Reducing human health impacts from power sector emissions with redispatch and energy storage

Qian Luo, Jeremiah X Johnson and Fernando Garcia-Menendez
North Carolina State University, Department of Civil, Construction, and Environmental Engineering, Raleigh, NC 27607, United States of America
* Author to whom any correspondence should be addressed.
E-mail: f.garcia@ncsu.edu

Keywords: power system, health impacts, energy storage, air pollution

Abstract
Emissions from the power sector significantly contribute to ambient air pollution and its associated adverse human health impacts. In this study, we explore how to cost-effectively reduce health impacts due to fine particulate matter (PM$_{2.5}$) attributable to power plant emissions by internalizing real-time health costs in plant dispatch decisions and re-optimizing the unit commitment and economic dispatch in light of these impacts. We show that internalizing the time- and location-varying health damage costs into power system operational decisions can reduce 61%–97% of adverse health impacts through decreases in coal generation and strategic shifts in the location and timing of pollutant releases. We also find that adding energy storage to the grid can mitigate health impacts by reducing wind power curtailment. Our findings demonstrate the need to consider temporal and spatial heterogeneity when determining the social cost of emissions.

1. Introduction
Air pollution is the fifth largest risk factor for global mortality, resulting in nearly 5 million deaths and 150 million years of healthy life lost annually [1]. The negative health effects associated with major indicators of air pollution, PM$_{2.5}$ and ground-level ozone, include cardiovascular disease, chronic respiratory disease, and respiratory infections [2, 3]. Ambient PM$_{2.5}$ exposure is estimated to have contributed to nearly 3 million premature deaths in 2017 [1]. In addition, PM$_{2.5}$ exposure has been associated with an average life expectancy reduction of 1 year worldwide [4]. Despite lower levels of exposure compared to developing countries, the United States still experiences significant adverse health impacts from PM$_{2.5}$, causing more than 85 000 deaths annually and making it the environmental risk factor with the largest burden of disease [5].

Emissions from the power sector are among the largest contributors to ambient PM$_{2.5}$ [6]. More than 20 000 annual premature deaths in the US have been estimated to be caused by power plant emissions, driven by secondary PM$_{2.5}$ [7]. As an important source of sulfur dioxide (SO$_2$) and nitrogen oxides (NO$_x$), power plants account for over 65% of ambient PM$_{2.5}$ sulfate and 30% of ambient PM$_{2.5}$ nitrate in Eastern US [8]. Thus, reducing emissions from the production of electricity, predominately due to the combustion of fossil fuels, would decrease secondary PM$_{2.5}$ formation and its associated health impacts, which is critical to environmental sustainability. The formation of secondary PM$_{2.5}$, however, is determined by atmospheric conditions, which differ greatly by time and location. Variations in temperature, wind, precipitation, absolute humidity, and mixing height influence PM$_{2.5}$ formation, transport, and, ultimately, ground-level concentrations [9, 10]. Different emission rates of SO$_2$ and NO$_x$ and power plant locations make the health impacts associated with their electricity generation vary greatly [11, 12]. Therefore, strategies to mitigate the adverse health impacts must consider atmospheric processes and power plant characteristics.

In this study, we develop a novel approach to demonstrate the potential to shift the time and location of these emissions through redispatch of existing power plants or deploying grid-scale energy storage in a manner that decreases PM$_{2.5}$ concentration and provides significant human health benefits. Several previous studies
have integrated air quality and power system models. Kerl et al demonstrated the potential to reduce health impacts by applying the community multiscale air quality (CMAQ) modeling system with the decoupled direct method (DDM) to a power system operational model [13]. If scaled-up to a regional level to match power system operations, however, the approach would be limited by the computational cost required to analyze the emissions from hundreds of electric generating units (EGUs) in a day-ahead power market. Other studies have taken advantage of reduced-complexity air quality models (RCMs) such as the Estimating Air pollution Social Impact Using Regression (EASIUR) model [14] and the Intervention Model for Air Pollution (InMAP) [15] to quantify the health impacts caused by a power system, but these efforts lack variation at fine temporal resolution due to the inherent limitations of reduced-form models. Through the development of the Air Pollution Emission Experiments and Policy (APEEP) model, Muller and Mendelsohn linked point source air emissions to damages, with the most significant costs resulting from human health impacts [16]. However, APEEP does not capture the temporal variation of air pollution formation and transport. McDonald-Buller et al suggested dynamic management of NOx and SO2 emissions but the emission prices are only differentiated for high ozone days [17]. Couzo et al proved that dynamic management of electricity generation was cost-effective in ozone reductions using three sensitivity techniques (brute force, DDM, and higher-order DDM), though this work solely focuses on a single high ozone episode [18].

Our study is the first to consider real-time health impacts caused by electricity generation at individual EGU level for an entire independent system operator. We reformulate a unit commitment and economic dispatch model with the addition of energy storage, with the goal of cost-effectively reducing these harms. We demonstrate the value of this approach on a realistic system: the independent system operator in Texas (ERCOT), which has a mix of coal, natural gas, and nuclear power plants and variable renewables (in particular, wind power), while providing suitable data availability and limited transmission interconnection to other regions. In addition, Texas has been a leading state in SO2 and NOx emissions from power sector for years [19]. We first use a dispersion model with simplified chemistry to quantify the marginal health impacts that would be attributable to secondary PM2.5 formed by SO2 and NOx emissions from every EGU in the power system the day before determining their operation. We then monetize these impacts based on an epidemiologically derived health impact function and the US Environmental Protection Agency’s recommended value of a statistical life (VSL), and internalize them into power plant dispatch decisions as part of the variable cost of electricity generation. Next, we optimize the power system operation with these costs internalized via the unit commitment and economic dispatch model in a day-ahead market to determine plant commitment and operation. In the presence of high spatial and temporal variation in the damages associated with power plant emissions, our approach identifies prime opportunities to reduce adverse human health impacts at a modest cost. By internalizing the adverse human health impacts and minimizing operational costs, if deployed globally, this approach could help meet the sustainable development goal of ensuring access to affordable and clean energy [20].

2. Methods

2.1. Overview and data collection

Our power system model represents all individual EGUs with a capacity greater than 25 MW, which results in 185 natural gas units, 32 coal units, and 4 nuclear units across the ERCOT region. Characteristics of each fossil-fueled EGU needed in this model include fuel types, fuel prices, capacity, heat rates, stack information, and emission rates data. Fuel type, capacity, and stack information are obtained from the US Energy Information Administration (EIA) Form-860 [21]. Heat rates and emission data are determined using data obtained from the Air Markets Program Data (AMPD) [22]. The emission rate (g MWh–1) is calculated for each generator based on the hourly emission data collected by AMPD. The location and capacity of wind turbines are from US wind turbine database [23] and hourly wind generation is from ERCOT [24]. The power system operational costs (except for health damage costs described below) include fossil fuel costs calculated from EIA Form-923 [25], variable operational and maintenance (O & M) costs [26], and start-up costs [27]. We deploy a zonal approach to capture the impacts of transmission constraints, relying on data from the US National Renewable Energy Laboratory (NREL) Regional Energy Deployment System Model (ReEDS) to define seven balancing zones (BZs) and their associated transmission line limits [28] (supplementary figure 1 and table 1 (https://stacks.iop.org/ERIS/1/025009/mmedia)). The system demand is acquired from ERCOT [29] and adjusted to conform to the BZ geographic boundaries by scaling based on the population. We run the Weather Research and Forecasting model (WRF4.1) to obtain the meteorological input required by the air quality model, CALPUFF. The WRF domain (lat: 25°–41°N and long: 90°–110°W) covers Texas and five other states neighboring Texas, with a horizontal resolution of 12 km. The census tract-level population and mortality rate are from the Center for Disease Control and Prevention Wide-ranging ONLine Data for Epidemiologic Research (CDC WONDER) dataset [30].
2.2. Marginal health impacts estimation

We estimate hourly marginal health damage costs for 217 fossil fuel-fired EGUs across the ERCOT region. All the EGUs modeled are located in Texas and each generator is assigned its own domain (approximately 800 km by 800 km, with the generator at the center) with a resolution of 3 km by 3 km to determine the health impacts caused by that generator (supplementary figure 1). A sensitivity analysis on the domain size indicates that the domain can capture most downwind exposure (supplementary table 2). CALPUFF (version 7), an advanced, integrated Lagrangian puff modeling system for the simulation of atmospheric pollution dispersion [31], is used to simulate the hourly change in ground-level secondary PM$_{2.5}$ concentration associated with sulfate and nitrate particles exclusively attributable to emissions (SO$_2$ and NO$_x$) from the power system. To include chemical transformations, the updated RIVAD scheme is selected in CALPUFF. CALPUFF uses hourly emission rates (g s$^{-1}$) calculated from the EGU emission rate (g MWh$^{-1}$) and the associated hourly electricity generation as an input to estimate the increase in secondary PM$_{2.5}$ after the release of pollutants. A monthly NH$_3$ background concentration of 2.5 ppb is used in CALPUFF [32].

After determining the resultant changes in ambient PM$_{2.5}$, we use a log-linear concentration-response function with a relative risk of 1.06 for all cause mortality [33] and an assumed value of a statistical life [34] to monetize the associated health impacts. Premature mortality has been found to account for the vast majority of the health damages associated with ambient PM$_{2.5}$ [35–38]. The concentration-response function and value of relative risk are derived from analyses based in the United States and have been used extensively in studies in quantifying the health impacts of PM$_{2.5}$ [39–41]. Health impacts are quantified for emissions from each EGU, emitted in each hour. To do this, CALPUFF is run using each hour’s SO$_2$ and NO$_x$ emissions as the input, and secondary PM$_{2.5}$ concentration is simulated for the following 48 h, which captures most of the air pollution increase (supplementary figure 2). Then, 48 h impacts are estimated by applying the concentration-response function and summing the impacts of each hour’s emissions across this time horizon. The two-day window is necessary to capture the time required for SO$_2$ and NO$_x$ to form particles and reach to the ground. Equations (1)–(3) are constructed to estimate the hourly marginal health damage costs ($/MW h$) for each EGU ($H_{j,h}$).

For each EGU, emissions (g s$^{-1}$) for each hour are determined by the hourly electricity generation. Thus, we run CALPUFF and calculate health impacts several times for every EGU’s one-hour emissions using various emission rates corresponding to different electricity generations within the EGU’s operational range. Then, a linear regression (with monetized health damage, $y$, as the dependent variable and electricity generation, $x$, as the independent variable) is performed to calculate the real-time marginal health damage cost ($H_{j,h}$ in $$/MW h$$) for that specific EGU $j$ at that specific hour $h$. To evaluate heterogeneities in a year, we examine 12 days, namely the first Monday of each month in 2015. Supplementary figure 5 shows that the 12 days selected can represent the range of annual variability in meteorology.

\[
\triangle y_{j,m,h} = y_0 \times \left(1 - \exp \left(-\frac{\ln RR}{10} \times \triangle c_{j,m,h}\right)\right) \tag{1}
\]

\[
\triangle HD_{j,h} = VSL \times \sum_{m=1}^{m=48} \triangle y_{j,m,h} \tag{2}
\]

\[
\triangle HD_{j,h} = f(z_{j,h}) = H_{j,h} \times z_{j,h} + c_{j,h}. \tag{3}
\]

Where,

- $\triangle y_{j,m,h}$: hourly mortality risk change due to the change in PM$_{2.5}$ concentration $m$ hours after SO$_2$ and NO$_x$ are emitted from EGU $j$ in hour $h$ (deaths)
- $\triangle c_{j,m,h}$: hourly average PM$_{2.5}$ concentration attributable to the power system emissions $m$ hours after SO$_2$ and NO$_x$ are emitted from EGU $j$ in hour $h$ ($\mu g/m^3$)
- $\triangle HD_{j,h}$: monetized total health damages due to SO$_2$ and NO$_x$ emissions from EGU $j$ in hour $h$ ($$/ hour$
- $y_0$: hourly baseline all-cause mortality incidence rate (deaths)
- $RR$: relative risk of all-cause mortality for a 10 $\mu g$ m$^{-3}$ increase in PM$_{2.5}$ concentrations, $RR = 1.06$
- $VSL$: value of a statistical life, $9.0$ million (2015 USD)
- $H_{j,h}$: marginal health damage costs of EGU $j$ at hour $h$ ($$/MW h$$)
- $z_{j,h}$: electricity generated by EGU $j$ at hour $h$ (MW h)
- $c_{j,h}$: intercept value of the regression function for EGU $j$ at hour $h$ ($$/hour$$)
2.3. Unit commitment and economic dispatch model

We develop a standard unit commitment and economic dispatch (UC/ED) model to determine the commitment of EGUs and their hourly dispatch. It is formulated and solved as a mixed-integer linear programming (MILP) problem.

Indices used in formulations:

- \( J \): set of all EGUs in ERCOT
- \( H \): set of hours
- \( K/I \): set of balancing zones (BZs) in ERCOT

Parameters:

- \( F_j \): fuel cost for unit \( j \) ($/MW h)
- \( R_j \): variable O & M cost for unit \( j \) ($/MW h)
- \( H_{j,h} \): marginal health damage costs of EGU \( j \) at hour \( h \) ($/MW h)
- \( T_j \): start-up costs of EGU \( j \) ($/start)
- \( s_{k,h} \): electricity from the energy storage located at BZ \( k \) at hour \( h \) (MW h)
- \( d_{k,h} \): energy demand in BZ \( k \) at hour \( h \) (MW h)
- \( SOC_{k,h} \): state of charge of the energy storage located at BZ \( k \) at hour \( h \) (MW h)
- \( ch_{k,h} \): amount of electricity charged by the energy storage located at BZ \( k \) at hour \( h \) (MW h)
- \( dis_{k,h} \): amount of electricity discharged from the energy storage located at BZ \( k \) at hour \( h \) (MW h)
- \( \eta_{ch/dis} \): charging and discharging efficiency of the energy storage (%)
- \( \mu_j \): ramping rates of EGU \( j \) (% of nameplate capacity)
- \( N_j \): nameplate capacity of EGU \( j \) (MW)
- \( DT_k \): duration of the energy storage located at BZ \( k \) (hours)
- \( E_k \): energy storage capacity located at BZ \( k \) (MW h)
- \( \tau_j \): minimum capacity requirement for EGU \( j \) (% of nameplate capacity)
- \( Pi_{i,k,h} \): energy flow from BZ \( i \) to BZ \( k \) at hour \( h \) (MW h)
- \( TC_{i,k} \): capacity of transmission line from BZ \( i \) to BZ \( k \) (MW)

Decision variables:

- \( z_{j,h} \): electricity generated by EGU \( j \) at hour \( h \) (MW h)
- \( v_{j,h} \): binary startup variable for EGU \( j \) in hour \( h \) (1 or 0)
- \( u_{j,h} \): binary unit commitment variable for EGU \( j \) in hour \( h \) (1 or 0)
- \( w_{j,h} \): binary shutdown variable for EGU \( j \) in hour \( h \) (1 or 0)

Because health damage costs are treated as linear to the electricity generation, they can be internalized into the UC/ED model as part of the variable system costs. Therefore, the objective of this model is to minimize total costs, inclusive of health damage costs and electricity generation costs (equation (4)). When not considering health impacts, \( H_{j,h} \) is removed from the objective function and the model will only minimize the electricity generation costs. The difference in the unit commitment and generator dispatch leads to differences in emissions and health impacts.

\[
\min \sum_{h=1}^{H} \sum_{j=1}^{J} [(F_j + R_j + H_{j,h}) \times z_{j,h} + T_j \times v_{j,h}].
\] (4)

The model includes most essential real-world constraints in a power system: ramping limitations, operational range, and minimum uptime or downtime are set for all EGUs in the system.

Unit commitment constraints

All variables in the unit commitment constraints are binary variables. The relationships for the startup, shutdown, and unit commitment binary variables can be defined and linearized as the following equations (equations (5)–(11)) [13]. We assume it takes 4 h for coal-fired EGUs to be turned on or off [42], which means that a coal-fired EGU can operate functionally 4 h after it is turned on and it cannot be turned on again within 4 h from the moment it is turned off. However, there are no such limitations for natural gas-fired EGUs. For natural gas EGUs:

\[
v_{j,h} - w_{j,h} = u_{j,h} - u_{j,h-1}
\] (5)

\[
v_{j,h} \leq u_{j,h}
\] (6)
For coal-fired EGUs:

\[ v_{j,h-4} - w_{j,h} = u_{j,h} - u_{j,h-1} \]  
\[ v_{j,h-4} \leq u_{j,h} \]  
\[ w_{j,h-4} \leq 1 - u_{j,h} \]  

During coal-fired EGUs’ startup process:

\[ v_{j,h-m} \leq 1 - u_{j,h} \quad (m = 0, 1, 2, 3). \]  

During coal-fired EGUs’ shutdown process:

\[ w_{j,h-m} \leq 1 - u_{j,h} \quad (m = 0, 1, 2, 3). \]  

**Balance between load and generation**

The fundamental constraint in a UC/ED model is the balance between the supply and demand in the grid. In BZ\(_k\), the energy generation from EGUs located at BZ\(_k\), energy consumption, and energy flow have to be balanced as equation (13). EGU\(_j\) is located at BZ\(_k\).

\[ s_{k,h} + \sum_j u_{j,h} \times z_{j,h} + \sum_{i \neq k} f_{i,k,h} - \sum_{i \neq k} f_{i,k,h} = d_{k,h} \]  

**Constraints on energy storage**

Electricity energy loss occurs during the charging and discharging processes for energy storage and a round-trip efficiency of 90% can be achieved for lithium-ion batteries [43]. We assume that the one-way charging or discharging efficiency of the energy storage is 95%. In addition, the state of charge for batteries needs to be maintained within the allowable range. We assume that the charge and discharge duration of the energy storage is 4 h at its rated capacity. To explore the impacts of using energy storage, we assume one energy storage device in each BZ.

\[ \text{SOC}_{k,h} = \text{SOC}_{k,h-1} + \text{ch}_{k,h} \times \eta_{ch} - \frac{\text{dis}_{k,h}}{\eta_{dis}} \]  
\[ \frac{\text{ch}}{\text{dis}_{k,h}} \leq \frac{E_k}{DT_k} \times 1 \text{ h}. \]  

**Ramping constraints**

Due to the flexibility of natural gas turbines within our one-hour time interval, we assume ramping constraints are only binding for coal-fired EGUs. For coal units, we assume a ramping rate of up to 33% per hour of their rated capacity [44]. This means that the maximum increase or reduction in output per hour for a 100 MW coal-fired EGU is 33 MW.

\[ z_{j,h} - z_{j,h-1} \leq \mu_j \times N_j \]  
\[ z_{j,h-1} - z_{j,h} \leq \mu_j \times N_j. \]  

**Operational range constraints**

Each EGU cannot exceed its available capacity. In addition, once committed, fossil-fueled power plants must maintain a minimum generation level. Here, an operational range of 40%–100% of the capacity is used for coal-fired EGUs, and a range of 25%–100% is used for natural gas-fired EGUs, consistent with ranges used in other studies [42, 45].

\[ z_{j,h} \leq u_{j,h} \times N_j \times 1 \text{ h} \]  
\[ z_{j,h} \geq \tau_j \times u_{j,h} \times N_j \times 1 \text{ h} \]  

**Minimum uptime and downtime constraints**

We assume that coal-fired EGUs also have to meet the minimum uptime and downtime requirements. Once committed, these EGUs must generate for at least 8 h [46]. In summary, here coal-fired EGUs need 4 h of preparation to start, at least eight working hours once committed and 4 h to power off. Therefore, coal-fired units need a minimum of 16 h to complete a cycle.

\[ v_{j,h-m} \leq u_{j,h} \quad (m = 5, 6, 7, 8, 9, 10, 11). \]
Transmission constraints

We construct the transmission network and set the capacity limitation between BAs based on the approach deployed in NREL’s ReEDS model [28]. Considering well-developed networks in each BZ, we assume there are no transmission limitations within them.

\[
f_{i,k,b} \leq TC_{i,k}. \tag{21}
\]

3. Results

3.1. Health damage costs from coal and natural gas generation

The health damage cost varies greatly across generators and the time of generation. Figure 1(a) shows the diurnal pattern in the distribution of health damages at the unit level for coal- and natural gas-fired EGUs. For many coal generators, the health damage cost far exceeds their operational cost. Despite the variability caused by meteorology and the location of population centers, in general, these damage costs are higher during the day and lower at night because the formation of PM$_{2.5}$ is affected by atmospheric variables, such as sunlight and mixing layer height. The primary oxidizer of SO$_2$ and NO$_x$, hydroxyl radicals, is produced from the photolysis of ozone and thus emissions released at night do not lead to formation of PM immediately. Additionally, the mixing height is higher during the day (figure 1(b)), creating a larger available volume for the dispersion of pollutants [47]. The emergence of sunlight and low mixing layer height contributes to the increase in damage costs during the early morning. Higher mixing layer height and the subsequent sunset drive the decrease later in the day.

We do not observe an evident seasonal pattern in the distribution of marginal health damages based on our selected days, which include one day for each month, but damage costs vary significantly across these days (figures 1(b) and (c)). We compare the damage costs to the meteorological conditions on day 1, day 3, and day 7 which have medium, high, and low health damage costs, respectively. We observe that the average mixing height over the domain is highest on day 7 and lowest on day 3, running counter to the damage costs on those two days (figure 1(c)). Most smokestacks remain below the mixing height and pollutants trapped in a thinner layer lead to higher surface concentration. In addition to the mixing layer, low wind speeds may also contribute to the high health damage costs on day 3. Wind speed below the mixing layer top is low (figure 1(d)) and its direction changes frequently on day 3 (supplementary figure 4), making PM$_{2.5}$ stagnate in regions with high
population for longer, resulting in higher health damage costs. We expect larger marginal health damage costs when mixing height and wind speed are lower, and when wind direction changes frequently.

With high SO\textsubscript{2} emission rates, coal-fired EGUs have higher damage costs than natural gas-fired EGUs. In extreme cases, the damages from coal-fired generators can reach $1000 per MW h, becoming the dominant contributor to generation cost when these externalities are considered. EGUs with the largest health damage costs, however, are not always those with the highest emission rates; often, EGUs with the largest damage costs are those in close proximity to large population centers in Texas (i.e., Austin, Dallas, Houston, and San Antonio). For example, supplementary figure 3 compares the marginal health damage costs of two EGUs with similar emission rates. The costs of one EGU are much higher because of its proximity to a population center (Austin) whereas the other is located in a more remote area with low population density. Despite having considerably lower emissions rates, the health impacts of natural gas-fired EGUs are not insignificant at all hours. For example, the marginal health damage cost of gas-fired EGUs in Decker Creek (shown in supplementary figure 3) is greater than $100 per MW h on day 12, making this natural gas facility a more expensive choice when taking health impacts into account.

Perhaps counterintuitively, we observe EGUs with negative marginal health damage costs. This can be driven by the exit temperature and velocity of emissions from the stack. Based on the data collected by continuous emission monitoring systems (CEMS), increasing generation usually leads to higher exit temperatures and velocities, which can in turn lead to higher plume rise. At higher altitudes, emitted pollutants can travel and disperse further in the atmosphere before reaching the ground, resulting in the potential for lower health impacts.

### 3.2. Mitigating health impacts through generator redispatch

In table 1, we present the resulting system-wide changes in emissions, health impacts, operational costs, and the benefit/cost ratio after internalizing health damages. By internalizing the marginal health damage cost to power system operation, we reduce several 100 tons of emissions per day across the system, decreasing health impacts by millions of dollars, mainly driven by the reduction of SO\textsubscript{2} emissions. On most of the selected days, the internalization of health damage costs reduces these adverse impacts by over 85%, while increasing the operational costs by less than 5%. Across the 12 days examined, the decrease in health damages exceeds the increase in operational costs by a factor greater than 20. Although the operational costs increase by over 8% on day 3, we find the highest benefit/cost ratio on that day due to high avoided health impacts. Lower reductions on day 7 and day 12 are related to the originally low health impacts prior to the internalization of damages.

These health benefits are achieved by increasing electricity generation at slightly more expensive but much cleaner EGUs based on a dynamic calculation of real-time health impacts. As shown in figure 2(a), electricity generation from most coal-fired EGUs is reduced, while selected natural gas plants increase their output. This results in a net reduction of SO\textsubscript{2} emissions (figure 2(b)) and, to a lesser extent NO\textsubscript{X} emissions (supplementary figure 8), yielding system-wide reductions in adverse health impacts (figure 2(c)). Due to high SO\textsubscript{2} emissions, the percentage of electricity generated by coal is reduced by more than 60% on average (33%–89%) on the 12 days examined after the health damage internalization (supplementary figure 7). The only coal-fired power plant that increases its generation has a very low emissions rate relative to other coal-fired power plants and low population density downwind (figure 3).

While we find that emission reductions are consistently followed by a decrease in health impacts (table 1), it is not a simple linear relationship between them because of the variability in marginal health damage costs.
Figure 2. Spatial distribution of changes at the unit level after internalizing health damage costs. On day 3, (a) shows the electricity generation change, (b) shows SO$_2$ emission change, (c) shows monetized health impacts change, and (d) shows health impacts monetized by generation. The size of symbols represents electricity generation from each EGU before the internalization.

Figure 3. Comparison of PM$_{2.5}$ impacts for two coal-fired EGUs. Cumulative PM$_{2.5}$ concentration increase for 48 h after release caused by SO$_2$ and NO$_x$ emitted during day 3. Electricity generation of the EGU on the left increases and that of the EGU on the right decreases after internalizing health damage costs.

Figure 2(d) also shows how the health impacts normalized by generation of each EGU (daily average health damage costs, $/MW h$) vary significantly within ERCOT and EGUs with the greatest emission reductions do not necessarily yield the greatest reduction in health impacts. This finding undermines the use of a static value for the social cost of emissions ($/ton$ or $/MW h$) to quantify impacts associated with the power sector and demonstrates the importance of considering both temporal and spatial heterogeneity in the public health cost of emissions.

When we internalize health damage costs, we find that several EGUs with comparatively high emissions but low generating costs are turned off on most of the days examined. This suggests that, if we were to deploy this method in power system operations, these EGUs would likely be retired. The approach can prioritize the retirement of coal generation based on health impacts criteria.

While we only internalize the health damages from SO$_2$ and NO$_x$ emissions in our optimization, the redispatch of ERCOT power plants also yields significant reductions in greenhouse gas emissions. The shift in
electricity generation from coal to natural gas reduces daily average CO₂ emissions by 9% (3%–14%, supplementary table 5). Assuming a social cost of CO₂ at $41/ton for year 2015 (2015 USD, with a discount rate at 3%) [48], these benefits would total $2 million/day ($1–$3 million/day), equal to 14% (3%–105%) of the health benefits that realized through the reduction of SO₂ and NOₓ, although we do not include them in our model. Further, the social cost of CO₂ may be much higher when including the full economic impacts of climate change [49].

3.3. Mitigating health impacts with energy storage

As an emerging technology capable of increasing grid reliability and integrating more renewable energy, energy storage can shift both the time and location of power sector emissions based on its charging and discharging strategies [50]. Here, we add energy storage to the grid, still internalizing health damages and minimizing the system’s overall operational costs. This formulation allows for both the redispatch of generators and the strategic charging/discharging of energy storage to reduce adverse health impacts.

With the establishment of competitive renewable energy zones (CEREZ) that connect areas with abundant wind resources to highly populated regions of Texas, total annual wind generation in the state nearly doubled from 2015 to 2019 (an 88% increase) [51]. To explore the impact of renewable energy on the ability of energy storage to reduce health damages, we consider three wind penetration scenarios. The basic wind penetration scenario considers the penetration level in 2015. The second scenario considers the wind penetration level in 2019 (1.88 × 2015 wind). The third scenario considered a penetration level 50% higher than that in 2019 (1.5 × 2019 wind), following projections of installed wind capacity in 2025 [52]. We incorporate energy storage with four-hour duration and various capacities (100 MW/400 MW h, 250 MW/1000 MW h, 500 MW/2000 MW h, and 1250 MW/5000 MW h) in each of the seven balancing zones in the unit commitment and dispatch model, yielding a total of 700–8750 MW of storage across ERCOT, to represent a wide range of potential future investments.

As shown in table 2, energy storage can be deployed to reduce operational costs and health damages by providing more flexibility to the grid. However, the overall effect on the system’s operational costs is small, with reductions limited to 1.3% assuming the 2015 wind generation, despite the introduction of very large energy storage systems. This is due, in part, to current grid flexibility to redispatch electricity generation without significant additional cost and low natural gas prices relative to coal, which yield a small range of operational costs across Texas power plants and limit storage-enabled arbitrage opportunities. In addition, we observe modest potential to reduce health damages (6.5% under 2015 wind) using energy storage. Because the optimization is formulated to minimize total system costs, the redispatch of coal and natural gas plants greatly reduces health damages, limiting the opportunity for further reductions from energy storage. Assuming the average across the 12 days selected is representative of a year, the health impacts reduced by the addition of energy storage will exceed $20 million annually under the 2015 wind scenario if a total storage capacity of 8750 MW is installed in ERCOT. In addition, energy storage cannot use a large fraction of the wind or solar energy lost due to limited renewable curtailment in the system under the 2015 wind scenario where wind generation only accounts for 11% of total electricity generation. However, with higher wind penetration, energy storage can reduce a greater share of the health damages and operational costs, and reduces 5 GW h of curtailment per day under the 1.5 × 2019 wind scenario, where 27% of total generation is from wind. Due to the transmission capacity limit between the western Texas where wind potential is high and the eastern Texas where electricity demand is high, most curtailment occurs in the western Texas and the addition of energy storage cannot reduce it significantly (supplementary table 6).

As more renewable energy is added to the grid, the health impacts can be reduced without energy storage by the increase in non-emitting energy sources [53]. Supplementary figure 9 shows that the health impacts reduced by energy storage do not rise with increasing wind penetration on all selected days. However, energy storage still helps reduce health impacts further under the 1.5 × 2019 Wind scenario for the 12 selected days due to more clean energy integrated to the grid.

Figure 4 compares the health impacts reduced by energy storage under different combinations of the wind penetration and four selected energy storage capacities across the 12 days selected. Reductions in health impacts
are normalized by the energy storage capacity (MW) installed in the power system. Here we observe a wide range in the health damage reductions from the deployment of energy storage. We also see instances of net increases in health damages (as indicated by negative values). Since the objective of this model is to minimize the overall system-wide costs, such cases can be caused by the use of storage to reduce operational costs at the expense of increasing health damages (supplementary figure 11). In all cases, the operation of energy storage reduces total system costs, inclusive of operational costs and health damages. This finding is consistent with previous studies that show the potential for energy storage to increase emissions [54, 55]. In addition, there is energy loss associated with the charging and discharging process, which yields a need for additional generation and emissions.

We also observe high values in the range of health damage reductions attributable to storage. These results show that there is a subset of days in which strategically charging and discharging energy storage can yield large health benefits by avoiding generating electricity at extremely high marginal costs. Because such instances are limited, these results suggest that mitigating adverse health impacts could be a secondary application for storage devices whose primary purpose is different grid service (e.g., frequency regulation). This attractive secondary approach is sometimes called 'stacked services'. Additionally, due to generally higher health damage costs during the morning, energy storage tends to be charged before sunrise and in the afternoon and then injects the electricity back to the grid in the morning and evening when those costs are higher (supplementary figure 12). This charging pattern may help reduce the grid’s peak load, displacing a need for generating capacity.

4. Conclusions and discussions

When internalizing marginal health damage costs at the unit level into the power system operation, we estimate that the reduction in adverse health impacts from redispatch could exceed $5 billion per year in ERCOT if we annualize our results. These benefits would require an annual increase in operational costs of less than $300 million and be achieved by reducing electricity generation at several power plants with extremely high marginal health costs while increasing generation at others with much lower impacts. In addition, we estimate that the strategic operation of large-scale energy storage could further reduce health damages by $20 million per year. Energy storage will play a critical role in any grid with high renewable energy penetration regardless of its potential to reduce health impacts. Through the strategic operation of this emerging technology, power sector emissions can be further reduced.

We compare the health impact estimates from this study to other RCMs, namely EASIUR, InMAP, and APEEP, by applying the county-level social costs ($/ton) of elevated emissions of SO2 and NOx from these RCMs to ERCOT [14, 15, 56, 57]. The average health damage costs across all EGUs from our model are close to those estimated by other RCMs (supplementary figure 6). All models demonstrate considerable variability across EGUs in average marginal health damage costs, driven by their proximity to population centers. However, in our model, the costs also reflect the temporal variations in PM_{2.5} formation and transport. We capture the diurnal pattern of the marginal health damages that other models are unable to represent. These
fast-changing marginal impacts suggest that using a constant social cost for emissions will overlook the key role of emission location and timing in determining the negative health effects of the power sector, introducing biases into the quantification of these impacts. The hourly resolution used in this approach enables redispatching electricity and strategically operating energy storage to yield large health benefits by reducing electricity generation during hours of extremely high health damage costs. Similar to the integration of renewables, the internalization of health damage costs tends to introduce variability into power systems. However, Oates and Jaramillo (2013) have found that increased cycling of coal-fired EGUs due to 20% of wind penetration results in a reduction in emissions and cost penalties associated with cycling when the full costs of coal cycling are considered [58]. Additionally, we find that ramping of EGUs after redispatching decreases compared with the dispatch without the internalization of health damages due to most coal-fired EGUs being turned off (supplementary figure 13).

To ensure that our approach could be deployed at the power system scale and that the modeling could be completed in the time frames in which unit commitment and dispatch decisions are made, computational costs must be considered. As such, the use of CALPUFF to simulate PM$_{2.5}$ concentrations meets the necessary criteria for computational tractability while still capturing the major factors that drive its spatial distribution and temporal variability. We run CALPUFF more than 10,000 times to estimate the hourly unit-level health damage costs for over 200 EGUs for a single day. With high-performance computing available at the system operator-level, this approach can be readily deployed in the day-ahead decision making time frames. In assuming pseudo-first-order reactions, though, CALPUFF overlooks some of the complexity of atmospheric chemistry and physics. While the simulations allow us to quantify the impacts of each power plant, modeling secondary PM$_{2.5}$ concentrations associated with specific power plants in isolation makes it infeasible to compare the model output with PM$_{2.5}$ concentrations observed at monitoring sites. However, CALPUFF has been evaluated independently through multiple studies which find it can adequately simulate PM$_{2.5}$ from point sources [59–61]. Some of these have reported underestimations of secondary particulate matter, suggesting that our results may provide a conservative estimate of the potential reduction in adverse health impacts.

In this study, we apply a reported association between mortality and long-term population exposure to PM$_{2.5}$. The association is used to approximate the full costs of emissions from electricity generation, which are dominated by mortality associated with chronic exposure to PM$_{2.5}$. The hourly damage cost used to make decisions in the UC/ED model is a fraction of the long-term impacts. In contrast, applying functions that model acute impacts would likely underestimate the negative impacts associated with electricity production given the cumulative and persistent nature of pollution from power plants. Additionally, there is uncertainty associated with applying concentration-response functions to relate mortality to changes in PM concentration directly. There are alternative approaches in life cycle health impact assessments that have been proposed based on quantifying the amount of PM inhaled and the health damages caused by the pollutant intake [62, 63]. These approaches using dose-response functions may reduce uncertainties associated with the use of a concentration-response function and are worth future exploration in quantifying the health impacts from the power plant emissions. Although relative risk varies across diseases and age ranges, here we only consider all-cause mortality based on complete population. The log-linear concentration function and the relative risk from Krewski et al (2009) are frequently used in studies quantifying the health impacts associated with air pollution (e.g., Heo et al 2016; Dedoussi et al 2020; Son et al 2020) [14, 33, 64, 65]. Compared with other epidemiological studies (e.g., Laden et al 2006; Lepeule et al 2012), the relative risk used here results in a relatively conservative estimate of health damages and health benefits from redispatch [66, 67]. Supplementary table 3 shows that higher relative risk values yield very similar rates of reduction in adverse health impacts. In addition, given uncertainty in the VSL, we test a range of VSL values based on Viscusi (2015) [68]. We find that higher VSL values lead to greater monetized health benefits, but comparable rates of reduction in mortality. Our results are not very sensitive to VSL as even lower values still make health damages a dominant variable in the optimization model (supplementary table 4).

Although ‘super polluting’ power plants have been defined in a previous study from the perspective of emissions [69], through this study, we are able to identify ‘super polluting’ power plants in terms of health impacts and show how operational changes to these plants could result in dramatic health benefits. While the retirement of coal plants would yield immediate health benefits, the reduction of coal capacity requires long-term planning and new investment to maintain a reliable grid operation. The approach presented in this study can help reduce the negative health impacts associated with coal-fired generation during this transition. This is particularly important in coal-dominated grids that will require longer periods to achieve full retirement. While coal is becoming a smaller portion in ERCOT’s fuel mix, many regions in the world still rely heavily on coal in their power sector. Those regions can adopt this approach to reduce health damages associated with electricity generation cost-effectively when large-scale clean energy is not yet available. In addition, as shown in supplementary figure 3, natural gas can also lead to negative health impacts with high diurnal and seasonal
variability. Therefore, the approach is also suitable for a coal-free grid and can be used to reduce electricity generation from specific natural gas-fired power plants during unfavorable weather conditions. More broadly, our findings provide useful insights on how to regulate emissions to mitigate health impacts efficiently under already stringent emission requirements and benefit human health without reforming the power sector in the short term.

**Author contribution statement**

Qian Luo: methodology, software, formal analysis, investigation, writing—original draft. Fernando Garcia-Menendez: conceptualization, methodology, writing—review editing, supervision. Jeremiah Johnson: conceptualization, methodology, writing—review editing, supervision.

**Acknowledgments**

This work was supported by the National Science Foundation (NSF) Environmental Sustainability Program under Grant No. 1934276 and NC State Research and Innovation Seed Funding (RISF) Program. The authors express their gratitude to Harrison Fell and Melinda Morrill for providing power system data, reviewing the manuscript, and providing valuable insights and recommendations.

**Data availability statement**

The data that support the findings of this study are available upon reasonable request from the authors.

**ORCID iDs**

Qian Luo https://orcid.org/0000-0003-3894-4093
Jeremiah X Johnson https://orcid.org/0000-0002-9602-4976
Fernando Garcia-Menendez https://orcid.org/0000-0003-0235-5692

**References**

[1] Health Effects Institute 2019 State of global air 2019: a special report on global pollution and its disease burden *Tech. Rep.* (Boston, MA: Health Effects Institute (HEI))

[2] Pope C A III and Dockery D W 2006 Health effects of fine particulate air pollution: lines that connect *J. Air Waste Manage. Assoc.* 56 709–42

[3] Jerrett M, Burnett R T, Pope C A, Ito K, Thurston G, Krewski D, Shi Y, Calle E and Thun M 2009 Long-term ozone exposure and mortality *New Engl. J. Med.* 360 1085–95

[4] Apte J S, Brauer M, Cohen A J, Ezzati M and Pope C A III 2018 Ambient PM_{2.5} reduces global and regional life expectancy *Environ. Sci. Technol. Lett.* 5 546–51

[5] GBD 2017 Risk Factor Collaborators 2018 Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017 *Lancet* 392 1923

[6] Fann N, Fulcher C M and Baker K 2013 The recent and future health burden of air pollution apportioned across US sectors *Environ. Sci. Technol.* 47 3580–9

[7] Penn S L, Arunachalam S, Woody M, Heiger-Bernays W, Triopoulos Y and Levy J I 2017 Estimating state-specific contributions to PM_{2.5} and O₃ health burden from residential combustion and electricity generating unit emissions in the United States *Environ. Health Perspect.* 125 324–32

[8] Zhang H, Hu J, Kleeman M and Ying Q 2014 Source apportionment of sulfate and nitrate particulate matter in the Eastern United States and effectiveness of emission control programs *Sci. Total Environ.* 490 171–81

[9] Jin J, Coull B A, Schwartz J, Hubbell B and Koutrakis P 2015 The impact of weather changes on air quality and health in the United States in 1994–2012 *Environ. Res. Lett.* 10 084009

[10] Dawson J, Adams P and Pandis S 2007 Sensitivity of PM_{2.5} to climate in the Eastern US: a modeling case study *Atmos. Chem. Phys.* 7 4295–309

[11] Strasert B, Teh S C and Cohan D S 2019 Air quality and health benefits from potential coal power plant closures in Texas *J. Air Waste Manage. Assoc.* 69 333–50

[12] Zvingilaite E 2011 Human health-related externalities in energy system modelling the case of the Danish heat and power sector *Appl. Energy* 88 535–44

[13] Kerl P Y, Zhang W, Moreno-Cruz J B, Nenes A, Realf M J, Russell A G, Sokol J and Thomas V M 2015 New approach for optimal electricity planning and dispatching with hourly time-scale air quality and health considerations *Proc. Natl Acad. Sci. USA* 112 10884–9

[14] Heo J, Adams P J and Gao H O 2016 Reduced-form modeling of public health impacts of inorganic PM_{2.5} and precursor emissions *Atmos. Environ.* 137 80–9

[15] Tessum C W, Hill J D and Marshall J D 2017 InMAP: a model for air pollution interventions *PLoS One* 12 e0176131
[54] Hittinger E S and Azevedo I M L 2015 Bulk energy storage increases United States electricity system emissions *Environ. Sci. Technol.* 49 3203–10

[55] Arbabi Zadeh M, Sioshans R, Johnson J X and Keoleian G A 2019 The role of energy storage in deep decarbonization of electricity production *Nat. Commun.* 10 1–11

[56] Center for Air, Climate and Energy Solutions 2020 Social costs of SO2 and NOx from APEEP, EASIUR, and InMAP, developed by the center for air, climate and energy solutions https://caces.us/data/ (accessed 5 June 2020)

[57] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[58] Arbabzadeh M, Sioshansi R, Johnson J X and Keoleian G A 2019 The role of energy storage in deep decarbonization of electricity production *Nat. Commun.* 10 1–11

[59] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[60] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[61] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[62] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[63] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[64] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[65] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[66] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[67] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[68] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4

[69] Muller N Z 2014 Boosting GDP growth by accounting for the environment *Science* 345 873–4