

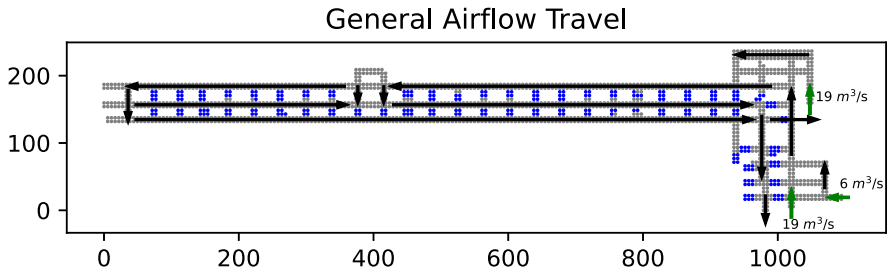

Fig. 11 Graphical representation of a portion of an underground room and pillar coal mine in Utah, USA

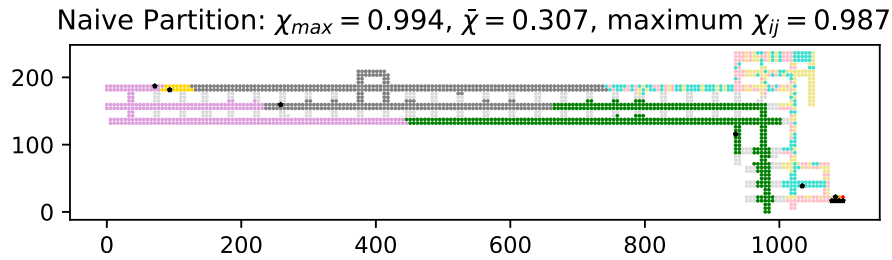

Fig. 12 Naive partition of the underground mine with 10 sensors

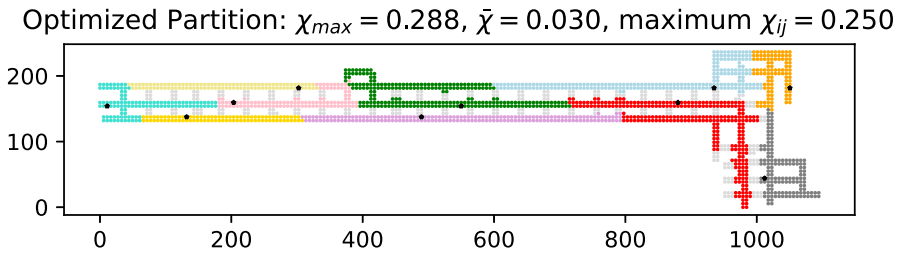

Fig. 13 Optimized partition of the underground mine with 10 sensors

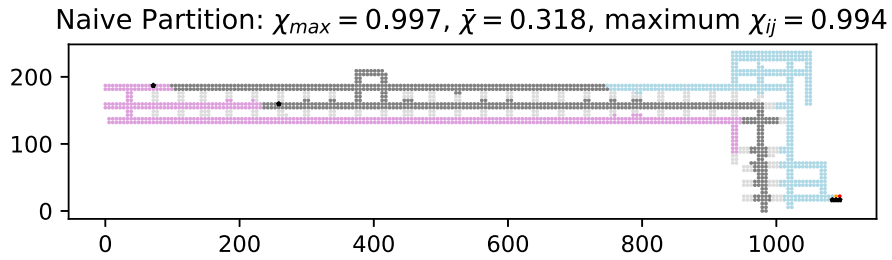

Fig. 14 Naive partition of the underground mine with five sensors

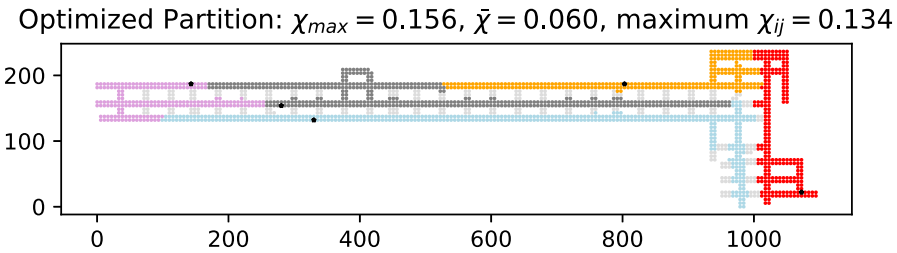

Fig. 15 Optimized partition of the underground mine with five sensors

Table 2 Placement of 5, 10, 20, and 50 sensors using the EPAQS algorithm

Number of sensors	Iterations	$\bar{\chi}_{ij}$	% Reduction $\bar{\chi}_{ij}$	$max(\chi_{ij})$	% Reduction $max(\chi_{ij})$	Time (minutes)
5	16	0.060	81.2	0.134	86.5	23.85
10	18	0.030	90.0	0.250	74.7	26.21
20	14	0.024	83.4	0.429	56.0	39.55
50	14	0.010	85.9	0.500	46.4	145.26

$\bar{\chi}$ =0.318 and a maximum pairwise overlap of χ_{ij} =0.994. After 16 iterations, the optimized partition has an average overlap of $\bar{\chi}$ =0.060 and a maximum pairwise overlap of χ_{ij} =0.134. Like the case of 10 sensors, the EPAQS algorithm returns an optimized solution with a 81.2% reduction in average pairwise overlap and a 86.5% reduction in maximum pairwise overlap (Figs. 14 and 15).

Table 2 summarizes the percent reductions in average pairwise overlap and maximum pairwise overlap alongside the number of iterations required for the placement of 5, 10, 20, and 50 sensors in the portion of the underground mine. As the number of sensors to be placed within the space increases, the minimum pairwise overlap increases. This is a feature of the space that is unavoidable, but the overall goal of providing even coverage with fixed sensors throughout the mine is still met. This is shown in the reductions in average overlap, $\bar{\chi}_{ij}$. As the partitions iteratively improve, $\bar{\chi}_{ij}$ tends to decrease.

Figure 16 shows the reduction in $max(\chi_{ij})$ and $\bar{\chi}_{ij}$ when using the EPAQS algorithm to partition the portion of the underground mine for 10 sensors. Maximum pairwise overlap necessarily decreases with each iteration, and average pairwise overlap tends to decrease as well.

In the case of 50 sensors, the reduction in $max(\chi_{ij})$ (Fig. 17) is perhaps less dramatic, but the decrease in $\bar{\chi}_{ij}$ shows that even this smaller reduction plays a large role in spacing sensors in the space such that there is more even spatial coverage of the space.

When compared to the random baseline, EPAQS significantly outperforms any random selection of sensor locations. With baseline testing generating 10,000 sets of locations and 100 sets of locations for 5, 10, 20, and 50 sensors, effect sizes remained within ±10%, showing the EPAQS algorithm's ability to optimize sensor placement. Table 3 summarizes the results of random baseline testing using 100 randomly generated sets of sensor placements compared to results generated by the EPAQS algorithm. Using the 100 randomly generated sets and as evidenced by the effect sizes in Table 3,

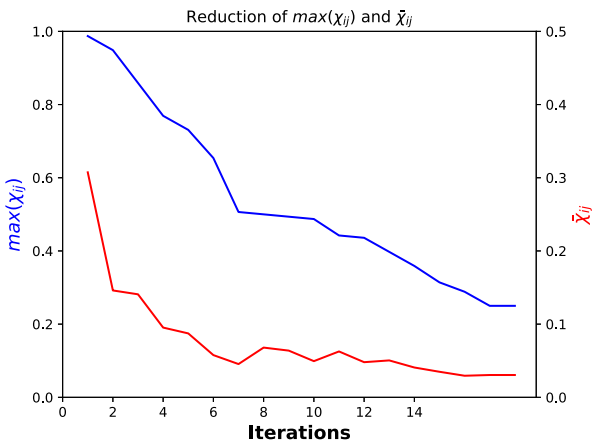

Fig. 16 Iterative reduction of $max(\chi_{ij})$ and $\bar{\chi}_{ij}$ for 10 sensors

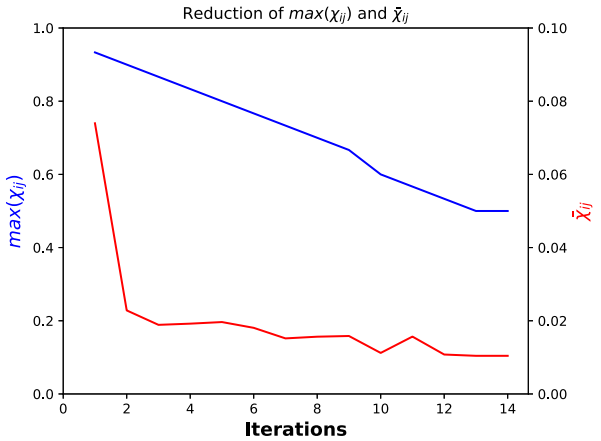

Fig. 17 Iterative reduction of $max(\chi_{ij})$ and $\bar{\chi}_{ij}$ for 50 sensors

EPAQS provides a statistically significant reduction in maximum pairwise and average pairwise overlaps ($p \ll 0.0005$).

With the iterative approach, the EPAQS algorithm is able to provide a suggested placement of air quality sensors for underground mining environments that provide even spatial coverage while respecting airflow. This is vital for any attempts to use spatial statistical methods for the estimation of contamination distributions. In order to proceed with statistical estimation of distributions, input data must be spaced such that it provides an even statistical representation across the estimation domain. When considering methods like geostatistics, even spatial resolution is relatively easy to obtain through declustering methods. In the case of a mine's ventilation system, however, declustering is not a viable option, as data is not nearly as abundant as seen in resource modeling applications. On average, AMSs used in US underground coal mines operate with 38 sensors across the entire mine [2]. When considering the large footprint of a mine, this poses a challenge when looking to expand non-regulatory monitoring of other areas, especially when real-time air quality sensors are still costly. With a constraint on the number of datapoints available, the development of a method to place sensors in the most effective way possible is preferred. Using the EPAQS algorithm, the number of sensors can be predetermined, and the algorithm can return a suggested placement scheme that provides an even statistical representation across the mine's ventilation system without the need for post-hoc data declustering.

4 Discussion

Declustering methods are not practical when considering air quality monitoring in underground mines. Air quality data is much less prevalent than that used as a basis for resource modeling, and the current cost of real-time air quality monitors prevents operations from installing the volume of sensors

Table 3 Baseline testing of the EPAQS algorithm using randomly selected sensor locations

Number of sensors	Random $\bar{\chi}_{ij}$	EPAQS $\bar{\chi}_{ij}$	Random $max(\chi_{ij})$	EPAQS $max(\chi_{ij})$
5	0.847	0.134	0.268	0.060
	Effect size	5.78	Effect size	1.98
10	0.918	0.250	0.131	0.030
	Effect size	7.83	Effect size	2.18
20	0.945	0.429	0.063	0.024
	Effect size	10.09	Effect size	2.20
50	0.946	0.500	0.021	0.010
	Effect size	21.11	Effect size	4.38

necessary to allow for traditional declustering techniques. Instead, care should be taken to place air quality monitors such that each monitor provides an even statistical coverage of the mine's ventilation system. When attempting to space input data evenly, this is relatively straightforward for applications like resource modeling, where a uniform grid can be established or modified as necessary to reflect the anisotropy of the estimation domain. This is not a possibility in an underground mine's ventilation system. It is crucial to consider the impacts of contamination dynamics and the movement of air within the mine, and how these factors can impact providing an even statistical coverage.

Using a modified A* pathfinding algorithm as outlined in Brown Requist and Momayez (2024), statistical relationships between two points in an underground mine can be approximated [23]. This allows for an attempt at even spacing of air quality sensors. Given that a typical grid approach is not reasonable in an underground mining environment, it is beneficial to explore other avenues to provide an even statistical coverage. The use of a capacity-constrained centroidal Voronoi diagram can work to approximate an even distribution of sensors. A capacity-constrained centroidal Voronoi diagram is a promising lens through which to view this problem, as any placement of sensors should see that the local neighborhood of relevant data collection is unique to one sensor (as seen in a Voronoi tessellation) and that the local neighborhood for one sensor is approximately equal in area to the local neighborhood of another sensor (via the capacity constraint).

The EPAQS algorithm provides an even statistical coverage across an underground mine's footprint by partitioning a complete, bidirected graph into multiple star graphs (complete, bipartite graphs of form $K_{1,k}$). This division allows for the approximation of the centroidal nature of a capacity-constrained centroidal Voronoi diagram. By iteratively minimizing the number of vertices shared between any two subgraphs and by requiring that all subgraphs contain the same number of vertices (locations within the underground mine), the EPAQS algorithm provides an even spatial coverage of the mine while considering contamination dynamics and airflow.

The algorithm approximates a solution to the capacity-constrained centroidal Voronoi diagram problem with a time complexity of $O(bn^3)$. This time complexity is not particularly efficient, especially when discretizing the underground mine's footprint into very small regions. Further work is needed to investigate the addition of heuristic methods to approximate solutions with improved time complexity to ensure that partitioning does not require an unreasonable amount of time. Currently, the best option would be to select smaller portions of an entire underground mine's footprint and apply the EPAQS algorithm to each portion. This is not necessarily ideal when attempting to provide an even statistical coverage across the mine's entire footprint, as there is no guarantee that the divided portions are evenly spaced themselves. Alternatively, the mine could be discretized into much larger portions using the EPAQS algorithm, and then the EPAQS algorithm could be iteratively applied to the computed portions to achieve sufficiently small cells. Neither approach is necessarily ideal. Additional work to reduce the time complexity of the algorithm is a promising start, but future attempts of this problem should likewise consider the use of graph representations of the space, as there may be more efficient heuristic methods that do not require the graph representation.

Because the EPAQS algorithm is the first algorithm developed specifically for the placement of air quality sensors that focuses on an even spacing with respect to airflow dynamics in underground mines, it is not practical to compare it to existing spatial optimization algorithms like those developed by Liu et al. (2022), Yan et al. (2023), or Muduli et al. (2017) [12–14]. These algorithms were developed for different purposes, complicating options for algorithm comparison. This is likewise true for existing capacity-constrained Voronoi diagram algorithms like those developed by Lloyd (1982) and Balzer, Schlömer, and Deussen (2009), which are constrained to applications in Euclidean spaces with continuous, integrable, and known weight fields [26, 27]. The same argument can be made for cell declustering, which requires a significantly higher volume of available data (installed sensors) than is reasonable in existing underground air quality monitoring systems. As such, the mathematical performance

of the EPAQS algorithm as presented is the best option to display the general effectiveness of the algorithm. With the overall goal of developing a sensor placement scheme that promotes numerical stability during spatial statistical estimation of contamination distributions in underground mines, more work is needed to compare the prevalence of positive-definite covariance matrices and negative kriging variances during estimation. This will require a comparison of sensors placed according to the EPAQS algorithm and spacing sensors with an equal distance or according to a grid. Preliminary work in this area indicates that sensor placement with EPAQS shows lower frequencies of negative kriging variances during estimation using ordinary kriging, but more analysis is needed across a larger sample size and will require a formalization of the statistical methods used for estimation.

5 Conclusion

The EPAQS algorithm aims to provide a method to optimize the placement of non-regulatory air quality sensors in the absence of federal regulation or industry-backed best practices as operations seek to expand their real-time air quality monitoring capabilities. With the eventual goal of using real-time air quality monitoring data for site-wide estimation of contamination concentration distributions, any statistical approach will require that input data is spaced such that the mine's ventilation system is evenly represented. Data spacing for statistical estimation is a common consideration, especially in resource modeling in the mining industry. Various methods exist, including cell declustering, but these method assume that the input data is abundant enough to support declustering. Data declustering requires grouping of input data, and it is vital to ensure that the volume of input data is sufficient such that declustering does not cause a meaningful change in statistical support for estimation within the domain.

The EPAQS algorithm presents an option for underground mining operations to computationally arrive at the best placement of air quality sensors in areas where regulation for sensor placement is not yet defined. In the absence of regulatory requirements or industry-accepted best practices, the EPAQS algorithm prioritizes placing sensors such that each sensor covers an area equal in size to all other sensors and that the sensors are placed such that there is minimal overlap of areas represented by collected data. As an approximation to the capacity-constrained centroidal Voronoi diagram that uses a graph representation of the mine's ventilation system, the EPAQS algorithm respects the statistical relationships between areas of the mine which is vital to providing an even statistical coverage. The development of this algorithm is part of ongoing work to estimate concentration distributions in underground ventilation systems using spatial statistics. The EPAQS algorithm plays an important role in ensuring that the input data sufficiently describes the behavior of contamination within the mine. This algorithm provides a means to place a set number of sensors throughout an underground mine, alleviating concerns of providing a clear picture of air quality throughout the mine with a number of sensors that is agreeable to an operation. This can minimize concerns of sensor placement, as the algorithm provides a computational means to place the sensors, and operations can tailor placement of sensors to the number of monitors they are intending to install. The EPAQS algorithm stands to ease and guide sensor placement with a computational approach in the absence of regulation or best practices as operations look to expand their non-regulatory air quality monitoring capabilities.

6 Supplemental Information

The most recent version of EPAQS algorithm is available for use at no cost as a Windows program on request from the corresponding author.

Acknowledgements Research reported in this publication was supported by the National Institute for Safety and Health (NIOSH) under award number U60OH012351.

Data Availability Data presented in this publication is available from the corresponding author upon reasonable request.

Declarations

Conflict of Interest The authors declare no competing interests.

References

1. Hopper CP, Zambrana PN, Goebel U, Wollborn J (2021) A brief history of carbon monoxide and its therapeutic origins. Nitric Oxide 111:45–63
2. Rowland J III, Harteis S, Yuan L (2018) A survey of atmospheric monitoring systems in us underground coal mines. Min Eng 70(2):37
3. United States Department of Labor: compliance guide for MSHA's safety standards for the use of Belth entry as an intake air course to ventilate working sections and areas where mechanized mining equipment is being installed or removed. https://arlweb.msha.gov/REGS/COMPLIAN/GUIDES/Belt%20Air%20Compliance%20Guide.pdf Accessed on 28 May 2024
4. Kowalski-Trakofler K, Alexander D, Brnich M, McWilliams L, Podlesny A, Lenart P (2009) Underground coal mining disasters and fatalities-United States, 1900–2006. MMWR 57(51):1379–1383
5. United States 73 Federal Register 251: flame-resistant conveyor belt, fire prevention and detection, use of air from the belt entry; final Rule. https://arlweb.msha.gov/REGS/FEDREG/FINAL/2008finl/E8-30639.pdf Accessed on 28 May 2024
6. Shriwas M, Pritchard C (2020) Ventilation monitoring and control in mines. Mining Metallurgy Explor 37(4):1015–1021

7. Journel AG (1989) Fundamentals of geostatistics in five lessons, vol 8. American Geophysical Union, Washington, DC
8. Cressie N (2015) Statistics for spatial data. John Wiley & Sons, New York
9. Journel AG (1983) Nonparametric estimation of spatial distributions. J Int Assoc Math Geol 15:445–468
10. Deutsch C (1989) DECLUS: a FORTRAN 77 program for determining optimum spatial declustering weights. Comput Geosci 15(3):325–332
11. Bourgault G (1997) Spatial declustering weights. Math Geology 29:277–290
12. Liu Y, Liu Z, Gao K, Huang Y, Zhu C (2022) Efficient graphical algorithm of sensor distribution and air volume reconstruction for a smart mine ventilation network. Sensors 22(6):2096
13. Yan Z, Wang Y, Fan J, Huang Y, Zhong Y (2023) An efficient method for optimizing sensors' layout for accurate measurement of underground ventilation networks. IEEE Access
14. Muduli L, Jana PK, Mishra DP (2017) A novel wireless sensor network deployment scheme for environmental monitoring in longwall coal mines. Process Saf Environ Prot 109:564–576
15. Li Y, He B, Wang Y (2021) Optimal sensor placement for underground tunnel monitoring via wireless sensor networks. Wirel Commun Mob Comput 2021(1):6621987
16. Jiang H, Qian J, Peng W (2009) Energy efficient sensor placement for tunnel wireless sensor network in underground mine. In: 2009 2nd international conference on Power Electronics and Intelligent Transportation System (PEITS), vol. 2. IEEE, pp 219–222
17. Ikeda H, Kolade O, Mahboob MA, Cawood FT, Kawamura Y (2021) Communication of sensor data in underground mining environments: an evaluation of wireless signal quality over distance. Mining 1(2):211–223
18. Li M, Liu Y (2009) Underground coal mine monitoring with wireless sensor networks. ACM Trans Sensor Netw 5(2):1–29
19. Chen G-Z, Zhu Z-C, Zhou G-B, Shen C-F, Sun Y-J (2008) Sensor deployment strategy for chain-type wireless underground mine sensor network. J China Univ Min Technol 18(4):561–566
20. Harb C, Rajapaksha R, Moya X, Roberts J, Hemp P, Uecker L, Rawson T, Roghanchi P (2021) Development of non-regulatory runtime respirable coal and silica dust monitor. In: Mine Ventilation, pp. 242–247. CRC Press, London
21. Medina A, Vanegas A, Madureira E, Roghanchi P, Rajapaksha R, Uecker L, Rawson T, Harb C (2023) Comparison of respirable coal and silica dust monitoring systems for underground mining applications. Underground Ventilation. CRC Press, London, pp 305–312
22. Rahimi E, Shekarian Y, Shekarian N, Roghanchi P (2023) Investigation of respirable coal mine dust (RCMD) and respirable crystalline silica (RCS) in the us underground and surface coal mines. Sci Rep 13(1):1767
23. Brown Requist K, Momayez M (2024) Minimum cost pathfinding algorithm for the determination of optimal paths under airflow constraints. Mining 4(2):429–446
24. Aurenhammer F (1991) Voronoi diagrams—a survey of a fundamental geometric data structure. ACM Comp Survey 23(3):345–405
25. Du Q, Faber V, Gunzburger M (1999) Centroidal Voronoi tessellations: applications and algorithms. SIAM Rev 41(4):637–676
26. Balzer M, Schlömer T, Deussen O (2009) Capacity-constrained point distributions: a variant of Lloyd's method. ACM Trans Graph 28(3):1–8
27. Lloyd S (1982) Least squares quantization in PCM. IEEE Trans Inf Theory 28(2):129–137
28. Yang K, Shekhar AH, Oliver D, Shekhar S (2015) Capacity-constrained network-Voronoi diagram. IEEE Trans Knowl Data Eng 27(11):2919–2932
29. Lantz B (2013) The large sample size fallacy. Scand J Caring Sci 27(2):487–492
30. Kühberger A, Fritz A, Lermer E, Scherndl T (2015) The significance fallacy in inferential statistics. BMC Res Notes 8:1–9