

A Predictive Framework for Transition Probability Shifts Using Environmental Data Via Lee-Carter Model Extension*

Kenneth Gunawan¹, Nicholas Andrew Tanubrata¹, Rex Aaron Holindo¹,
Ferry Vincenttius Ferdinand^{1,†}, and Helena Margaretha^{1,**}

Abstract—This study addresses the challenge of incorporating environmental data, particularly air pollution, into practical actuarial models. This study extends the Lee-Carter model using environmental variables and employs LASSO regression to address potential multicollinearity among pollution indicators. The results are closed-form approximations for estimating the expected changes in transition probabilities from the previous transition probabilities forecast within both multiple-state and multiple-decrement models, which are calculated using current environmental data. These approximations enable actuaries to quantify the expected difference in transition probabilities from the initial forecast by utilizing the newest environmental data that was previously unavailable. The framework is empirically tested using a case study on a dementia-related multiple-state model in the UK. Results show significant improvements in predictive accuracy, measured by reductions in Mean Absolute Percentage Error (MAPE), reducing MAPE by over 40% in several age groups. These results emphasize the value of environmental integration in actuarial modeling, particularly in the context of health-related insurance products.

I. INTRODUCTION

ACCURATE mortality and morbidity projections are fundamental to actuarial practice because they affect the valuation of life and health insurance liabilities, pension obligations, and long-term risk management. The Lee-Carter model remains one of the most widely used stochastic mortality models because it provides a parsimonious representation of age-specific mortality through an age pattern and a common time-varying mortality index [1]. However, the Lee-Carter model is primarily extrapolative: future mortality is projected from patterns identified in historical mortality data. Consequently, changes in external conditions that are not reflected in the historical mortality trend may contribute to deviations between projected and subsequently observed mortality.

*This work was part of the research supported by the Directorate of Research and Community Service, Directorate General of Research and Development, Ministry of Higher Education, Science, and Technology of the Republic of Indonesia, under the Contract No. 0636/LL3/AL.04/2026

**corresponding author.

¹Department of Mathematics, Faculty of Science and Technology, Universitas Pelita Harapan Kenneth Gunawan: 01112220011@student.uph.edu, Nicholas Andrew Tanubrata: 01112220017@student.uph.edu, Rex Aaron Holindo: 01112220003@student.uph.edu, Ferry Vincenttius Ferdinand: ferry.ferdinand@uph.edu, Helena Margaretha: helena.margaretha@uph.edu

[†]Deceased on May 29, 2026.

Manuscript received July 31, 2026; accepted September 30, 2026.

Several studies have extended stochastic mortality models to improve flexibility and forecasting performance. Examples include neural-network extensions of the Lee-Carter framework [9] and age-coherent specifications designed to preserve plausible mortality relationships across ages [10]. Other studies have incorporated observable external information into mortality models. Niu and Melenberg [2] extended the Lee-Carter framework by including economic growth, while French and O'Hare [3] considered economic, health-expenditure, and lifestyle-related determinants in mortality forecasting. More recently, Dimai [5] examined economic, environmental, and lifestyle variables jointly in a stochastic mortality framework. These studies indicate that external covariates can provide information that is not contained directly in historical mortality rates and can support more interpretable mortality projections.

Environmental conditions are particularly relevant because they may change over time and are associated with important health outcomes. Pollution-related disease was estimated to account for approximately 9 million premature deaths worldwide in 2019 [7]. For ambient air pollution specifically, a systematic review and meta-analysis found consistent evidence that long-term exposure to particulate matter is associated with increased all-cause and several cause-specific mortality outcomes [4]. Air pollution has also been investigated in relation to neurological outcomes. A systematic review and meta-analysis reported evidence consistent with an association between ambient PM_{2.5} exposure and dementia, with more limited evidence for nitrogen dioxide and nitrogen oxides [6]. Evidence from the UK Biobank has likewise examined associations between ambient air pollution and incident dementia [18]. These findings provide a practical motivation for examining whether routinely observed air-pollution indicators can help explain deviations from mortality and morbidity patterns captured by a baseline stochastic model.

Air pollution was selected as the external data source in this study for both substantive and practical reasons. First, the pollutants considered in the empirical application have documented associations with mortality and dementia-related outcomes, making them plausible candidates for explaining part of the variation that remains after fitting a mortality-only model. Second, air-pollution concentrations are routinely measured by public monitoring systems and are observed separately from mortality data, which makes them suitable for a framework intended to update an existing actuarial projection when more recent external information becomes available.

Air pollution is not assumed to be the only determinant of mortality or morbidity. Economic conditions, health-care access and expenditure, lifestyle factors, smoking, temperature, and other environmental or demographic variables may also provide useful information, as considered in previous mortality studies [3], [5]. These variables were not included because the present study is designed to evaluate a focused environmental extension of the Lee–Carter framework rather than to construct a comprehensive causal model of mortality. Their omission therefore reflects the scope of the study, and the proposed framework can in principle accommodate other external covariates when suitable and temporally consistent data are available.

Most mortality models that incorporate external covariates focus on improving mortality-rate forecasts or explaining the evolution of latent mortality factors. In actuarial applications, however, mortality and morbidity risks are often represented through transition intensities in multiple-state and multiple-decrement models, where the quantities required for valuation are transition probabilities rather than intensities alone. Less attention has been given to translating newly observed external information into an explicit adjustment of previously projected transition probabilities. Such an adjustment is particularly relevant when a baseline forecast has already been produced, but more recent environmental information becomes available.

Accordingly, this study extends the Lee–Carter framework by modeling the residual component of the baseline model using lagged environmental indicators. Rather than replacing the original Lee–Carter structure, the proposed approach preserves its fitted parameters and introduces an environmental adjustment to the residual mortality component. LASSO regression is used to obtain a parsimonious environmental correction from several potentially related pollution indicators. The resulting adjustment is then propagated through the actuarial transition intensities to derive the expected changes in transition probabilities, using analytical expressions for the multiple-decrement model and second-order approximations for the multiple-state model.

The objective is therefore not merely to improve statistical prediction by introducing additional covariates, but to integrate observable environmental information into an actuarial mortality framework and quantify how the resulting mortality adjustment affects actuarially relevant transition probabilities. This distinction is particularly important in the multiple-state setting considered in this study, where the transition-probability relationships are governed by a system of coupled equations and the required adjustments do not have a simple analytical form. The environmental component is therefore kept parsimonious and interpretable so that its effect can be incorporated consistently into the subsequent actuarial derivation.

The empirical application considers a dementia-related multiple-state model in the UK, where air-pollution data are used to adjust transitions between healthy, dementia, and death states. The application illustrates how the environmental adjustment obtained from the statistical model is translated into changes in actuarial transition probabilities. The contribution of this study therefore lies not only in incorporating environ-

mental information into the Lee–Carter framework, but also in translating the resulting adjustment into explicit changes in transition probabilities for multiple-state and multiple-decrement actuarial models.

II. THEORETICAL FRAMEWORK

A. Lee-Carter Model

The Lee-Carter model is one of the most widely used stochastic mortality models and represents the logarithm of the central mortality rate as

$$\ln \mu_{x,t} = a_x + b_x k_t + \epsilon_{x,t}, \quad (1)$$

where $\mu_{x,t}$ denotes the central mortality rate at age x and year t ; a_x captures the average mortality level at age x ; b_x represents the age-specific sensitivity to changes in the mortality index; k_t is the time-varying index capturing the general mortality trend; and $\epsilon_{x,t}$ is the residual term, assumed to be normally distributed with mean zero.

Recent developments have extended and adapted the Lee-Carter framework to account for heterogeneity across populations, cohort effects, and structural changes in mortality trends. For instance, a previous study proposed a functional data approach that generalizes the Lee-Carter model for age-specific fertility and age-specific mortality [11]. These extensions are particularly useful in actuarial science, as reliable mortality projections enable insurers to address challenges associated with lifespan and death risk [12].

Throughout this study, the Lee-Carter component is written consistently as $a_x + b_x k_t$. The environmental extension introduced later is applied to the residual component $\epsilon_{x,t}$ while retaining the original Lee-Carter parameters.

B. LASSO Regression

LASSO regression (Least Absolute Shrinkage and Selection Operator) is a regularized regression technique that performs both variable selection and coefficient shrinkage, improving model interpretability and prediction accuracy, especially in high-dimensional settings. LASSO regression estimates the regression coefficients by minimizing the residual sum of squares subject to a penalty on the sum of the absolute values of the coefficients:

$$\hat{\beta} = \arg \min_{\beta} \left\{ \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\}, \quad (2)$$

where $\lambda \geq 0$ is the regularization parameter that controls the degree of shrinkage: as λ increases, more coefficients are forced to zero, resulting in a sparser model.

LASSO regression is particularly useful when a model contains several potentially related predictors because the L_1 penalty performs coefficient shrinkage and variable selection simultaneously [13], [14]. As the regularization parameter increases, coefficients associated with predictors that contribute limited information may be reduced to zero, resulting in a more parsimonious model. This property is useful when

environmental variables and their lagged values may contain overlapping information.

In the present study, interpretability and parsimony are particularly important because the regression model is not the final component of the framework. Its estimated coefficients are subsequently used to construct an environmental adjustment to the Lee–Carter mortality rate, which is then propagated through the transition intensities and transition probabilities of the actuarial model. A sparse linear specification therefore provides a transparent connection between the environmental predictors and the subsequent actuarial derivation.

C. Disability Income Multiple State Model

The Disability Income Multiple State Model (DIMSM) is an actuarial modeling framework designed to represent the transitions of individuals between multiple health states - typically including active, disabled, and dead - and to assess the associated financial risks, such as disability income benefits and life insurance liabilities. This model extends the classical two-state (alive/dead) life insurance framework by incorporating intermediate health states and allowing transitions between them over time.

A common structure of a DIMSM includes several defined states, such as:

- State 0: Active
- State 1: Disabled
- State 2: Death

Transitions between these states are governed by age- and time-dependent transition intensities or probabilities denoted as $\mu_{ij}(x, t)$, where i and j represent the originating and destination states, respectively. For example, $\mu_{01}(x, t)$ is the transition intensity from active to disabled at age x and time t . The multiple-state approach allows for a richer representation of morbidity dynamics and can incorporate recovery, recurrent disability, and cause-specific mortality. The DIMSM is particularly useful in analyzing the financial impact of disability and mortality under uncertainty and time dependence.

Unlike a simple two-state survival model, the multiple-state framework requires the simultaneous consideration of several transition intensities. The resulting transition probabilities depend jointly on the possible movements between states. This structure becomes important in the proposed environmental extension because an adjustment to one or more transition intensities must be propagated consistently through the entire transition system.

D. Multiple Decrement Model

The Multiple Decrement Model is a classical actuarial framework used to represent the probability of exit from an initial state due to several competing risks or causes of decrement, such as death, disability, or withdrawal. Unlike single decrement models, which consider only one mode of exit, the multiple decrement model accounts for the fact that multiple, mutually exclusive causes may lead to a transition out of the initial state.

Formally, let $q_x^{(j)}$ denote the probability that an individual aged x exits due to cause j within one year, and let q_x be the

total probability of exit from the initial state due to all causes combined. The model assumes:

$$q_x = \sum_{j=1}^k q_x^{(j)}, \quad (3)$$

where k is the number of decrement causes. The decrement probabilities are determined by the corresponding cause-specific forces of decrement, denoted by $\mu_x^{(j)}$, such that:

$$q_x^{(j)} = \int_x^{x+1} p_x(t) \mu_t^{(j)} dt, \quad (4)$$

where $p_x(t)$ is the probability of surviving (not exiting) to time t .

In this study, the Multiple Decrement Model is employed to disaggregate total risk into its constituent causes, enabling a clearer understanding of how different pathways, such as death from different originating states, contribute to the overall exit probability. This decomposition is crucial for the accurate valuation of insurance liabilities and for the analysis of public health policies.

III. METHODOLOGY

A. Environmental Modeling of Lee-Carter Residuals

The proposed extension preserves the original Lee-Carter structure and models the component of the observed mortality rate that is not explained by the baseline model. From Equation (1), the Lee-Carter residual is

$$\epsilon_{x,t} = \ln \mu_{x,t} - (a_x + b_x k_t). \quad (5)$$

A positive residual indicates that the observed mortality rate is higher than the mortality rate represented by the baseline Lee–Carter component, whereas a negative residual indicates the opposite.

Rather than modifying the Lee-Carter parameters a_x , b_x , and k_t , the proposed framework models this residual using lagged environmental variables. This preserves the baseline mortality structure and allows the environmental adjustment to be propagated explicitly through the subsequent actuarial transition model.

The environmental model is defined as

$$\epsilon_{x,t} = c_x^0 + \sum_{j=1}^k (c_x^{1,j} y_{t-j}^1 + c_x^{2,j} y_{t-j}^2 + \dots + c_x^{n,j} y_{t-j}^n) + \epsilon'_{x,t}, \quad (6)$$

where c_x^0 is the intercept of the environmental model, $c_x^{p,j}$, $p \neq 0$ is the coefficient of the p^{th} environmental index, and y_{t-j}^p is the p^{th} environmental index at time $t-j$. Also, n is the number of environmental variables, k is the maximum lag considered, and $\epsilon'_{x,t}$ is the residual term of the environmentally extended model.

The coefficients $c_x^{p,j}$ are estimated using LASSO regression. LASSO is selected because its sparse linear structure simultaneously performs coefficient shrinkage and variable selection while retaining an explicit mapping between each environmental predictor and the resulting mortality adjustment. This interpretability is important because the estimated

environmental coefficients are subsequently incorporated into the transition intensities and propagated through the multiple-state and multiple-decrement models. More complex nonlinear machine-learning models could be considered for prediction alone, but would not provide the same direct analytical link between the environmental predictors and the subsequent actuarial derivation.

B. Defining Fitted Mortality Rates

The new mortality rate of the extended model will be defined as

$$\mu'_{x,t} = \exp(a_x + b_x k_t + c_x^0 + \sum_{j=1}^k (c_x^{1,j} y_{t-j}^1 + c_x^{2,j} y_{t-j}^2 + \dots + c_x^{n,j} y_{t-j}^n) + \epsilon'_{x,t}). \quad (7)$$

Assuming that $\epsilon_{x,t}$ is normally distributed with mean zero and variance $\sigma_{\epsilon_{x,t}}^2$, the fitted mortality rate from the original Lee–Carter model is

$$\begin{aligned} \bar{\mu}_{x,t} &= E[\mu_{x,t}] \\ &= E[\exp(a_x + b_x k_t + \epsilon_{x,t})] \\ &= \exp\left(a_x + b_x k_t + \frac{1}{2} \sigma_{\epsilon_{x,t}}^2\right). \end{aligned} \quad (8)$$

Similarly, assuming that $\epsilon'_{x,t}$ is normally distributed with mean zero and variance $\sigma_{\epsilon'_{x,t}}^2$, the fitted mortality rate of the environmentally extended model is

$$\begin{aligned} \bar{\mu}'_{x,t} &= \exp\left(a_x + b_x k_t + c_x^0 + \sum_{j=1}^k \sum_{p=1}^n (c_x^{p,j} y_{t-j}^p) + \frac{1}{2} \sigma_{\epsilon'_{x,t}}^2\right) \\ &= \bar{\mu}_{x,t} \tilde{c}_{x,t}, \end{aligned} \quad (9)$$

where

$$\tilde{c}_{x,t} = \exp\left[c_x^0 + \sum_{j=1}^k \sum_{p=1}^n (c_x^{p,j} y_{t-j}^p) + \frac{1}{2} (\sigma_{\epsilon'_{x,t}}^2 - \sigma_{\epsilon_{x,t}}^2)\right]. \quad (10)$$

C. Model Estimation and Predictive Evaluation

The model is estimated and evaluated chronologically to preserve the temporal structure of the data. Observations from 1990 to 2023 are used as the training period for estimating the baseline Lee–Carter model and the LASSO-based environmental adjustment. The observations for 2024 are excluded from model estimation and are reserved as the out-of-sample test data.

For the environmental extension, the 2024 adjustment uses only lagged environmental observations available prior to the prediction year. In the empirical application, where one- and two-year lags are considered, the environmental predictors for the 2024 prediction are therefore obtained from 2023 and 2022, respectively. No 2024 outcome information is used in estimating either the Lee–Carter parameters or the environmental regression coefficients.

The predictive contribution of the environmental information is evaluated by comparing the 2024 prediction errors of the original Lee–Carter model and the environmentally extended model using the same observed 2024 data. The Mean Absolute Percentage Error (MAPE) is used as the performance measure. Consequently, a lower MAPE for the environmentally extended model reflects improved predictive performance on observations that were not used during model estimation, rather than merely an improvement in in-sample fit.

D. Integration of the Environmental and Actuarial Components

The statistical and actuarial components of the proposed framework are sequential. The LASSO regression is used only to estimate the environmental coefficients in Equation (6), from which the adjustment factor $\tilde{c}_{x,t}$ in Equation (10) is obtained. No further machine-learning estimation is performed in the subsequent transition-probability derivations.

For a transition from state i to state j , the corresponding fitted baseline transition intensity $\bar{\mu}_{x,t}^{ij}$ is adjusted as

$$\bar{\mu}_{x,t}^{ij'} = \bar{\mu}_{x,t}^{ij} \tilde{c}_{x,t}^{ij}. \quad (11)$$

The adjusted intensities are then substituted into the actuarial multiple-state or multiple-decrement model. Thus, the environmental regression produces the adjustment factor, while the equations that follow propagate this factor into the transition probabilities required for actuarial applications.

E. Probability Changes Formulas

In this study, we will derive the analytical formula for the probability changes in a multiple-state and a multiple-decrement model. The goal of deriving the formula is to determine the expected change in the transition probabilities by utilizing the environmental indices data from previous periods.

1) *Multiple-state Model*: In this section, we derive the changes in transition probabilities for a single sickness state disability income model. We will denote ${}_s p_{x,t}^{ij}$ as the probability that a person of age x in state i at time t will be in state j at time $t+s$ using the original Lee–Carter formula, and ${}_s p_{x,t}^{ij'}$ will be the transition probabilities calculated using the extended formula. The calculation of the transition probabilities will utilize the Kolmogorov forward equations defined by:

$$\frac{d}{ds} {}_s p_{x,t}^{00} = {}_s p_{x,t}^{01} \bar{\mu}_{x+s,t}^{10} - {}_s p_{x,t}^{00} (\bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02}), \quad (12)$$

$$\frac{d}{ds} {}_s p_{x,t}^{01} = {}_s p_{x,t}^{00} \bar{\mu}_{x+s,t}^{01} - {}_s p_{x,t}^{01} (\bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12}), \quad (13)$$

$$\frac{d}{ds} {}_s p_{x,t}^{02} = {}_s p_{x,t}^{00} \bar{\mu}_{x+s,t}^{02} + {}_s p_{x,t}^{01} \bar{\mu}_{x+s,t}^{12}, \quad (14)$$

$$\frac{d}{ds} {}_s p_{x,t}^{10} = {}_s p_{x,t}^{11} \bar{\mu}_{x+s,t}^{10} - {}_s p_{x,t}^{10} (\bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02}), \quad (15)$$

$$\frac{d}{ds} {}_s p_{x,t}^{11} = {}_s p_{x,t}^{10} \bar{\mu}_{x+s,t}^{01} - {}_s p_{x,t}^{11} (\bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12}), \quad (16)$$

$$\frac{d}{ds} s p_{x,t}^{12} = s p_{x,t}^{10} \bar{\mu}_{x+s,t}^{02} + s p_{x,t}^{11} \bar{\mu}_{x+s,t}^{12}. \quad (17)$$

Closed-form expressions for the transition probabilities become cumbersome in the multiple-state setting. To obtain tractable expressions that allow the environmental adjustment to be propagated explicitly, this study employs a second-order Taylor approximation. Under the CFM assumption, the second derivatives of the transition probabilities are

$$\begin{aligned} \frac{d^2}{ds^2} s p_{x,t}^{00} &= s p_{x,t}^{00} \left[\bar{\mu}_{x+s,t}^{01} \bar{\mu}_{x+s,t}^{10} + (\bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02})^2 \right] \\ &\quad - s p_{x,t}^{01} \bar{\mu}_{x+s,t}^{10} \left(\bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12} + \bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02} \right). \end{aligned} \quad (18)$$

$$\begin{aligned} \frac{d^2}{ds^2} s p_{x,t}^{01} &= s p_{x,t}^{01} \left[\bar{\mu}_{x+s,t}^{10} \bar{\mu}_{x+s,t}^{01} + (\bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12})^2 \right] \\ &\quad - s p_{x,t}^{00} \bar{\mu}_{x+s,t}^{01} \left(\bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02} + \bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12} \right). \end{aligned} \quad (19)$$

$$\begin{aligned} \frac{d^2}{ds^2} s p_{x,t}^{02} &= s p_{x,t}^{00} \left[-\bar{\mu}_{x+s,t}^{02} (\bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02}) + \bar{\mu}_{x+s,t}^{01} \bar{\mu}_{x+s,t}^{12} \right] \\ &\quad + s p_{x,t}^{01} \left[\bar{\mu}_{x+s,t}^{10} \bar{\mu}_{x+s,t}^{02} - \bar{\mu}_{x+s,t}^{12} (\bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12}) \right]. \end{aligned} \quad (20)$$

$$\begin{aligned} \frac{d^2}{ds^2} s p_{x,t}^{10} &= s p_{x,t}^{10} \left[\bar{\mu}_{x+s,t}^{10} \bar{\mu}_{x+s,t}^{01} + (\bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02})^2 \right] \\ &\quad - s p_{x,t}^{11} \bar{\mu}_{x+s,t}^{10} \left(\bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12} + \bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02} \right). \end{aligned} \quad (21)$$

$$\begin{aligned} \frac{d^2}{ds^2} s p_{x,t}^{11} &= s p_{x,t}^{11} \left[\bar{\mu}_{x+s,t}^{10} \bar{\mu}_{x+s,t}^{01} + (\bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12})^2 \right] \\ &\quad - s p_{x,t}^{10} \bar{\mu}_{x+s,t}^{01} \left(\bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02} + \bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12} \right). \end{aligned} \quad (22)$$

$$\begin{aligned} \frac{d^2}{ds^2} s p_{x,t}^{12} &= s p_{x,t}^{10} \left[-\bar{\mu}_{x+s,t}^{02} (\bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02}) + \bar{\mu}_{x+s,t}^{01} \bar{\mu}_{x+s,t}^{12} \right] \\ &\quad + s p_{x,t}^{11} \left[\bar{\mu}_{x+s,t}^{10} \bar{\mu}_{x+s,t}^{02} - \bar{\mu}_{x+s,t}^{12} (\bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12}) \right]. \end{aligned} \quad (23)$$

Let h denote the projection interval. Under the Constant Force of Mortality (CFM) assumption, each transition intensity is treated as constant over $[t, t+h)$. Considering that ${}_0p_{x,t}^{ij}$ is equal to 1 if $i = j$ and is equal to 0 if $i \neq j$, the Taylor approximations of ${}_s p_{x,t}^{ij}$ are

$$\begin{aligned} {}_h p_{x,t}^{00} &\approx 1 - h \left(\bar{\mu}_{x,t}^{01} + \bar{\mu}_{x,t}^{02} \right) \\ &\quad + \frac{1}{2} h^2 \left[\bar{\mu}_{x,t}^{01} \bar{\mu}_{x,t}^{10} + (\bar{\mu}_{x,t}^{01} + \bar{\mu}_{x,t}^{02})^2 \right]. \end{aligned} \quad (24)$$

$$\begin{aligned} {}_h p_{x,t}^{01} &\approx h \bar{\mu}_{x,t}^{01} - \frac{1}{2} h^2 \bar{\mu}_{x,t}^{01} \\ &\quad \times \left(\bar{\mu}_{x+s,t}^{01} + \bar{\mu}_{x+s,t}^{02} + \bar{\mu}_{x+s,t}^{10} + \bar{\mu}_{x+s,t}^{12} \right). \end{aligned} \quad (25)$$

$$\begin{aligned} {}_h p_{x,t}^{02} &\approx h \bar{\mu}_{x,t}^{02} + \frac{1}{2} h^2 \left[-\bar{\mu}_{x,t}^{02} (\bar{\mu}_{x,t}^{01} + \bar{\mu}_{x,t}^{02}) + \bar{\mu}_{x,t}^{01} \bar{\mu}_{x,t}^{12} \right]. \end{aligned} \quad (26)$$

$$\begin{aligned} {}_h p_{x,t}^{10} &\approx h \bar{\mu}_{x,t}^{10} - \frac{1}{2} h^2 \bar{\mu}_{x,t}^{10} \\ &\quad \times \left(\bar{\mu}_{x,t}^{10} + \bar{\mu}_{x,t}^{12} + \bar{\mu}_{x,t}^{01} + \bar{\mu}_{x,t}^{02} \right). \end{aligned} \quad (27)$$

$$\begin{aligned} {}_h p_{x,t}^{11} &\approx 1 - h \left(\bar{\mu}_{x,t}^{10} + \bar{\mu}_{x,t}^{12} \right) \\ &\quad + \frac{1}{2} h^2 \left[\bar{\mu}_{x,t}^{10} \bar{\mu}_{x,t}^{01} + (\bar{\mu}_{x,t}^{10} + \bar{\mu}_{x,t}^{12})^2 \right]. \end{aligned} \quad (28)$$

$$\begin{aligned} {}_h p_{x,t}^{12} &\approx h \bar{\mu}_{x,t}^{12} + \frac{1}{2} h^2 \left[\bar{\mu}_{x,t}^{10} \bar{\mu}_{x,t}^{02} - \bar{\mu}_{x,t}^{12} (\bar{\mu}_{x,t}^{10} + \bar{\mu}_{x,t}^{12}) \right]. \end{aligned} \quad (29)$$

To calculate the expected change in transition probabilities, we will define the expected change as

$${}_h \Delta_{x,t}^{ij} := {}_h p_{x,t}^{ij'} - {}_h p_{x,t}^{ij}. \quad (30)$$

The values of ${}_h p_{x,t}^{ij}$ can be calculated using Equations (24)-(29). The values of ${}_h p_{x,t}^{ij'}$ can also be calculated using Equations (24)-(29), by replacing every $\bar{\mu}_{x,t}^{ij}$ into $\bar{\mu}_{x,t}^{ij} \cdot \tilde{c}_{x,t}^{ij}$. The final formulas for the expected changes are

$$\begin{aligned} {}_h \Delta_{x,t}^{00} &= h \left[(\tilde{c}_{x,t}^{01} - 1) \bar{\mu}_{x,t}^{01} + (\tilde{c}_{x,t}^{02} - 1) \bar{\mu}_{x,t}^{02} \right] \\ &\quad + \frac{1}{2} h^2 \left[\bar{\mu}_{x,t}^{01} \bar{\mu}_{x,t}^{10} (\tilde{c}_{x,t}^{01} \tilde{c}_{x,t}^{10} - 1) + (\tilde{c}_{x,t}^{01} \bar{\mu}_{x,t}^{01} + \tilde{c}_{x,t}^{02} \bar{\mu}_{x,t}^{02})^2 \right. \\ &\quad \left. - (\bar{\mu}_{x,t}^{01} + \bar{\mu}_{x,t}^{02})^2 \right]. \end{aligned} \quad (31)$$

$$\begin{aligned}
{}_h\Delta_{x,t}^{01} &= h(\tilde{c}_{x,t}^{01} - 1)\bar{\mu}_{x,t}^{01} \\
&+ \frac{1}{2}h^2\bar{\mu}_{x,t}^{01}(1 - \tilde{c}_{x,t}^{01}) \left[\right. \\
&\quad \bar{\mu}_{x,t}^{01}(1 - \tilde{c}_{x,t}^{01}) + \bar{\mu}_{x,t}^{02}(1 - \tilde{c}_{x,t}^{02}) \\
&\quad \left. + \bar{\mu}_{x,t}^{10}(1 - \tilde{c}_{x,t}^{10}) + \bar{\mu}_{x,t}^{12}(1 - \tilde{c}_{x,t}^{12}) \right]. \quad (32)
\end{aligned}$$

$$\begin{aligned}
{}_h\Delta_{x,t}^{02} &= h\bar{\mu}_{x,t}^{02}(\tilde{c}_{x,t}^{02} - 1) \\
&+ \frac{1}{2}h^2 \left[\bar{\mu}_{x,t}^{02}(1 - \tilde{c}_{x,t}^{02}) \left\{ \right. \right. \\
&\quad \bar{\mu}_{x,t}^{01}(1 - \tilde{c}_{x,t}^{01}) + \bar{\mu}_{x,t}^{02}(1 - \tilde{c}_{x,t}^{02}) \left. \right\} \\
&\quad \left. + \bar{\mu}_{x,t}^{01}\bar{\mu}_{x,t}^{12}(\tilde{c}_{x,t}^{01}\tilde{c}_{x,t}^{12} - 1) \right]. \quad (33)
\end{aligned}$$

$$\begin{aligned}
{}_h\Delta_{x,t}^{10} &= h\bar{\mu}_{x,t}^{10}(\tilde{c}_{x,t}^{10} - 1) \\
&+ \frac{1}{2}h^2\bar{\mu}_{x,t}^{10}(1 - \tilde{c}_{x,t}^{10}) \left[\right. \\
&\quad \bar{\mu}_{x,t}^{10}(1 - \tilde{c}_{x,t}^{10}) + \bar{\mu}_{x,t}^{12}(1 - \tilde{c}_{x,t}^{12}) \\
&\quad \left. + \bar{\mu}_{x,t}^{01}(1 - \tilde{c}_{x,t}^{01}) + \bar{\mu}_{x,t}^{02}(1 - \tilde{c}_{x,t}^{02}) \right]. \quad (34)
\end{aligned}$$

$$\begin{aligned}
{}_h\Delta_{x,t}^{11} &= h \left[\bar{\mu}_{x,t}^{10}(1 - \tilde{c}_{x,t}^{10}) \right. \\
&\quad \left. + \bar{\mu}_{x,t}^{12}(1 - \tilde{c}_{x,t}^{12}) \right] \\
&+ \frac{1}{2}h^2 \left[\bar{\mu}_{x,t}^{10}\bar{\mu}_{x,t}^{01}(\tilde{c}_{x,t}^{10}\tilde{c}_{x,t}^{01} - 1) \right. \\
&\quad \left. + (\tilde{c}_{x,t}^{10}\bar{\mu}_{x,t}^{10} + \tilde{c}_{x,t}^{12}\bar{\mu}_{x,t}^{12})^2 \right. \\
&\quad \left. - (\bar{\mu}_{x,t}^{10} + \bar{\mu}_{x,t}^{12})^2 \right]. \quad (35)
\end{aligned}$$

$$\begin{aligned}
{}_h\Delta_{x,t}^{12} &= h\bar{\mu}_{x,t}^{12}(\tilde{c}_{x,t}^{12} - 1) \\
&+ \frac{1}{2}h^2 \left[\bar{\mu}_{x,t}^{10}\bar{\mu}_{x,t}^{02}(\tilde{c}_{x,t}^{10}\tilde{c}_{x,t}^{02} - 1) \right. \\
&\quad \left. + \bar{\mu}_{x,t}^{12}(1 - \tilde{c}_{x,t}^{12}) \left\{ \right. \right. \\
&\quad \left. \left. \bar{\mu}_{x,t}^{10}(1 - \tilde{c}_{x,t}^{10}) + \bar{\mu}_{x,t}^{12}(1 - \tilde{c}_{x,t}^{12}) \right\} \right]. \quad (36)
\end{aligned}$$

2) *Multiple-decrement Model*: For the multiple decrement model, this study will derive the expected change formulas in a double decrement framework, namely, the expected change of the first decrement and the second decrement. The notation ${}_h q_{x,t}^{(i)}$ refers to the probability that a person surviving until age x dies due to the i^{th} decrement. Using the CFM assumption, the values of ${}_h q_{x,t}^{(i)}$ can be calculated analytically as

$${}_h q_{x,t}^{(1)} = \frac{\bar{\mu}_{x,t}^{(1)}}{\bar{\mu}_{x,t}^{(1)} + \bar{\mu}_{x,t}^{(2)}} \left(1 - e^{-h(\bar{\mu}_{x,t}^{(1)} + \bar{\mu}_{x,t}^{(2)})} \right), \quad (37)$$

$${}_h q_{x,t}^{(2)} = \frac{\bar{\mu}_{x,t}^{(2)}}{\bar{\mu}_{x,t}^{(1)} + \bar{\mu}_{x,t}^{(2)}} \left(1 - e^{-h(\bar{\mu}_{x,t}^{(1)} + \bar{\mu}_{x,t}^{(2)})} \right). \quad (38)$$

To calculate the expected change in the decrement probabilities, we will define the expected change as

$${}_h\Delta_{x,t}^{(i)} = {}_h q_{x,t}^{(i)'} - {}_h q_{x,t}^{(i)}. \quad (39)$$

The values of ${}_h q_{x,t}^{(i)}$ can be calculated using Equations (37) and (38). The values of ${}_h q_{x,t}^{(i)'}$ can be calculated by using (37) and (38), as well as replacing $\bar{\mu}_{x,t}^{(i)}$ into $\bar{\mu}_{x,t}^{(i)} \cdot \tilde{c}_{x,t}^{(i)}$. The final expected change formulas can be calculated by

$$\begin{aligned}
\Delta_{x,t}^{(1)} &= \frac{c_{x,t}^{(1)}\mu_{x,t}^{(1)}}{c_{x,t}^{(1)}\mu_{x,t}^{(1)} + c_{x,t}^{(2)}\mu_{x,t}^{(2)}} \\
&\times \left\{ 1 - \exp \left[-\frac{c_{x,t}^{(1)}\mu_{x,t}^{(1)} + c_{x,t}^{(2)}\mu_{x,t}^{(2)}}{12} \right] \right\} \\
&- \frac{\mu_{x,t}^{(1)}}{\mu_{x,t}^{(1)} + \mu_{x,t}^{(2)}} \\
&\times \left\{ 1 - \exp \left[-\frac{\mu_{x,t}^{(1)} + \mu_{x,t}^{(2)}}{12} \right] \right\}. \quad (40)
\end{aligned}$$

$$\begin{aligned}
\Delta_{x,t}^{(2)} &= \frac{c_{x,t}^{(2)}\mu_{x,t}^{(2)}}{c_{x,t}^{(1)}\mu_{x,t}^{(1)} + c_{x,t}^{(2)}\mu_{x,t}^{(2)}} \\
&\times \left\{ 1 - \exp \left[-\frac{c_{x,t}^{(1)}\mu_{x,t}^{(1)} + c_{x,t}^{(2)}\mu_{x,t}^{(2)}}{12} \right] \right\} \\
&- \frac{\mu_{x,t}^{(2)}}{\mu_{x,t}^{(1)} + \mu_{x,t}^{(2)}} \\
&\times \left\{ 1 - \exp \left[-\frac{\mu_{x,t}^{(1)} + \mu_{x,t}^{(2)}}{12} \right] \right\}. \quad (41)
\end{aligned}$$

IV. RESULTS AND DISCUSSION

The proposed framework is evaluated using a multiple-state dementia model for the UK, with mortality and morbidity data obtained from the Office for National Statistics¹ and air-pollution data obtained from the UK government official database [15]. The permanent disability model considered in the application is a special case of the disability income model, with $\mu_{x,t}^{10} = 0$ for all x and t .

As described in the Methodology, the models are estimated using data from 1990–2023, while the 2024 observations are reserved for out-of-sample evaluation. All MAPE values are reported for the UK application therefore refer to the 2024 predictions. The predictive performance of the original Lee–Carter model is compared with that of the environmentally extended Lee–Carter model using the same observed 2024 data. A lower MAPE for the extended model indicates that the lagged environmental variables provide additional predictive information beyond that captured by the baseline Lee–Carter model for the 2024 test observations.

The environmental variables considered are SO₂, NO, PM₁₀, PM_{2.5}, ammonia, and non-methane hydrocarbons, with one-

¹<https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsregistrationsummarytables/2022>

and two-year lags. The corresponding changes in transition probabilities are obtained using Equations (31)–(36) and the environmental adjustment factor $\tilde{c}_{x,t}$ defined in Equation (10). The analysis is presented separately for each transition and sex because these represent distinct actuarial risk pathways, and the environmental adjustment is not expected to affect all transitions, ages, and sexes uniformly.

An additional application to respiratory and non-respiratory mortality in the US is presented in Appendix B. This additional case illustrates the use of the same environmental extension in a different mortality setting and geographical context.

A. Results

Table I summarizes the main MAPE reduction patterns across the six sex-specific transition models. The transition-specific results are discussed separately because each transition represents a different actuarial risk process and the patterns vary across age and sex.

1) *Male Population:* For male subjects transitioning from healthy to dementia, the environmentally extended Lee–Carter model reduced the 2024 MAPE by more than 55% for the age groups 45–49 through 95+, compared with the original Lee–Carter model. The improvement was smaller for the 40–44 age group, with a MAPE reduction of less than 30%. Figure 1 shows the age-specific comparison between the two models.

The LASSO estimates show that the selected pollutants and coefficient directions vary across age groups. Detailed age-specific coefficients are reported in Appendix Table II. In the Appendix tables, “+” denotes a positive estimated LASSO coefficient, “-” denotes a negative estimated coefficient, and “0” indicates that the coefficient was shrunk to zero by the LASSO penalty.

For male subjects transitioning from healthy to death, the environmentally extended model reduced the 2024 MAPE by more than 40% for the age groups 50–54 through 75–79 and for the 95+ group. Smaller reductions of less than 20% were observed for ages 45–49 and 90–94, while the reduction for ages 40–44 was less than 0.5%. Figure 2 presents the age-specific comparison.

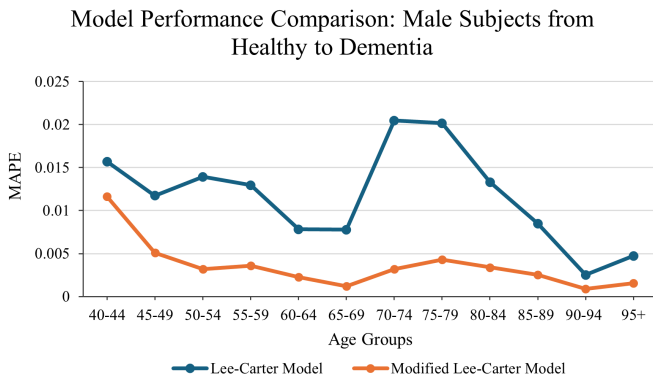


Fig. 1: Model performance (MAPE) for male subjects transitioning from healthy to dementia.

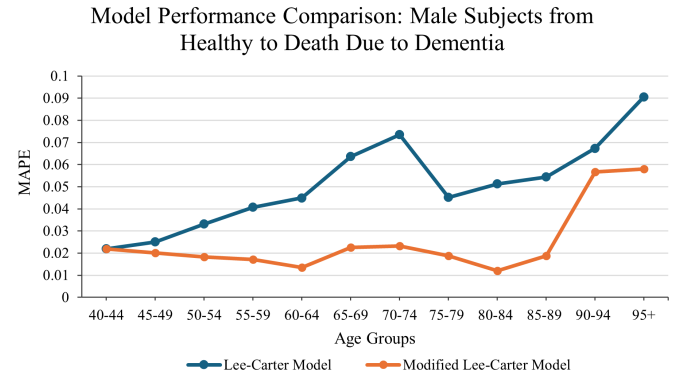


Fig. 2: Model performance (MAPE) for male subjects transitioning from healthy to death.

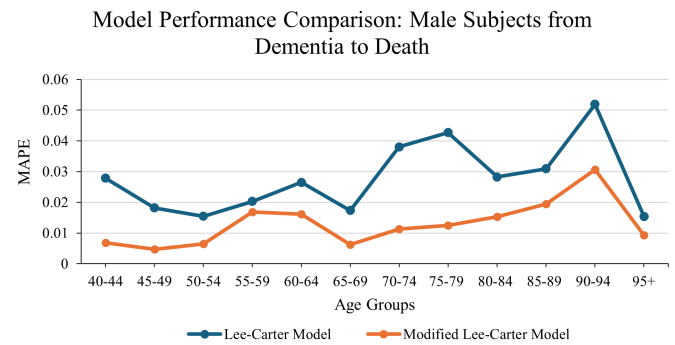


Fig. 3: Model performance (MAPE) for male subjects transitioning from dementia to death.

The corresponding LASSO coefficient patterns also differ across age groups and pollutant lags. Detailed results are provided in Appendix Table III.

For male subjects transitioning from dementia to death, the environmentally extended model reduced the 2024 MAPE by more than 40% for ages 40–44, 45–49, 50–54, and 65–69 through 80–84. Reductions of approximately 37–39% were observed for ages 60–64, 85–89, and 95+, whereas the 55–59 age group showed a smaller reduction of less than 17%. Figure 3 shows the corresponding age-specific comparison.

The selected environmental variables and coefficient directions again vary across ages, indicating that the environmental adjustment differs from those observed for the healthy-to-dementia and healthy-to-death transitions. Detailed LASSO results are reported in Appendix Table IV.

2) *Female Population:* For female subjects transitioning from healthy to dementia, the environmentally extended model reduced the 2024 MAPE by more than 40% for ages 40–44 through 50–54, 65–69 through 80–84, and 90+. Smaller reductions of less than 20% were observed for ages 55–59, 60–64, and 85–89. The age-specific comparison is shown in Figure 4.

The age-specific LASSO coefficients also differ from those obtained for the corresponding male transition, supporting separate analysis by sex. Detailed results are reported in

TABLE I: Summary of MAPE reductions for the 2024 out-of-sample predictions

Sex	Transition	Larger MAPE reductions	Smaller MAPE reductions
Male	Healthy → Dementia	More than 55% for ages 45–49 to 95+	Less than 30% for ages 40–44
Male	Healthy → Death	More than 40% for ages 50–54 to 75–79 and 95+	Less than 20% for ages 45–49 and 90–94; less than 0.5% for ages 40–44
Male	Dementia → Death	More than 40% for ages 40–54 and 65–84	37–39% for ages 60–64, 85–89, and 95+; less than 17% for ages 55–59
Female	Healthy → Dementia	More than 40% for ages 40–54, 65–84, and 90+	Less than 20% for ages 55–64 and 85–89
Female	Healthy → Death	More than 40% for ages 40–44 to 85–89	Less than 20% for ages 90+
Female	Dementia → Death	More than 45% for ages 40–54 and most older age groups	Less than 20% for ages 55–59; less than 32% for ages 70–74

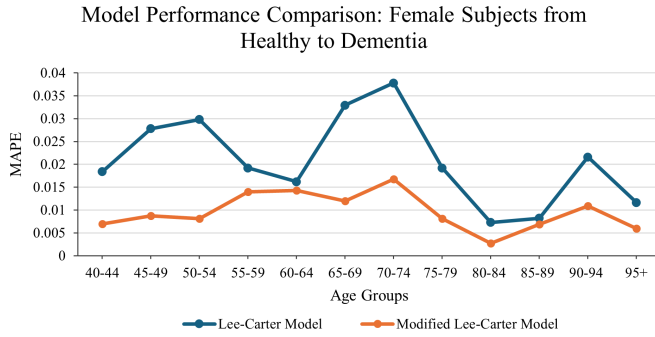


Fig. 4: Model performance (MAPE) for female subjects transitioning from healthy to dementia.

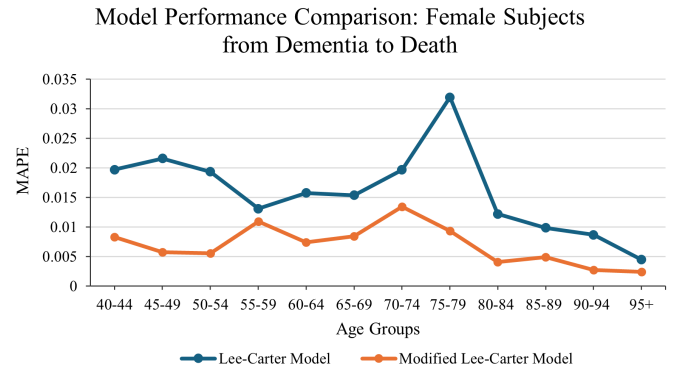


Fig. 6: Model performance (MAPE) for female subjects transitioning from dementia to death.

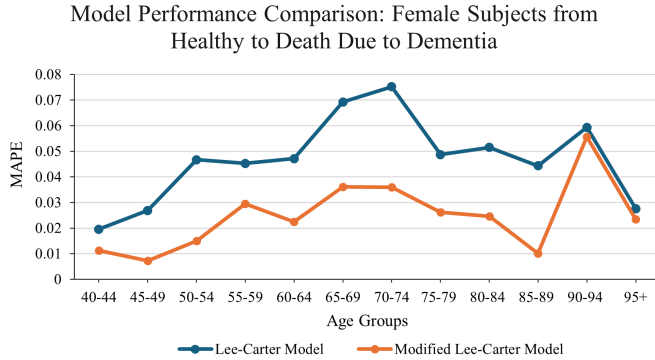


Fig. 5: Model performance (MAPE) for female subjects transitioning from healthy to death.

Appendix Table V.

For female subjects transitioning from healthy to death, the environmentally extended model reduced the 2024 MAPE by more than 40% across ages 40–44 through 85–89. For ages 90+, the improvement was smaller, with a MAPE reduction of less than 20%. Figure 5 presents the age-specific comparison.

The associated LASSO coefficient patterns are reported in Appendix Table VI. As in the other transitions, the selected environmental variables vary by age, indicating that a common environmental adjustment across all age groups would not capture the observed pattern.

For female subjects transitioning from dementia to death,

the environmentally extended model produced substantial reductions in the 2024 MAPE across most age groups. Reductions exceeded 45% for ages 40–44 through 50–54 and across most older age groups. The reduction was less than 20% for ages 55–59 and less than 32% for ages 70–74. Figure 6 presents the complete age-specific comparison.

Detailed pollutant- and age-specific LASSO coefficients are reported in Appendix Table VII.

B. Discussion

The results show that the contribution of environmental information is not uniform across transition pathways, age groups, or sex. This heterogeneity is important from an actuarial perspective because the healthy-to-dementia, healthy-to-death, and dementia-to-death transitions represent different components of the multiple-state risk process. The environmental adjustment should therefore be evaluated separately for each transition rather than applied uniformly to the entire multiple-state model. The differences between male and female results further indicate that the same environmental information may contribute differently to sex-specific transition intensities.

To illustrate the pollutant-specific patterns in greater detail, the male healthy-to-dementia transition is considered. For SO_2 , the lag-1 coefficient was positive for ages 65–69 through 90–94, negative for ages 55–59 and 95+, and shrunk to zero for the

remaining age groups. The lag-2 coefficient was positive for ages 70–74 through 90–94 and negative for several younger age groups and the 95+ group. Previous research reported an association between higher SO₂ exposure and lower MMSE (Mini-Mental State Examination) scores among individuals with mild cognitive impairment [17]. The variation in coefficient direction across ages in the present model indicates that the environmental contribution is not uniform across the age range.

For NO, the lag-1 coefficient was positive for ages 65–69 through 90–94 and negative for ages 50–54 through 60–64 and 95+, while the coefficients for the remaining age groups were shrunk to zero. The lag-2 pattern differed, with positive coefficients across several younger and middle-to-older age groups. These findings are consistent with previous evidence reporting an increased risk of dementia associated with higher nitrogen-oxide exposure among older individuals in the UK [18].

The particulate-matter variables also show age-dependent patterns. For PM₁₀, positive lag-1 and lag-2 coefficients occur predominantly among several older age groups, whereas negative coefficients are more frequent among younger age groups and at age 95+. Previous research using two-sample multivariable Mendelian randomization reported an association between higher PM₁₀ exposure and Alzheimer disease risk [19]. For PM_{2.5}, the coefficient directions differ from those of PM₁₀, including positive coefficients for several younger age groups and negative coefficients across several older groups. Previous evidence has also linked increased PM_{2.5} exposure with dementia risk among older individuals [18]. These differences reinforce that the estimated environmental adjustment depends on both pollutant type and age rather than following a single common exposure pattern.

For ammonia, the lag-1 coefficient was positive for ages 40–44, 45–49, 65–69 through 75–79, and 90–94, while the coefficient was shrunk to zero for ages 80–84 and was negative for the remaining age groups. The lag-2 coefficient was negative for ages 40–44 through 65–69 and 95+, but positive for ages 70–74 through 90–94. The alternating coefficient directions indicate that the estimated relationship varies across both age and lag period. Previous research has suggested that ammonia is unlikely to be a primary cause of Alzheimer's disease, although it is recognized as a neurotoxic compound that can affect mitochondrial and cognitive functions [20].

For non-methane hydrocarbons, the lag-1 coefficient was positive for ages 50–54 through 60–64 and 95+, negative for ages 65–69 through 90–94, and shrunk to zero for ages 40–49. For lag 2, positive coefficients were obtained for ages 65–69 and 70–74, while negative coefficients occurred for ages 50–54, 55–59, and 90–94; the remaining coefficients were shrunk to zero. These patterns partially correspond to previous findings suggesting that exposure to total hydrocarbons and non-methane hydrocarbons may contribute to dementia development among adults aged 50 years and above [21].

Across the six transition models, the differences in MAPE reduction and LASSO coefficient patterns illustrate why the transition pathways should be analyzed separately. From an actuarial perspective, an environmental adjustment to the

healthy-to-dementia intensity does not have the same implication as an adjustment to the healthy-to-death or dementia-to-death intensity. Each adjusted intensity enters the multiple-state system differently and consequently produces a different change in the corresponding transition probabilities.

Overall, the environmental variables contribute differently across ages, sexes, and transition pathways. The purpose of the proposed framework is therefore not to assign a single environmental effect to mortality or morbidity, but to provide a mechanism through which updated environmental information can modify individual transition intensities and subsequently the corresponding transition probabilities. This is particularly relevant for multiple-state actuarial models, where changes in different transition pathways have different implications for projected state occupancy and future insurance liabilities.

The present analysis focuses on air pollution because it provides a well-documented and routinely observed environmental source of information. This does not imply that air pollution is the only relevant external determinant of mortality or morbidity. Economic, demographic, lifestyle, health-care, and other environmental variables may also contribute useful information and can, in principle, be incorporated within the same framework.

V. CONCLUSION

This study develops an extension of the Lee–Carter mortality framework that incorporates environmental information while preserving the original actuarial mortality structure. Lagged air-pollution variables are used to model the excess-mortality component through LASSO regression, providing a parsimonious and interpretable environmental adjustment that can be incorporated into the subsequent actuarial derivation.

The main contribution is not limited to incorporating external predictors into mortality forecasting. The estimated environmental adjustment is propagated through transition intensities to derive expected changes in transition probabilities for multiple-state and multiple-decrement models. In the multiple-state setting, second-order approximations provide tractable expressions for the transition-probability changes, while the multiple-decrement formulation permits analytical expressions under the Constant Force of Mortality assumption.

For the UK dementia application, the models are estimated using 1990–2023 data and evaluated on the 2024 observations. Across the six sex-specific transition models, the environmentally extended model produces lower MAPE than the original Lee–Carter model, although the magnitude of the reduction varies across transitions, ages, and sex. These differences reinforce the importance of treating each transition intensity as a distinct actuarial risk process rather than imposing a common environmental adjustment across the multiple-state model. The additional US application further illustrates that the framework can be applied to a different mortality setting and environmental data structure.

Overall, the proposed framework provides a way to connect observable environmental conditions with actuarially relevant transition probabilities. Future work may extend the framework to other external variables, including socio-demographic,

lifestyle, and health-care indicators, and may examine how the resulting changes in transition probabilities affect premiums, reserves, and other insurance quantities in practical applications.

ACKNOWLEDGMENT

This article is dedicated to the memory of Dr. Ferry Vincent-tius Ferdinand. The authors gratefully acknowledge the late Dr. Ferry Vincenttius Ferdinand for his valuable contributions to the supervision and validation of this study and to the preparation of the original manuscript draft.

This research forms the preliminary part of a research project funded by the Directorate of Research and Community Service, Directorate General of Research and Development, Ministry of Higher Education, Science, and Technology of the Republic of Indonesia, under the Contract No. 0636/LL3/AL.04/2026, dated 27 April 2026.

REFERENCES

- [1] R. D. Lee and L. R. Carter, "Modeling and forecasting U.S. mortality," *Journal of the American Statistical Association*, vol. 87, no. 419, pp. 659–671, 1992.
- [2] G. Niu and B. Melenberg, "Trends in mortality decrease and economic growth," *Demography*, vol. 51, no. 5, pp. 1755–1773, 2014.
- [3] D. French and C. O'Hare, "Forecasting death rates using exogenous determinants," *Journal of Forecasting*, vol. 33, no. 8, pp. 640–650, 2014.
- [4] J. Chen and G. Hoek, "Long-term exposure to PM and all-cause and cause-specific mortality: A systematic review and meta-analysis," *Environment International*, vol. 143, Art. no. 105974, 2020.
- [5] M. Dimai, "Modeling and forecasting mortality with economic, environmental and lifestyle variables," *Decisions in Economics and Finance*, vol. 48, pp. 1161–1195, 2025.
- [6] E. H. Wilker, M. Osman, and M. G. Weisskopf, "Ambient air pollution and clinical dementia: Systematic review and meta-analysis," *BMJ*, vol. 381, Art. no. e071620, 2023.
- [7] R. Fuller et al., "Pollution and health: A progress update," *The Lancet Planetary Health*, vol. 6, no. 6, pp. e535–e547, 2022.
- [8] Z. M. Dong, H. L. Shang, and A. Bruhn, "Air pollution and mortality impacts," *Risks*, vol. 10, no. 6, Art. no. 126, 2022.
- [9] R. Richman and M. V. Wüthrich, "A neural network extension of the Lee–Carter model to multiple populations," *Annals of Actuarial Science*, vol. 15, no. 2, pp. 346–366, 2021.
- [10] G. Gao and Y. Shi, "Age-coherent extensions of the Lee–Carter model," *Scandinavian Actuarial Journal*, vol. 2021, no. 10, pp. 998–1016, 2021.
- [11] R. J. Hyndman and M. S. Ullah, "Robust forecasting of mortality and fertility rates: A functional data approach," *Computational Statistics & Data Analysis*, vol. 51, no. 10, pp. 4942–4956, 2007.
- [12] D. Atance and E. Navarro, "A simplified model for measuring longevity risk for life insurance products," *Financial Innovation*, vol. 10, no. 1, Art. no. 61, 2024.
- [13] S. Altelbany, "Evaluation of ridge, elastic net and LASSO regression methods in precedence of multicollinearity problem: A simulation study," *Journal of Applied Economics and Business Studies*, vol. 5, no. 1, pp. 131–142, 2021.
- [14] E. Konzen and F. A. Ziegelmann, "LASSO-type penalties for covariate selection and forecasting in time series," *Journal of Forecasting*, vol. 35, no. 7, pp. 592–612, 2016.
- [15] GOV.UK, "Air pollutants data from GOV.UK." [Online]. Available: <https://www.gov.uk/government/statistical-data-sets/env01-emissions-of-air-pollutants>
- [16] X. Shang et al., "Leading mediators of sex differences in the incidence of dementia in community-dwelling adults in the UK Biobank: A retrospective cohort study," *Alzheimer's Research & Therapy*, vol. 15, no. 1, Art. no. 7, 2023.
- [17] M.-C. Chen et al., "Air pollution is associated with poor cognitive function in Taiwanese adults," *International Journal of Environmental Research and Public Health*, vol. 18, no. 1, Art. no. 316, 2021.
- [18] Y. Zhang et al., "Associations of ambient air pollution exposure and lifestyle factors with incident dementia in the elderly: A prospective study in the UK Biobank," *Environment International*, vol. 190, Art. no. 108870, 2024.
- [19] Z. Huang, G. He, S. Sun, Y. Feng, and Y. Huang, "Causal associations of ambient particulate matter 10 and Alzheimer's disease: Result from a two-sample multivariable Mendelian randomization study," *Archives of Medical Science*, vol. 20, no. 5, pp. 1604–1618, 2024.
- [20] A. Adlimoghaddam, M. G. Sabbir, and B. C. Albeni, "Ammonia as a potential neurotoxic factor in Alzheimer's disease," *Frontiers in Molecular Neuroscience*, vol. 9, Art. no. 57, 2016.
- [21] H.-W. Zhang et al., "Long-term exposure to ambient hydrocarbons increases dementia risk in people aged 50 years and above in Taiwan," *Current Alzheimer Research*, vol. 16, no. 14, pp. 1276–1289, 2019.
- [22] U.S. Environmental Protection Agency, "Air quality data from U.S. EPA." [Online]. Available: <https://www.epa.gov/outdoor-air-quality-data/download-daily-data>
- [23] Centers for Disease Control and Prevention, "U.S. mortality data from CDC WONDER." [Online]. Available: <https://wonder.cdc.gov/ucd-icd10.html>

APPENDIX A DETAILED LASSO REGRESSION RESULTS FOR THE UK DEMENTIA APPLICATION

This appendix presents the age-specific LASSO regression results for the six sex-specific transition models considered in the UK dementia application. The symbol “+” denotes a positive estimated LASSO coefficient, “-” denotes a negative estimated coefficient, and “0” indicates that the corresponding coefficient was shrunk to zero by the LASSO penalty.

APPENDIX B ADDITIONAL US APPLICATION

Besides the dementia case using the UK dataset explained in this paper, we also applied the methodology to another country (i.e., the United States) using monthly environmental data from the Environmental Protection Agency (EPA) [22] and mortality data from CDC WONDER [23] over the period 2000–2020. The environmental variables include air pollution and meteorological indicators, while the air-pollution series are represented using weighted exponential averages with a one-month delay (Lag-1). For the US case, we choose respiratory disease, and we implemented the multiple decrement model as described in this paper. For respiratory disease mortality, the environmentally extended Lee–Carter model reduces MAPE by approximately 20–33% for individuals aged 54–62. Figure 7 presents the age-specific model comparison, while the detailed LASSO coefficient patterns are reported in Table VIII.

For non-respiratory disease mortality, the extended model produces MAPE reductions of more than 30% in several younger age groups, whereas the reductions are generally below 20% for individuals aged 70 years and above. Figure 8 presents the age-specific comparison, and the corresponding LASSO coefficient patterns are reported in Table IX. These differences indicate that the incremental contribution of the environmental variables varies across age groups and mortality outcomes. In Tables VIII and IX, “+” denotes a positive estimated LASSO coefficient, “-” denotes a negative estimated coefficient, and “0” indicates that the coefficient was shrunk to zero by the LASSO penalty.

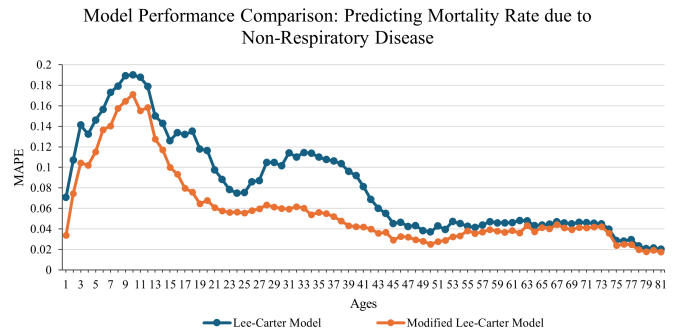


Fig. 8: Model performance based on MAPE for mortality due to non-respiratory disease in the US.

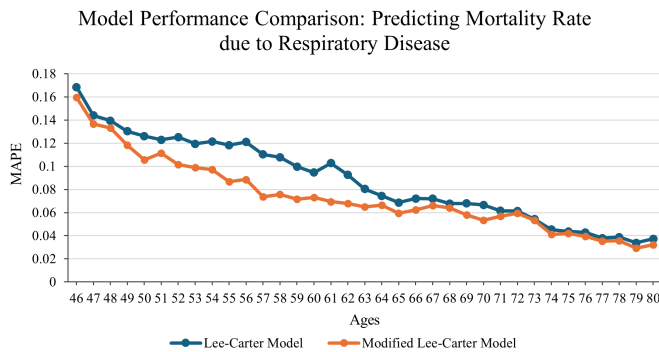


Fig. 7: Model performance based on MAPE for mortality due to respiratory disease in the US.

TABLE II: LASSO Regression Results for Male Subjects Transitioning from Healthy to Dementia, by Pollutant and Age Group

Variable	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95+
(Intercept)	+	+	-	-	-	-	-	-	+	+	+	-
SO ₂ lag 1	0	0	0	-	0	+	+	+	+	+	+	-
SO ₂ lag 2	0	-	-	-	-	-	+	+	+	+	+	-
NO lag 1	0	0	-	-	-	+	+	+	+	+	+	-
NO lag 2	+	+	+	+	0	+	+	+	+	+	0	0
PM ₁₀ lag 1	0	-	-	-	-	+	+	+	+	+	+	-
PM ₁₀ lag 2	-	-	-	-	-	+	+	+	+	+	+	-
PM _{2.5} lag 1	0	+	+	+	+	-	-	-	-	-	-	+
PM _{2.5} lag 2	0	+	+	+	+	-	-	-	-	-	-	+
Ammonia lag 1	+	+	-	-	-	+	+	+	0	-	+	-
Ammonia lag 2	-	-	-	-	-	-	+	+	+	+	+	-
Non-methane lag 1	0	0	+	+	+	-	-	-	-	-	-	+
Non-methane lag 2	0	0	-	-	0	+	+	0	0	0	-	0

TABLE III: LASSO Regression Results for Male Subjects Transitioning from Healthy to Death from Dementia, by Pollutant and Age Group

Variable	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95+
Intercept	-	-	-	-	-	-	-	-	-	-	-	+
SO ₂ lag 1	0	0	-	-	+	0	+	+	+	-	+	0
SO ₂ lag 2	0	0	-	-	-	-	-	-	-	-	0	0
NO lag 1	0	0	-	-	-	-	-	-	+	-	0	0
NO lag 2	0	0	+	0	+	+	+	+	+	+	-	0
PM ₁₀ lag 1	0	0	0	0	-	-	-	-	+	-	0	0
PM ₁₀ lag 2	0	+	+	0	-	0	-	-	+	-	-	0
PM _{2.5} lag 1	0	0	+	+	+	+	+	+	-	+	-	-
PM _{2.5} lag 2	0	0	0	+	+	+	+	+	0	+	-	0
Ammonia lag 1	0	-	-	-	-	-	0	-	-	-	+	+
Ammonia lag 2	0	+	-	-	-	-	-	-	-	-	-	+
Non-methane lag 1	0	0	0	0	-	0	-	0	-	+	-	0
Non-methane lag 2	0	0	0	+	0	0	+	+	+	-	0	0

TABLE IV: LASSO Regression Results for Male Subjects Transitioning from Dementia to Death, by Pollutant and Age Group

Variable	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95+
(Intercept)	+	-	+	+	+	-	+	+	+	-	+	-
SO ₂ lag 1	-	+	0	0	0	+	+	0	+	+	+	-
SO ₂ lag 2	-	-	+	0	0	+	+	+	+	0	+	-
NO lag 1	+	-	0	0	+	+	-	-	0	0	+	0
NO lag 2	0	+	+	-	0	-	0	-	-	-	0	+
PM ₁₀ lag 1	-	-	-	0	+	+	0	0	-	+	-	0
PM ₁₀ lag 2	0	-	-	0	+	+	-	-	-	-	-	+
PM _{2.5} lag 1	-	+	+	0	0	-	-	-	-	0	0	0
PM _{2.5} lag 2	-	+	0	0	0	-	+	+	+	+	0	0
Ammonia lag 1	-	+	+	+	0	-	+	+	+	+	+	-
Ammonia lag 2	-	+	-	0	+	-	+	+	+	-	-	+
Non-methane lag 1	+	-	+	0	0	-	-	+	-	-	-	+
Non-methane lag 2	+	+	-	0	-	+	-	0	-	0	-	+

TABLE V: LASSO Regression Results for Female Subjects Transitioning from Healthy to Dementia, by Pollutant and Age Group

Variable	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95+
(Intercept)	-	-	-	-	+	-	+	+	-	-	-	+
SO ₂ lag 1	+	+	+	+	0	0	0	-	+	+	0	-
SO ₂ lag 2	+	+	+	0	0	-	-	-	+	0	+	-
NO lag 1	+	-	-	0	0	+	0	+	+	0	+	0
NO lag 2	0	+	+	-	+	0	+	+	0	-	0	0
PM ₁₀ lag 1	0	0	-	0	0	0	0	0	+	0	0	0
PM ₁₀ lag 2	-	-	-	0	0	0	0	+	+	0	0	+
PM _{2.5} lag 1	-	-	+	0	0	0	0	+	-	0	0	+
PM _{2.5} lag 2	0	+	+	0	0	0	0	-	-	0	0	0
Ammonia lag 1	+	+	+	+	0	-	-	-	+	+	+	-
Ammonia lag 2	+	+	+	-	+	-	-	-	+	-	+	-
Non-methane lag 1	-	-	+	0	-	0	0	+	-	0	0	+
Non-methane lag 2	-	-	-	0	-	0	-	0	-	0	-	+

TABLE VI: LASSO Regression Results for Female Subjects Transitioning from Healthy to Death from Dementia, by Pollutant and Age Group

Variable	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95+
Intercept	-	-	-	-	-	-	-	-	-	-	-	+
SO ₂ lag 1	-	+	0	0	0	0	0	+	+	-	+	-
SO ₂ lag 2	-	-	-	0	0	0	0	-	0	-	0	0
NO lag 1	+	+	+	+	+	+	+	+	+	-	0	-
NO lag 2	0	-	+	0	0	0	0	+	0	+	0	0
PM ₁₀ lag 1	-	+	-	0	0	0	0	-	+	-	0	0
PM ₁₀ lag 2	0	+	-	0	0	0	0	-	0	-	0	0
PM _{2.5} lag 1	0	-	+	+	+	+	+	+	-	+	0	0
PM _{2.5} lag 2	0	+	0	0	0	0	0	0	-	+	0	0
Ammonia lag 1	-	-	-	-	-	-	-	-	-	-	0	+
Ammonia lag 2	-	-	-	-	-	-	-	-	-	-	0	+
Non-methane lag 1	0	-	0	0	0	0	0	0	-	+	0	0
Non-methane lag 2	+	+	0	+	0	0	0	+	0	-	0	0

TABLE VII: LASSO Regression Results for Female Subjects Transitioning from Dementia to Death, by Pollutant and Age Group

Variable	40-44	45-49	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85-89	90-94	95+
(Intercept)	-	-	+	+	+	+	+	+	+	-	-	+
SO ₂ lag 1	+	+	-	0	0	0	0	-	0	-	+	0
SO ₂ lag 2	+	-	+	-	+	-	0	-	+	-	+	+
NO lag 1	+	-	-	0	-	+	-	-	0	-	+	-
NO lag 2	-	+	+	0	+	0	0	0	-	0	+	0
PM ₁₀ lag 1	+	+	-	0	0	-	0	0	+	-	-	0
PM ₁₀ lag 2	-	-	-	0	0	0	0	-	+	-	0	0
PM _{2.5} lag 1	-	-	+	0	-	0	0	0	-	+	+	0
PM _{2.5} lag 2	0	+	0	0	0	0	+	0	0	+	0	0
Ammonia lag 1	+	+	+	+	+	-	+	0	+	-	+	+
Ammonia lag 2	-	+	-	+	+	-	+	+	+	-	+	+
Non-methane lag 1	-	-	+	0	-	0	+	+	-	0	-	0
Non-methane lag 2	+	+	-	0	-	+	0	+	0	-	-	0

TABLE VIII: LASSO Regression by Pollutants from Healthy to Death Due to Respiratory Disease

Age	Intercept	AQI	Wind	Temp.	Humid.	SO ₂	CO	NO ₂	O ₃	PM ₁₀	PM _{2.5}	Butad.	Acetal.
46	-	+	-	-	-	+	+	-	-	-	+	-	+
47	-	0	-	0	+	-	-	0	0	-	0	-	0
48	-	0	-	-	+	0	-	0	+	-	0	+	+
49	-	0	-	-	-	-	0	0	0	0	0	-	+
50	-	+	-	-	-	-	+	-	-	+	-	-	+
51	-	0	-	-	-	-	0	-	0	+	0	-	+
52	-	+	-	0	-	-	+	-	+	+	+	-	+
53	-	0	-	-	0	-	-	+	-	0	+	+	+
54	-	+	-	-	-	0	-	-	-	-	-	+	+
55	-	+	0	-	+	-	+	+	-	-	+	+	+
56	-	-	+	+	-	-	+	+	0	-	+	+	+
57	-	+	-	-	-	-	+	-	-	-	+	-	-
58	-	+	-	-	-	-	+	+	0	+	0	+	+
59	-	+	-	-	+	-	+	-	-	+	-	+	+
60	-	+	-	+	+	-	+	-	+	-	+	+	-
61	-	+	-	-	+	-	+	-	0	-	0	+	-
62	-	+	-	-	-	0	+	-	-	+	-	-	+
63	-	0	-	0	-	0	0	+	0	+	0	0	0

TABLE VIII: LASSO Regression by Pollutants from Healthy to Death Due to Respiratory Disease (Continued)

Age	Intercept	AQI	Wind	Temp.	Humid.	SO ₂	CO	NO ₂	O ₃	PM ₁₀	PM _{2.5}	Butad.	Acetal.
64	-	-	-	-	+	+	-	+	+	-	+	-	-
65	-	-	-	-	+	0	-	+	+	-	+	-	-
66	-	-	-	-	-	+	-	+	-	0	-	-	+
67	-	+	-	-	-	+	+	-	-	-	-	-	+
68	-	-	+	+	-	-	-	+	+	+	+	+	-
69	-	0	-	+	-	+	-	-	-	-	+	+	-
70	-	-	+	-	+	0	-	0	0	-	+	-	+
71	-	+	+	+	+	-	+	-	-	-	-	+	+
72	-	-	+	0	-	0	-	+	-	+	+	0	-
73	-	-	-	-	-	+	-	-	+	-	+	-	-
74	-	+	-	-	-	+	-	+	-	+	-	+	-
75	-	+	+	+	-	+	-	+	-	-	-	+	+
76	-	-	+	+	+	-	-	+	+	-	-	+	+
77	-	0	+	+	+	0	+	+	-	+	0	-	-
78	-	-	+	+	0	+	+	0	0	-	+	-	-
79	-	+	+	-	+	0	0	-	+	+	0	+	-
80	-	-	+	+	+	-	-	+	+	+	+	+	-

TABLE IX: LASSO Regression by Pollutants from Healthy to Death Due to Non-Respiratory Disease

Age	Intercept	AQI	Wind	Temp.	Humid.	SO ₂	CO	NO ₂	O ₃	PM ₁₀	PM _{2.5}	Butad.	Acetal.
0	-	+	+	-	+	+	-	+	+	+	-	+	-
1	-	-	-	-	-	+	-	+	+	+	-	+	-
2	-	+	+	-	+	+	-	+	+	+	-	+	-
3	-	+	-	-	-	+	-	+	+	+	-	+	-
4	-	-	+	-	-	+	-	+	+	-	-	-	-
5	-	+	+	-	+	+	0	+	+	+	-	-	-
6	-	+	+	-	+	+	-	+	+	+	-	-	-
7	-	+	-	-	+	+	-	0	+	+	-	+	-
8	-	-	+	-	-	+	-	+	+	+	-	+	-
9	-	0	+	+	-	0	0	0	0	+	+	-	-
10	-	+	+	+	+	-	+	+	-	+	-	-	-
11	-	-	+	-	-	+	-	-	+	+	-	+	-
12	-	-	+	+	-	-	+	+	+	+	+	-	-
13	-	-	-	-	0	+	+	+	+	+	+	0	-
14	-	-	-	-	+	+	+	+	+	+	-	-	-
15	-	-	+	-	+	+	-	+	+	+	-	+	-
16	-	-	+	-	-	+	-	+	+	+	+	+	-
17	-	-	+	-	+	+	-	+	+	+	+	+	-
18	-	-	-	-	-	+	-	-	+	+	+	+	-
19	-	-	-	-	-	+	-	0	+	+	+	+	-
20	-	-	-	-	-	+	-	0	+	+	+	+	-
21	-	-	-	-	-	+	-	-	+	+	+	+	-
22	-	-	-	-	+	+	-	0	+	+	+	+	-
23	-	-	-	-	-	+	-	-	+	+	+	+	-
24	-	-	-	-	+	+	-	0	+	+	+	+	-
25	-	-	-	-	-	+	-	+	+	+	+	+	-
26	-	-	-	-	-	+	-	0	+	+	0	+	-

TABLE IX: LASSO Regression by Pollutants from Healthy to Death Due to Non-Respiratory Disease (Continued)

Age	Intercept	AQI	Wind	Temp.	Humid.	SO ₂	CO	NO ₂	O ₃	PM ₁₀	PM _{2.5}	Butad.	Acetal.
27	-	-	-	-	+	+	-	-	+	+	+	+	-
28	-	-	-	-	+	+	-	+	+	+	+	+	-
29	-	-	-	-	+	+	-	-	+	+	+	+	-
30	-	-	-	-	-	+	-	0	+	+	+	+	-
31	-	-	-	-	-	+	-	-	+	+	+	+	-
32	-	-	-	-	-	+	-	-	+	+	+	+	-
33	-	-	-	-	-	+	-	-	+	-	+	+	-
34	-	-	-	-	-	+	+	-	+	-	+	+	-
35	-	-	+	-	-	0	+	-	+	+	+	-	-
36	-	-	-	-	-	+	-	-	+	+	+	+	-
37	-	-	-	-	+	+	-	-	+	+	+	+	-
38	-	-	-	-	+	+	-	-	+	+	+	+	-
39	-	-	-	-	-	+	+	-	+	+	-	+	-
40	-	-	-	-	-	+	0	0	+	+	+	+	-
41	-	-	-	-	-	+	0	-	+	+	-	+	-
42	-	-	-	-	-	+	+	+	+	+	+	-	-
43	-	-	-	-	+	+	+	-	+	-	+	-	-
44	-	-	-	-	+	+	+	-	+	-	+	+	-
45	-	-	-	-	-	+	+	-	+	+	+	+	-
46	-	+	-	-	-	+	-	-	+	+	-	+	-
47	-	-	-	-	-	+	-	-	+	+	+	+	-
48	-	-	-	-	+	-	-	+	+	+	+	+	-
49	-	-	-	-	+	+	+	0	-	-	+	-	0
50	-	-	-	-	-	+	+	-	+	-	+	-	-
51	-	-	-	-	+	-	+	-	+	-	+	-	-
52	-	-	-	-	-	+	+	-	+	+	+	-	-
53	-	-	-	-	+	0	-	-	+	+	+	+	+

TABLE IX: LASSO Regression by Pollutants from Healthy to Death Due to Non-Respiratory Disease (Continued)

Age	Intercept	AQI	Wind	Temp.	Humid.	SO ₂	CO	NO ₂	O ₃	PM ₁₀	PM _{2.5}	Butad.	Acetal.
54	-	+	-	-	+	-	-	+	+	-	-	+	+
55	-	-	+	+	+	-	-	+	-	-	+	-	-
56	-	-	+	+	-	-	-	+	+	-	+	+	-
57	-	0	-	-	+	-	+	-	+	+	-	-	+
58	-	-	+	+	+	-	-	+	+	+	+	-	+
59	-	+	-	0	+	-	0	-	+	-	-	+	-
60	-	-	-	+	+	+	-	0	+	-	+	+	-
61	-	+	-	-	+	0	+	-	+	-	0	+	-
62	-	+	+	-	-	-	+	-	-	+	-	-	+
63	-	-	+	+	-	+	-	+	-	+	-	-	+
64	-	-	-	-	+	+	-	+	+	-	+	+	-
65	-	-	-	-	+	-	-	+	+	-	+	+	-
66	-	+	+	-	-	+	+	-	-	+	-	-	+
67	-	+	+	-	-	+	+	-	-	+	-	-	+
68	-	-	+	+	-	-	-	+	+	+	+	+	-
69	-	-	-	+	-	+	-	-	+	-	+	+	-
70	-	-	+	+	+	-	+	+	-	+	-	-	+
71	-	+	+	+	+	-	+	-	-	-	-	+	+
72	-	-	-	+	-	0	-	0	-	+	+	-	-
73	-	-	-	-	-	+	-	-	+	+	+	-	-
74	-	+	-	-	-	+	-	-	-	-	+	+	-
75	-	+	+	+	-	+	-	+	-	-	-	+	+
76	-	-	+	+	+	-	-	+	-	-	-	+	+
77	-	-	-	+	-	-	+	-	-	-	+	-	+
78	-	+	+	+	+	0	+	-	-	+	+	-	+
79	-	+	-	-	+	-	+	-	-	-	-	+	+
80	-	-	+	+	-	-	-	+	-	+	+	-	-