

# 1 Executive summary

This research investigates the relationship between minimum and maximum temperatures in London, exploring the possibility of a linear dependence. Utilizing data visualization and statistical modeling, we initially identified a strong linear connection between these temperature variables.

Further analysis, incorporating radiation levels, revealed an even better-fit model. We employed JAGS and R to fit and assess the models, ensuring convergence and validating assumptions through MCMC diagnostics and residual analyses. The superior model was selected based on the Deviance Information Criterion (DIC). Our results provide a comprehensive understanding of the linear dependency, contributing valuable insights for temperature prediction.

## 2 Introduction

The research aims to explore the linear dependence between minimum and maximum temperatures in London. This investigation was prompted by initial data visualization suggesting a potential relationship. To address this, we conducted a comprehensive analysis, considering the inclusion of radiation levels to enhance the predictive power of our model. The primary objective is to determine the extent of linear dependence and ascertain whether additional variables improve model accuracy.

## 3 Data

The open dataset is available on kaggle via the URL: <https://www.kaggle.com/datasets/emmanuelfwerr/london-weather-data>. The selected data comprises weather information for London, including such features as «date», «cloud\_cover», «sunshine», «global\_radiation», «max\_temp», «mean\_temp», «min\_temp», «precipitation», «pressure» and «snow\_depth». Graphical exploration in R (via `pairs()`) revealed patterns and trends that guided subsequent modeling decisions — to investigate «min\_temp» and «max\_temp» features (Fig. 1). All the plots, obtained in R, are redrawn in this report by L<sup>A</sup>T<sub>E</sub>X tools: `tikz` and `pgfplots` packages.

## 4 Model

Our statistical model is a hierarchical specification involving minimum and maximum temperatures. Initially, a simple linear model was postulated and justified based on the observed linear relationship in the data.

To enhance the model, radiation levels were incorporated, and the resulting hierarchical specification was fitted using JAGS and R. Prior distributions were chosen to reflect our understanding of the data and ensure model stability. To be short, JAGS model (Table 1) completely reflects an example, given in the course.

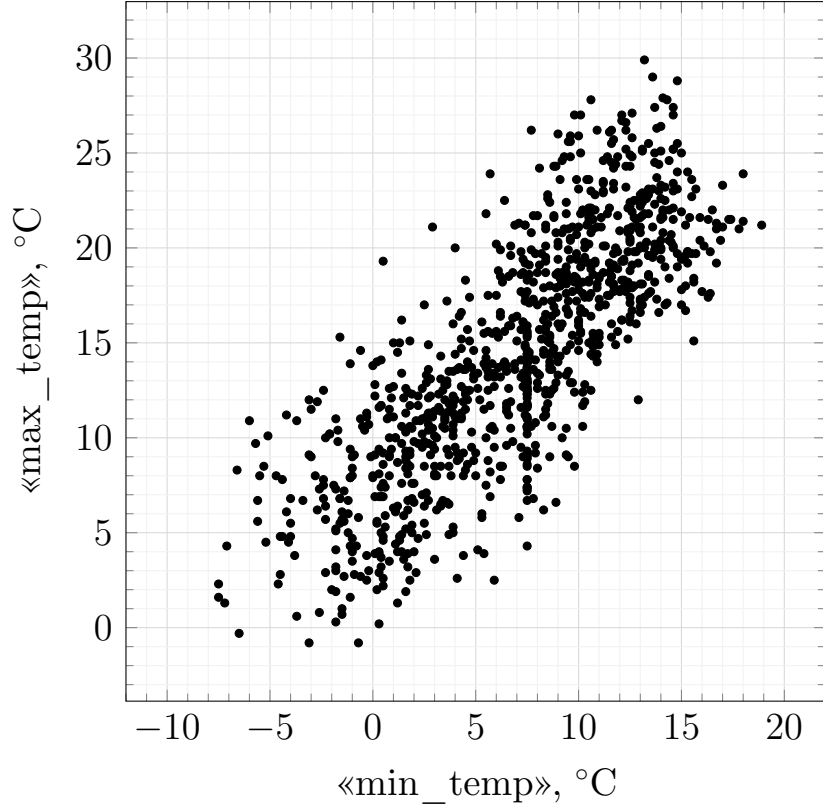


Figure 1: Pairplot of min and max temperatures for 1000 days (Model №1)

	Assumption	DIC
Model №1	$\max\_temp_i \sim N(\mu_i, \sigma^2), i = \overline{1, 1000}$ $\mu_i = b_1 + b_2 \cdot \min\_temp_i$ $b_j \sim N(0, 10^{-6}), j = \overline{1, 2}$ $\sigma^2 \sim Gamma(2.5, 25)$	5496
Model №2	$\max\_temp_i \sim N(\mu_i, \sigma^2), i = \overline{1, 1000}$ $\mu_i = b_1 + b_2 \cdot \min\_temp_i + b_3 \cdot \text{global\_radiation}_i$ $b_j \sim N(0, 10^{-6}), j = \overline{1, 3}$ $\sigma^2 \sim Gamma(2.5, 25)$	5176

Table 1: Models comparison using DIC

MCMC convergence was assessed, and residual analyses confirmed the model’s adequacy (Fig. 2). Model comparison using DIC favored the enhanced model, validating its superiority (Table 1).

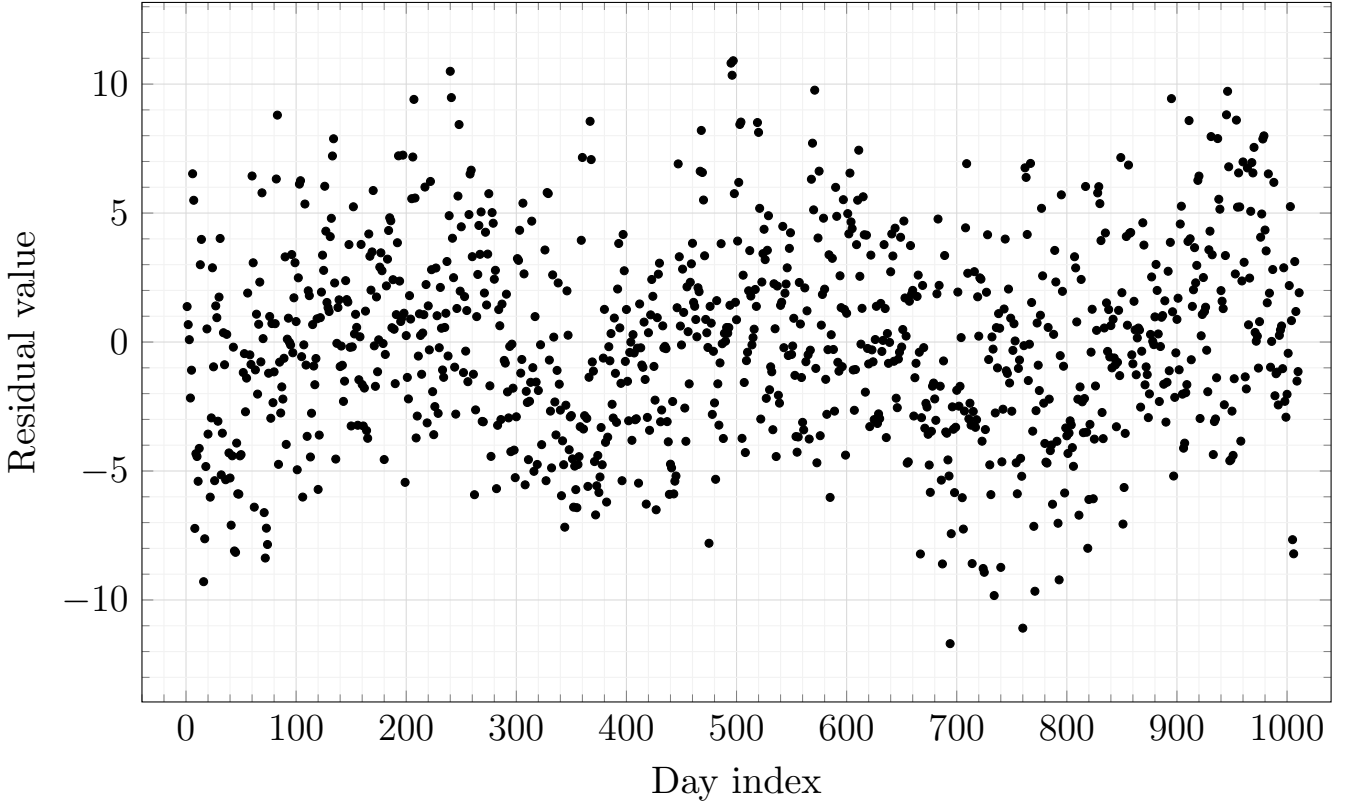


Figure 2: Residual analyses (Model №1)

## 5 Results

The results indicate a robust linear dependence between minimum and maximum temperatures, further improved by incorporating radiation levels. MCMC diagnostics confirm the convergence of the fitted models. The Deviance Information Criterion (DIC) comparison favored the model including radiation levels, signifying its superior predictive performance. Residual analyses revealed no discernible patterns, affirming the model’s adequacy in capturing the linear relationship.

## 6 Conclusions

In conclusion, our research substantiates a strong linear dependence between minimum and maximum temperatures in London. The inclusion of radiation levels significantly enhances the predictive accuracy of the model. Our findings provide valuable insights into temperature patterns, aiding in more accurate predictions.