

# New Emulators for Monthly Gas Consumption

This document is a concise summary of the construction and validation of the monthly gas consumption emulators. The 1000 design runs are divided into a group of 800 used for training, and a group of 200 used for validation. For reference, [Table 1](#) lists the meaning and range of the 8 inputs which are varied from one simulation to the other.

## 1 Emulators' Training

In order to emulate the monthly gas consumption, the range of each of the eight input variables  $V_1, \dots, V_8$  has been linearly rescaled into  $[-1, 1]$ . Given the outputs  $y_i$  of the 800 training runs for a month, I have first built a linear regression model of  $y_i$  as a function of (some of the) linear, quadratic and interaction terms of the  $V_i$ . Hence, I have built an emulator of the linear regression residuals. Details of the two steps follow.

### Linear Regression

Let  $\mathcal{R}$  be the set of all linear, quadratic and interaction terms of the 8 inputs  $V_1, \dots, V_8$ , all mutually orthogonal. As it turns out, for all months, the set of the “best” 10 regressors in  $\mathcal{R}$  (the set maximising the adjusted  $R^2$ ) explains the data extremely well, with an adj- $R^2$  of about 0.999 in all cases. Given the excellent result, and to use an approach which is uniform among months and simple to explain, for each month I have built a linear regression model by selecting the best 10 regressors according to the above criterion. *(previously, for May, I was using some 3rd-order terms to improve the linear model. Probably not worth it)*

**Table 1:** Meaning and range of the eight parameters varied among the  $n$  simulations of this work. (*Mohammad may provide more appropriate descriptions for the middle column.*)

Short Name	Physical Meaning	Range
$V_1$	Heating Setpoint	$[17.5, 20.5]$ °C
$V_2$	Boiler Efficiency	$[0.6, 0.75]$
$V_3$	External wall thickness	$[4, 6.3]$ cm
$V_4$	Roof quilt thickness	$[15, 21]$ cm
$V_5$	Floor insulation thickness	$[4.5, 5.5]$ cm
$V_6$	Infiltration rate	$[0.2, 0.95]$ ac/h
$V_7$	DHW Consumption	$[6.15, 22] \times 10^{-6}$ litre/day
$V_8$	Cooking	$[1.05, 6.3]$ W/m <sup>2</sup>

Table 2 shows which 10 regressors have been selected for each month, the adjusted  $R^2$  of the associated linear model, and the variance of the residuals.

## Bayes Linear Emulators of Residuals

In order to build an emulator of the residuals, choices about the following are to be made: active inputs, correlation lengths, prior emulator variance and nugget variance.

To choose these, for each month, I have followed the general procedure which I highlight below. However, some of the choices necessarily follow from a considerable and necessarily subjective work of “try  $\rightleftharpoons$  validate, then make a decision”, rather than from the “blind” application of a single rule. Hope this still sounds good for the paper, one we should validation plots? (details in the next section)

- **Active Inputs:** To identify which inputs were significant in explaining the residuals, I have looked at the best selection of second- (and sometimes third-) order terms in a linear model. Terms which would appear alone with a significant  $t$ -value ( $t > 8$ ) would generally be included as an active input, while the inclusion of

**Table 2:** For each of the nine months considered, from left to right: the set of covariates used to build the linear regression model, the adjusted  $R^2$  of the model, the variance of the residuals. In the second column, the “\*” symbol is used as shorthand notation to denote all linear and interaction terms, *e.g.*  $a * b * c = \{a, b, c, ab, ac, bc\}$ .

Month	Regressors	Adj. $R^2$	Resid. Var. $\sigma^2$
Jan	$V_1 * V_2 * V_6, V_3, V_4, V_2^2, V_6^2$	0.9998	84.86
Feb	$V_1 * V_2 * V_6, V_3, V_4, V_2^2, V_6^2$	0.9998	50.63
Mar	$V_1 * V_2 * V_6, V_3, V_4, V_2^2, V_6^2$	0.9997	49.14
Apr	$V_1 * V_2 * V_6, V_3, V_4, V_2^2, V_6^2$	0.9998	49.58
May	$V_1 * V_2 * V_6, V_3, V_8, V_1^2, V_6^2$	0.9989	206.70
Sep	$V_1 * V_2 * V_6, V_3, V_4, V_8, V_6^2$	0.9995	0.88
Oct	$V_1 * V_2 * V_6, V_3, V_4, V_8, V_6^2$	0.9996	23.20
Nov	$V_1 * V_2 * V_6, V_3, V_4, V_2^2, V_6^2$	0.9998	57.22
Dec	$V_1 * V_2 * V_6, V_3, V_4, V_2^2, V_6^2$	0.9998	67.65

terms appearing only in interaction with other terms, or with a low  $t$ -value would be considered case by case, with a process of trial and error.

- **Correlation lengths:** Once the active inputs have been chosen, for simplicity I have used the same correlation length for all of them. The correlation length has been chosen by a repeated trial-and-validation procedure, adjusted to the number of active inputs used (if more active inputs are used, higher correlation lengths are required to maintain a similar level of correlation across the input space).
- **Prior Variances:** The cumulative prior variance (the one of the second-order stochastic process over which we make specifications, and the one of the “nugget” process) has been imposed to be the variance of the residuals which are being fitted. This overall variance has been split into the two components in the proportions of 95% and 5% respectively.

Table 3 shows the choices of the above quantities for the nine months of interest. [Figure 1](#)

**Table 3:** For each of the nine months considered, the table below shows which active inputs have been chosen and the value of the correlation length (same for all active inputs).

Month	Active Inputs	Corr. Length
Jan	$V_1, V_2, V_3, V_6, V_8$	0.8
Feb	$V_1, V_2, V_3, V_7, V_8$	0.5
Mar	$V_1, V_2, V_3, V_6, V_7, V_8$	1.2
Apr	$V_1, V_2, V_3, V_6, V_8$	0.7
May	$V_1, V_2, V_3, V_4, V_6, V_8$	1.2
Sep	$V_1, V_2, V_3, V_6, V_7, V_8$	1.0
Oct	$V_1, V_2, V_3, V_6, V_7, V_8$	1.2
Nov	$V_1, V_2, V_3, V_6, V_7, V_8$	1.2
Dec	$V_1, V_2, V_3, V_6, V_7, V_8$	1.3

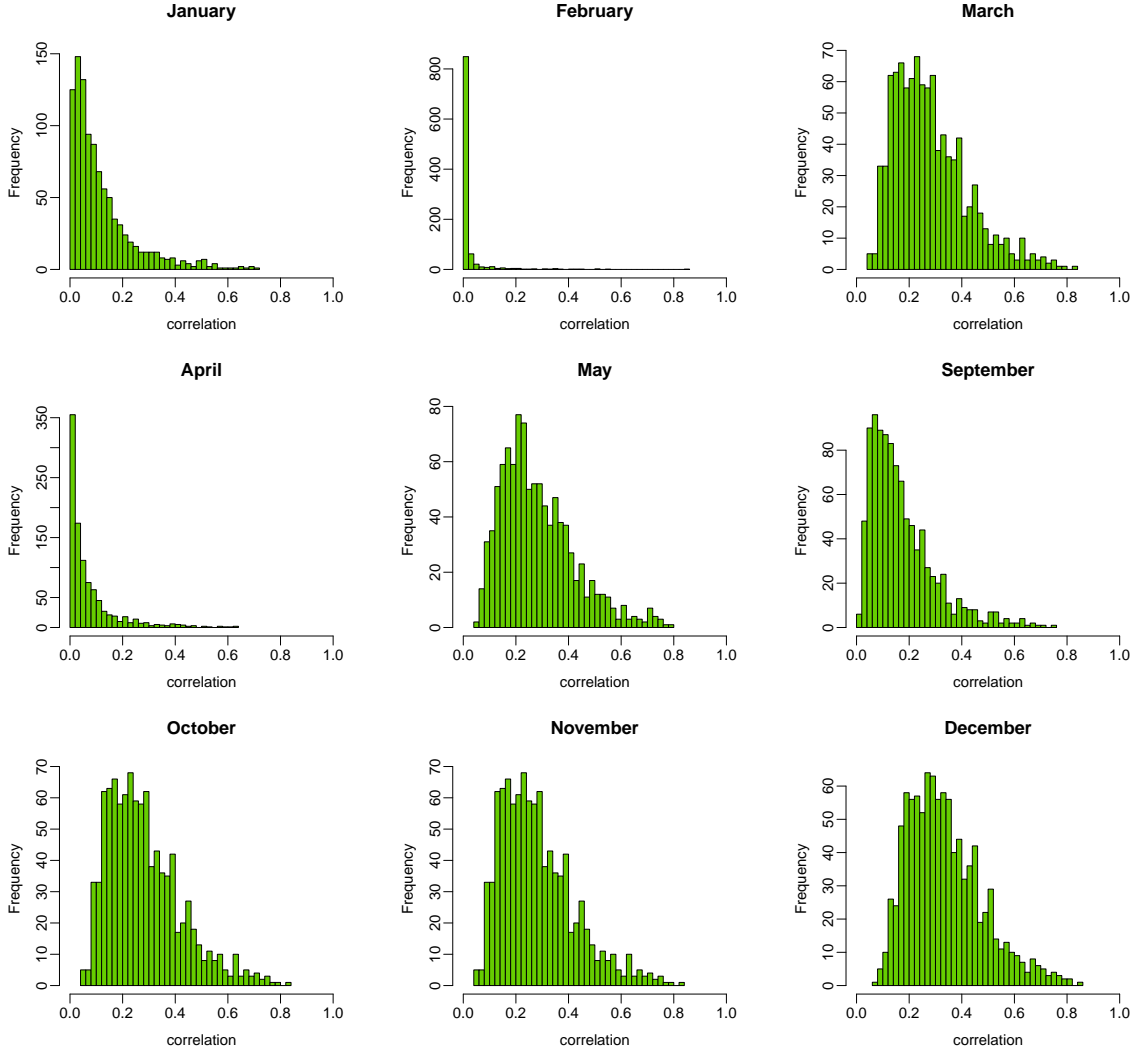
instead gives an idea of the level of correlation across the input space induced by the above choices. Specifically, called  $P$  the point at the centre of the hypercube with edges  $[-1, 1]$ , the figure shows histograms of the correlation between the point  $P$  and the 1000 design points.

## 2 Emulators' Validation and Exploration

This section shows some plots concerning the validation of the previously built emulators. The emulators are validated on the results of 200 simulations not used for training.

Figure 2 shows the standardised errors on the 200 validation point. For each of the 200 points, if  $y_i$  is the simulator output,  $\hat{y}_i$  the emulator prediction and  $\hat{\sigma}_i$  the emulator standard deviation, the standardised error is computed as follows:

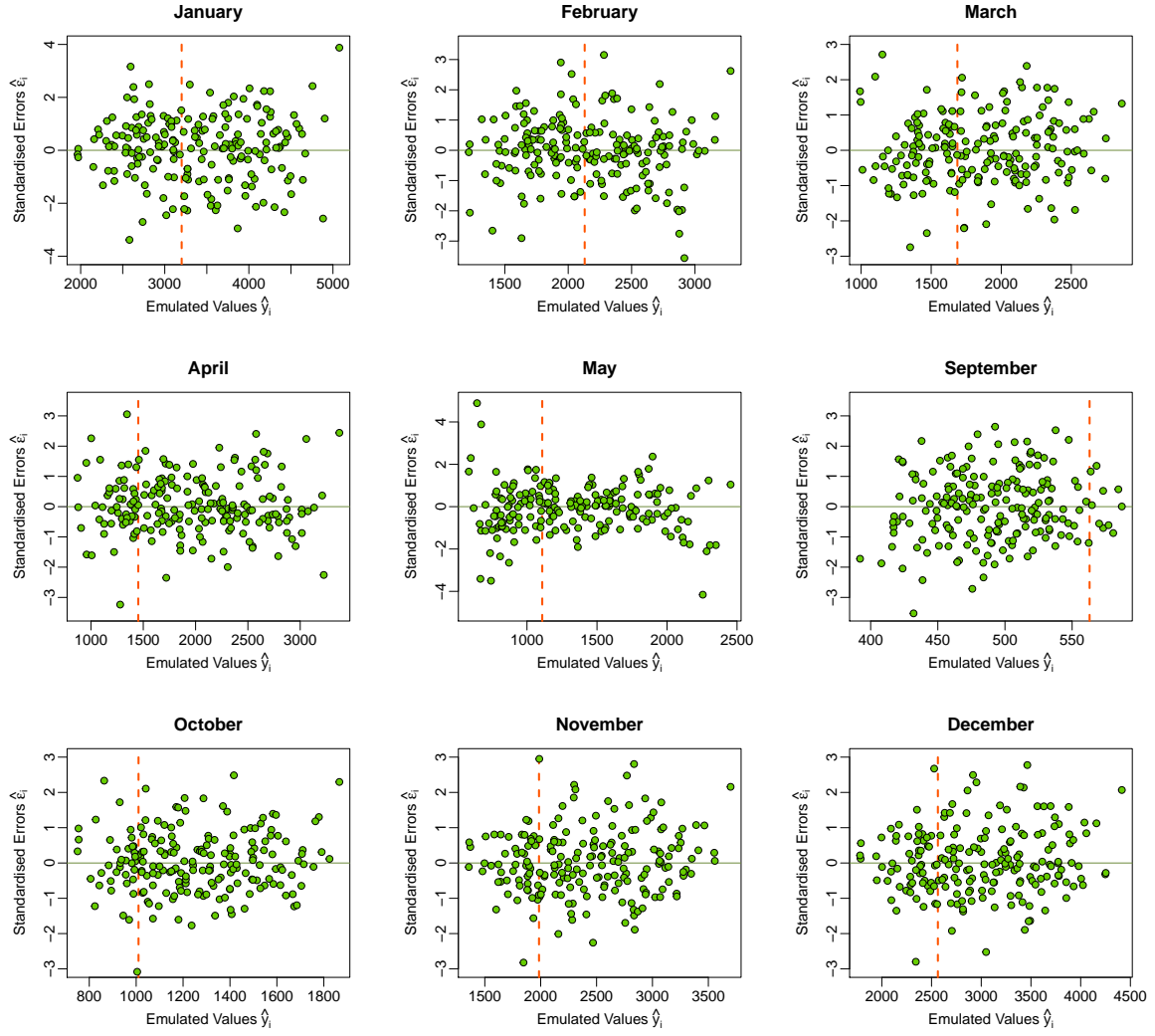
$$\hat{\varepsilon}_i = \frac{y_i - \hat{y}_i}{\hat{\sigma}_i}. \quad (1)$$



**Figure 1:** Histograms of the correlations between the point at the center of the hypercube  $[-1, 1]^N$  (where, for each month,  $N$  is the number of active inputs) and the 1000 design points.

## 2.1 Comparison Between LR and Trained Emulators

For each month, [Table 4](#) compares the prior variance of the emulator (*i.e.*, the variance of the original linear regression residuals) with the adjusted variance of the emulator predictions on the 200 validation points. I am reporting the 2.5 and 97.5 empirical percentiles of these, to get a sense of the difference in prediction uncertainty between the original linear regression and the trained emulators.



**Figure 2:** Validation of the emulators on a set of 200 points. For the different months, the standardised errors in (1) are plotted against the emulator’s fitted values. In each plot, the vertical dashed line identifies the value on the  $x$ -axis corresponding to the observed gas consumption for the month.

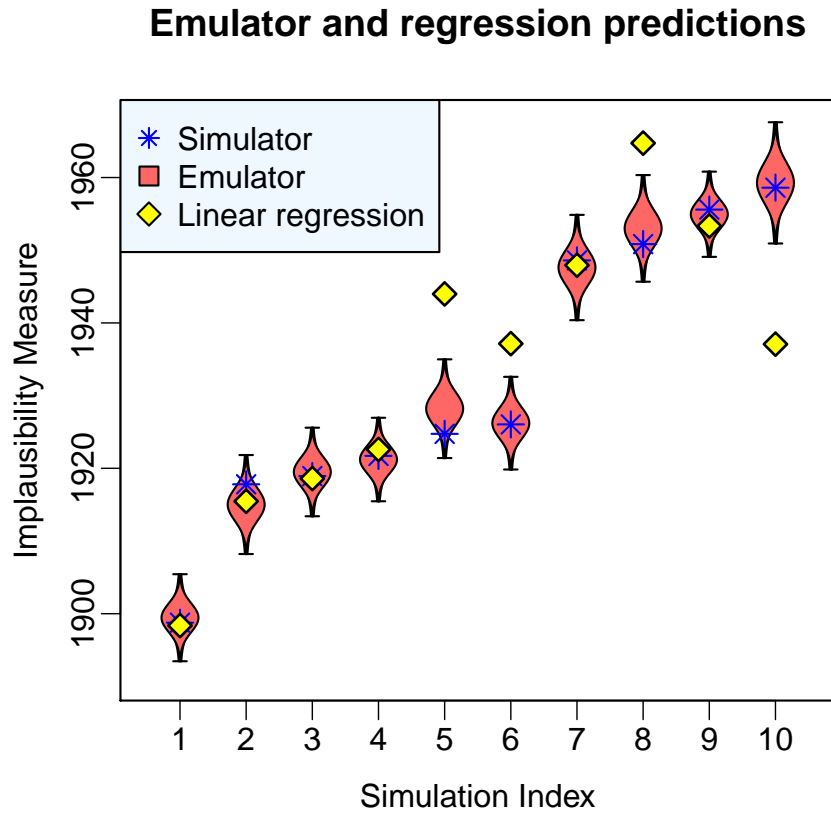
Finally, [Figure 3](#) allows to compare graphically the accuracy of the emulator’s predictions versus the linear regression’s ones. The plot concerns the month of March, and only results for 10 of the 200 validation points are shown (10 consecutive ones in terms of the simulator outputs), to avoid cluttering which would prevent details to be appreciated.

The red “cat’s eye” boxes represent the emulator predictions for these points, within

**Table 4:** For each month: the first line of the table shows the variance of the linear regression residuals (prior cumulative variance of the emulator). The second and third lines show the 2.5 and 97.5 empirical percentiles of the emulator variances at the 200 validation points.

	Jan	Feb	Mar	Apr	May	Sep	Oct	Nov	Dec
$\sigma^2$	84.86	50.63	49.14	49.58	206.70	0.88	23.20	57.22	67.65
[2.5%,	6.4	10.6	3.2	4.5	13.3	0.07	1.51	3.7	4.17
97.5%] CI	29.1	43.4	11.5	22.6	53.8	0.35	5.41	13.35	12.67

a range of three standard deviations from the mean prediction. The blue star and the yellow squares represent instead the real simulator outputs and the linear regression predictions, respectively. It can be appreciated that the emulator is often much more precise than the linear model, and has a high level of confidence in its prediction (small uncertainty bands).



**Figure 3:** Visual comparison between emulator (red cat's eye boxes) and linear regression (yellow squares) predictions. The plot concerns 10 validation points of the month of March. The height of the emulator cat's eye boxes cover a range of three standard deviations from the emulator's mean prediction.