

1. Linear and polynomial regression

We aim to model daily average temperatures in Linköping using the following quadratic Bayesian regression model:

$$temp = \beta_0 + \beta_1 \cdot time + \beta_2 \cdot time^2 + \epsilon, \quad \epsilon \sim N(0, \sigma^2)$$

Here, *time* is defined as the number of days since the start of the observation period divided by 365, i.e., a scaled time variable ranging from 0 to 1.

We place a conjugate prior on the parameters β and the error variance σ^2 :

$$\beta | \sigma^2 \sim N(\mu_0, \sigma^2 \Omega_0^{-1}), \quad \sigma^2 \sim \text{Inv} - \chi^2(\nu_0, \sigma_0^2)$$

We began with the following prior hyperparameters: $\mu_0 = (0, -100, 100)^T$, $\Omega_0 = 0.01 \cdot I_3$, $\nu_0 = 1$ and $\sigma_0^2 = 1$.

However, when sampling from this prior and visualizing the resulting regression curves, we found that most curves did not align with our prior belief that temperature should vary smoothly within a realistic range (-20°C to +30°C) over the course of the year. Many curves were too extreme due to overly vague prior variances.

To better reflect plausible seasonal temperature patterns, we adjusted the priors to: $\mu_0 = (20, -100, 100)$ (reflecting a mean intercept of 20°C) and $\Omega_0 = 50 \cdot I_3$ (wider spread but still regularizing). These new priors produce more reasonable prior predictive regression curves.

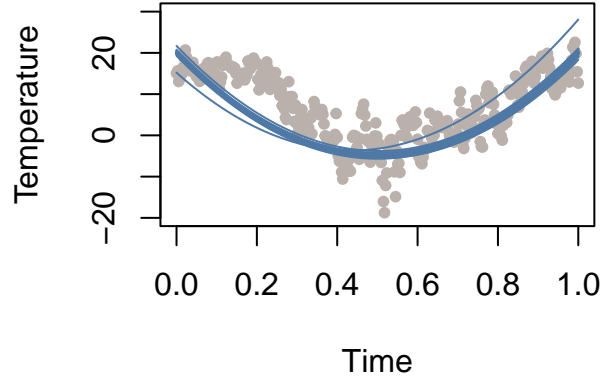


Figure 1: Prior Predictive Regression Curves

The prior predictive curves reflect a wide but plausible range of behaviors for seasonal temperature trends in Linköping. Most curves follow a quadratic seasonal pattern, peaking or dipping around the center of the time range, consistent with expected annual temperature fluctuations. This confirms that the revised priors are now better aligned with prior beliefs.

We now aim to simulate draws from the joint posterior distribution of the model parameters $\beta_0, \beta_1, \beta_2$ and σ^2 . Based on conjugate Bayesian linear regression theory, the posterior distributions are given by:

$$\beta|\sigma^2 \sim N(\mu_n, \sigma^2 \Omega_n^{-1}), \quad \sigma^2|y \sim \text{Inv} - \chi^2(\nu_n, \sigma_n^2)$$

with updated hyperparameters:

$$\mu_n = (X'X + \Omega_0)^{-1}(X'X\hat{\beta} + \Omega_0\mu_0)$$

$$\Omega_n = X'X + \Omega_0$$

$$\nu_n = \nu_0 + n$$

$$\sigma_n^2 = \frac{1}{\nu_n}(\nu_0\sigma_0^2 + (y'y + \mu_0'\Omega_0\mu_0 + \mu_n'\Omega_n\mu_n))$$

We simulate from the posterior and visualize the marginal distributions of each parameter:

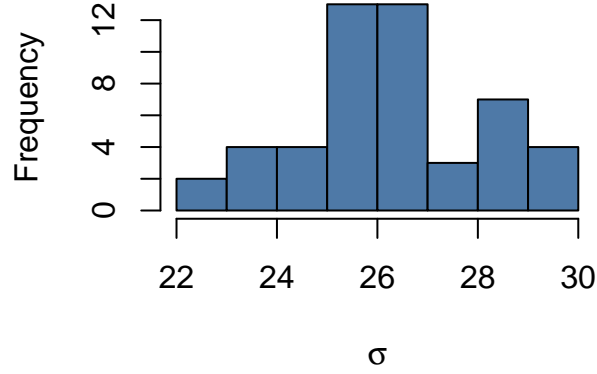


Figure 2: Marginal posterior of σ

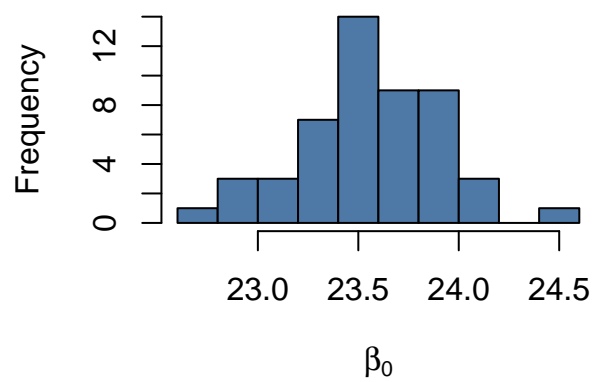


Figure 3: Marginal posterior of β_0

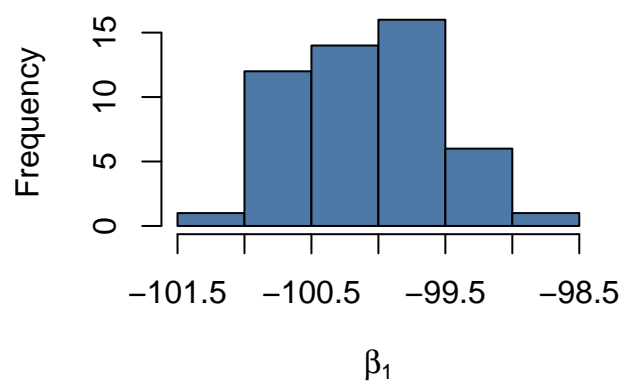


Figure 4: Marginal posterior of β_1

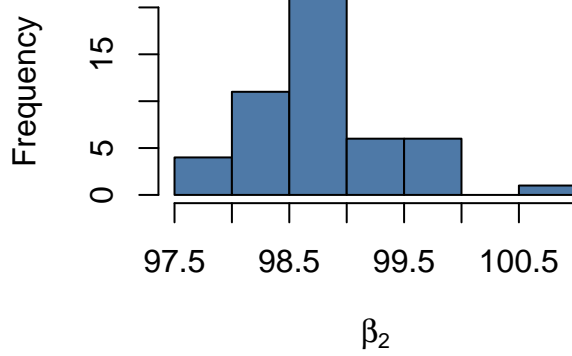


Figure 5: Marginal posterior of β_2

To summarize the uncertainty in the regression function $f(\text{time}) = E[\text{temp}|\text{time}] = \beta_0 + \beta_1 \cdot \text{time} + \beta_2 \cdot \text{time}^2$, we compute the posterior median curve for $f(\text{time})$ and the 90% equal-tail credible interval.

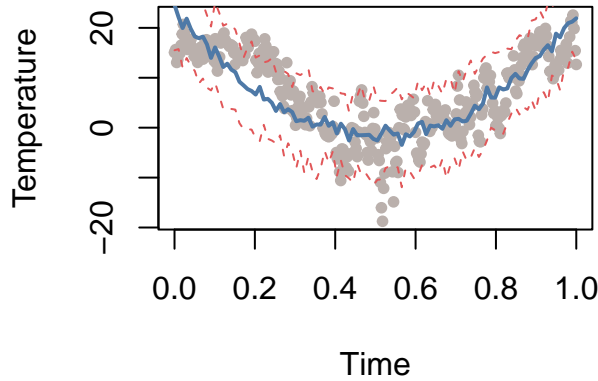


Figure 6: Posterior Regression Curve with 90% Credible Intervals

The posterior median curve closely follows the seasonal temperature pattern suggested by the data. The 90% credible intervals cover most observed data points, as expected. Since the regression model accounts for observation noise ($\epsilon \sim N(0, \sigma^2)$) we expect a large proportion of data points to fall within the interval—but not necessarily all. Roughly 90% of them should, under the assumption that the model is well-calibrated.

We are interested in identifying the time \tilde{x} at which the expected temperature reaches its minimum. Since the regression function is a quadratic polynomial, the minimum occurs at

$$\tilde{x} = -\frac{\beta_1}{2\beta_2}$$

Using the posterior draws of β_1 and β_2 obtained previously, we simulate the posterior distribution of \tilde{x} .

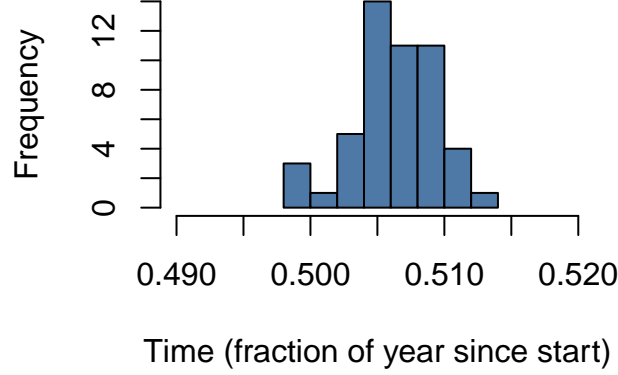


Figure 7: Posterior Distribution of \tilde{x}

As shown in the histogram, the posterior distribution of \tilde{x} is concentrated around the middle of the winter period (January), which aligns well with prior expectations for the coldest time of year in Linköping. The narrow distribution suggests high certainty in the estimate, given the data.

We now estimate a polynomial regression of order 10. However, since higher-order polynomial terms may not contribute significantly to explaining the data—and could even lead to overfitting—we choose a shrinkage prior to regularize the model and mitigate this risk.

We again use the conjugate prior:

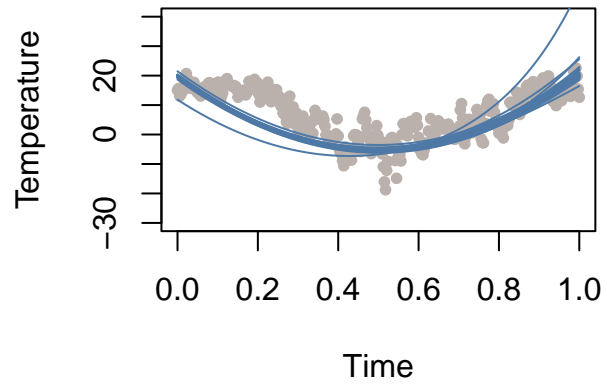
$$\beta|\sigma^2 \sim N(\mu_0, \sigma^2 \Omega_0^{-1}), \quad \sigma^2 \sim \text{Inv} - \chi^2(\nu_0, \sigma_0^2)$$

To reflect our belief that only the lower-order terms (e.g., constant, linear, quadratic) are likely to be influential, we set: $\mu_0 = (20, -100, 100, 0, 0, 0, 0, 0, 0, 0)^T$, $\Omega_0 = 50 \cdot I_1 1$, $\nu_0 = 1$ and $\sigma_0^2 = 1$.

This prior centers the first three coefficients on meaningful values while shrinking the remaining higher-order terms toward zero, discouraging unnecessary model complexity.

We now simulate from the joint prior to generate prior predictive regression curves:

Prior Predictive Regression Curve:



As shown in the plot, the shrinkage prior successfully constrains the model's flexibility. Most curves resemble smooth, low-degree polynomials, and avoid the erratic behavior typically associated with unregularized high-order models. This behavior is consistent with our prior belief that only the lower-degree terms are essential for capturing the seasonal temperature trend.