

Da

The seal of the University of Padua, featuring two figures (a monk and a scholar) under a gothic arch, surrounded by the Latin text "UNIVERSITAS STUDII PADUAE".

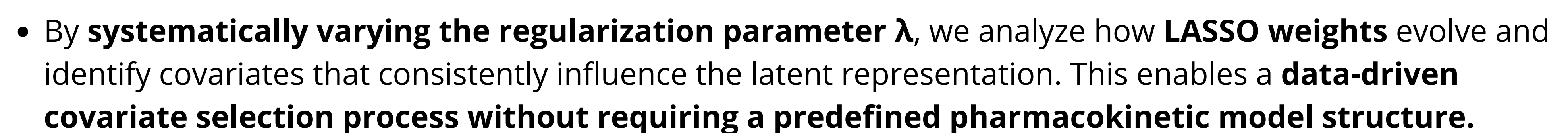
² School of Computer Science and Digital Technologies, London South Bank University, London, United Kingdom (UK)



Dataset

$$C(t) = \frac{D}{V} \cdot \frac{k_a}{k_a - k_e} \cdot \left(e^{-k_e(t-t_D-t_{\text{lag}})} - e^{-k_a(t-t_D-t_{\text{lag}})} \right)$$
$$CL_{ij} = \theta_1 \cdot SNP_{ij}^{\theta_2} \cdot \left(\frac{age_{ij}}{47}\right)^{\theta_3} \cdot \left(\frac{albi_{ij}}{4.1}\right)^{\theta_4} \cdot \left(\frac{hgbi_{ij}}{12.5}\right)^{\theta_5} \cdot e^{\eta_{CL}}$$

Extract informative pharmacokinetic **covariates** by mapping **patient characteristics** to **VAE-derived latent representations** of drug profiles using **LASSO regression**.



- The model architecture employs a **Variational Autoencoder (VAE)** to **learn a structured latent space representation** of the simulated pharmacokinetic (PK) profiles. The encoder maps the input profile to a latent distribution $N(\mu, \sigma)$, from which **latent vectors z are sampled**. The Decoder reconstructs the original PK signal from z . The **LASSO regression model estimate the VAE's latent variables, using as input the patient-specific covariates**. The **L1 penalty** encourages sparsity in the regression coefficients, **forcing non-relevant covariates toward zero**.

- The **VAE reconstructed tacrolimus PK** profiles with high accuracy, achieving a **MAPE of 2.26%** and a **MAE of 0.0009 mg/L** on the test set as can be seen in **Table 1**.

- **LASSO regression** was applied to **predict the VAE's latent variables from covariates**, and covariate weights were analyzed across **λ values ranging from 0.0001 to 1.0**. The model consistently identified **SNP, age, hemoglobin, and albumin** as relevant features (**Fig. 2**), matching the covariates used to determine **inter-individual variability** in the **clearance equation**.
- As **λ increased**, the model **progressively eliminated weak predictors**. **Random noise covariates were effectively discarded at all λ values**. At low regularization ($\lambda \leq 0.005$), additional covariates (e.g., weight, ethnicity) appeared with minor contributions. At $\lambda = 1.0$, as over-penalization case, all covariates were excluded.

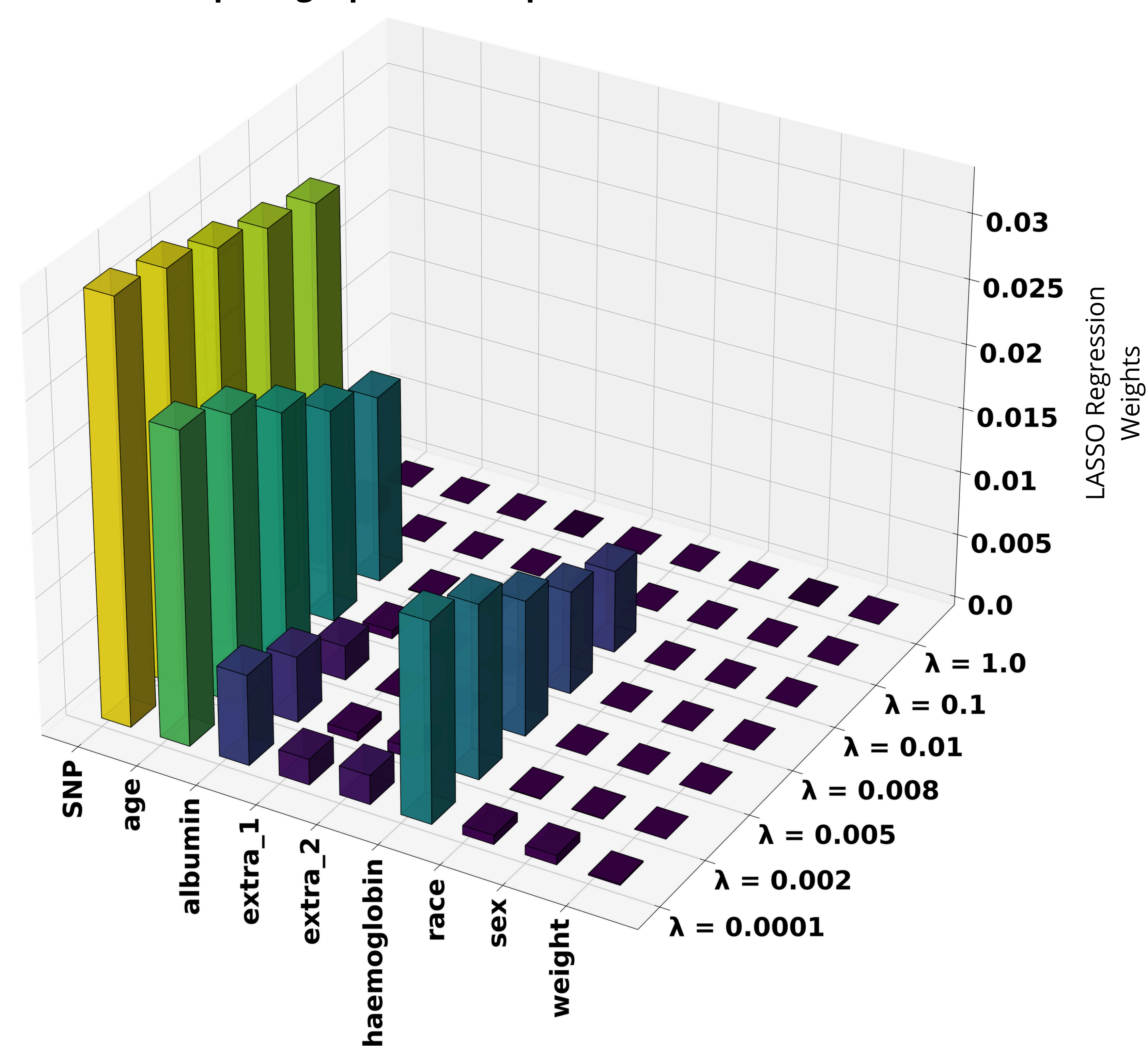


Fig. 2. 3D visualization of LASSO regression weights across covariates and regularization strengths (λ). Bar height and color intensity reflect covariate importance: brighter and taller bars indicate stronger influence. The x-axis shows covariates, the y-axis the λ values, and the z-axis the corresponding LASSO weight.

- The **VAE-LASSO framework** enables **model-free, data-driven covariate selection** in population pharmacokinetics by **linking patient characteristics to latent representations of PK profiles**.
- It reliably identifies clinically relevant covariates while discarding irrelevant ones, **without requiring prior knowledge of the underlying PK model**.
- Compact **latent space representation of PK profiles provides a structured and interpretable foundation for covariate analysis**, supporting precision pharmacotherapy.

Metric	VAE model
MAE (Mean Absolute Error)	0.0009 (mg/L)
MAPE (Mean Absolute Percentage Error)	2.26 %

Table 1. Reconstruction performance of the Variational Autoencoder (VAE) on pharmacokinetic profiles, reported as Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).



Diego Perazzolo: *diego.perazzolo.1@studenti.unipd.it*

Enrico Grisan: *enrico.grisan@lsbu.ac.uk*

Chiara Castellani: chiara.castellani@unipd.it