

Identifying systematic bias in model predictions



Shamin Saffian¹, Daniel Wright², Stephen Duffull²



Faculty of Pharmacy¹, The National University of Malaysia, Kuala Lumpur, Malaysia
School of Pharmacy², University of Otago, Dunedin, New Zealand

Background

- An assessment of predictive performance is required to demonstrate that a model is fit for purpose and has utility in the intended setting
- For non-repeated measures data, we expect a linear relationship between the predicted and observed data
- Sheiner and Beal [1] proposed a method for assessing predictive performance using mean prediction error (MPE)
- MPE may not be sensitive to bias patterns that display a systematic trend over the range of the observations

Aim

- To illustrate a method for detecting trends in bias patterns from model predictions

Methods

- The proposed method is a generalised form of MPE. It was developed and evaluated in three steps;

1. Binning and MPE slope;

- n bins were created and MPE determined for each
- MPE was regressed against the bin number;

$$MPE(k) = \beta_0 + \beta_1 \times bin_k;$$

where β_1 is the slope, β_0 is the intercept, k is the bin number in the n th iteration

Let $n = \{2, \dots, N_{max}\}$, let $k = \{1, \dots, n\}$ where N_{max} is the total possible bins

- Binning repeated until $n=N_{max}$.

2. Asymptotic slope (S_∞);

- Slope values (β_1) were regressed against the bin number
- The asymptotic slope (at infinite bins) was given by;

$$\beta_1 = S_\infty + (S_0 - S_\infty)(\exp^{-\theta \times nj}); j \in \{2, \dots, N_{max}\}$$

where S_∞ is the slope at infinite bins, S_0 is the initial slope, θ is a rate constant, and j is the number of bin

- If the 95% CI of S_∞ included zero, then no bias was concluded

3. Evaluation

- The method was evaluated using six different patterns of bias including bidirectional, off-set, non-linear and non-monotonic
- Data sets of $n=100$ were simulated using MATLAB
- MPE and S_∞ were determined for each scenario
- Random noise was simulated assuming; $\varepsilon_i \sim N(0, \sigma^2)$ where ε_i is the random error of the i th data point and σ^2 set to 0.25

Results

- Unlike MPE, our method correctly detected systematic bias in Scenarios C, D, and E
- MPE detected a constant positive bias (Scenario B)
- Neither method could detect curvilinear bias (Scenario F)

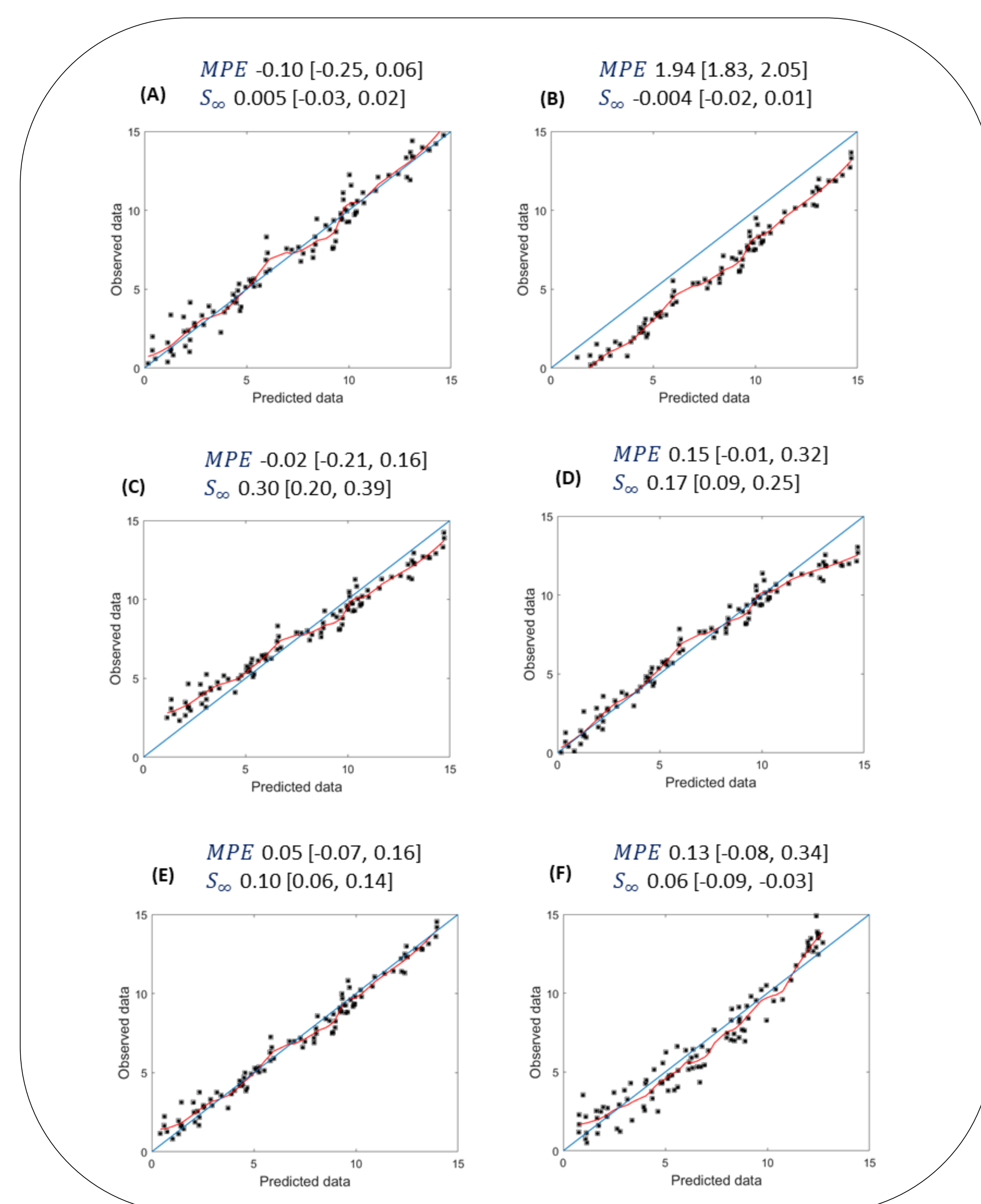


Figure 1. Simulated bias scenarios including MPE [95% CI] and asymptotic slope (S_∞) [95% CI]. Scenarios as follows;

- Unbiased predictions (control)
- Linear, off-set predictions (shifted below the line of identity)
- Linear, bidirectional, bias
- Non-linear, unidirectional bias at larger observations
- Non-linear, unidirectional bias at smaller observations
- U-shaped, non-monotonic, bias

Conclusions

- A method for numerically detecting systematic deviation in model-predictions has been proposed
- It provides an additional interpretation to the MPE approach
- The method has similarities to a LOESS regression but provides numerical and statistical quantities associated with any deviation
- The method could be automated and implemented in statistics software programmes

Reference

1. Sheiner & Beal. *J Pharmacokinet Biopharm* 1981 9:503-512