# Read-me file for the R code producing time series forecasts using Moore's law and Wright's law*

François Lafond

University of Oxford

**Abstract**

This note accompanies the R code that produces time series forecasts according to the methods developed by Farmer & Lafond (2016) and Lafond et al. (2018).

## Contents



*Contact: francois.lafond@inet.ox.ac.uk

# 1 Methods

In Farmer & Lafond (2016) and Lafond et al. (2018), we devised methods to produce time series forecasts. This note and the accompanying code gives details on how to generate the density forecasts, or the point forecasts and prediction intervals.

Throughout, we assume that there is no missing data. The note follows the notation from Lafond et al. (2018).

## 1.1 Moore's law

**Model.** The dependent variable (cost) is modelled as a geometric random walk with drift

$$y_t = y_{t-1} + \mu + n_t, \tag{1}$$

where $y \equiv \log \text{cost}$, and the noise follows a MA(1) process

$$n_t = v_t + \theta v_{t-1},$$

$$v_t \sim \mathcal{N}(0, \sigma_v^2).$$

It is helpful to define a shorthand for the first-differenced variables

$$Y_t = y_t - y_{t-1}.$$

**Parameter estimation.** Farmer & Lafond (2016) estimate the parameters

$$\hat{\mu} = \frac{1}{m} \sum_{i=2}^{m+1} Y_i = \frac{y_{m+1} - y_1}{m} \tag{2}$$

and

$$\hat{K}^2 = \frac{1}{m-1} \sum_{i=2}^{m+1} (Y_i - \hat{\mu})^2, \tag{3}$$

where $m + 1 = T$ is the number of available years, so that $m$ is the number of available growth rates (first differences of the log).

The parameter $-1 < \theta < 1$ is calibrated separately, and Farmer & Lafond (2016) suggest to use the value

$$\theta = 0.63.$$

**Forecasts.** Denoting by $y_T$ the last in-sample observation, the predictions at horizon $\tau$ are

$$\hat{y}_{T+\tau} \sim \mathcal{N}\Big(E[\hat{y}_{T+\tau}], \text{Var}[\hat{y}_{T+\tau}]\Big), \tag{4}$$

where $\sim \mathcal{N}$ means that the random variable follows a normal distribution. Its parameters are

$$E[\hat{y}_{T+\tau}] = y_T + \hat{\mu}\tau, \tag{5}$$

and

$$\mathrm{Var}[\hat{y}_{T+\tau}] = \frac{\hat{K}^2\Big[-2\theta + \Big(1 + \frac{2(m-1)\theta}{m} + \theta^2\Big)\Big(\tau + \frac{\tau^2}{m}\Big)\Big]}{1+\theta^2}. \tag{6}$$

**Comments.** Note that if $\theta = 0$,

$$\mathrm{Var}[\hat{y}_{T+\tau}] = \hat{K}^2\Big(\tau + \frac{\tau^2}{m}\Big)$$

To understand the effect of $\theta$, it is helpful to simplify the formula using the assumptions that $m$ is large enough, so $m-1 \approx m$, and $\tau$ is large enough, so that the term $-2\theta$ is negligible as compared to the term that includes the factor $(\tau + \tau^2/m)$ (i.e.$-2\theta \approx 0$). Then,

$$\mathrm{Var}[\hat{y}_{T+\tau}] \approx \frac{(1+\theta)^2}{1+\theta^2}\hat{K}^2\Big(\tau + \frac{\tau^2}{m}\Big).$$

In the relevant case in Farmer & Lafond (2016) ($0 < \theta < 1$), the prefactor involving $\theta$ is a monotonic increasing and concave function of $\theta$. It varies from 1 to 2 as $\theta$ varies from 0 to 1. Thus, increasing $\theta$ increases the prediction intervals. In practice, this is how Farmer & Lafond (2016) adjusted the null model so that it would imply theoretical prediction intervals large enough for the empirically measured forecast errors.

## 1.2 Wright's law

**Model.** Wright's law is a simple linear relationship between the log of costs and the log of experience. Lafond et al. (2018) propose a version in first differences,

$$y_t - y_{t-1} = \omega(x_t - x_{t-1}) + \eta_t, \tag{7}$$

where $y \equiv \log \mathrm{cost}$, $x \equiv \log \mathrm{experience}$ and the noise follows a MA(1) process

$$\eta_t = u_t + \rho u_{t-1},$$

$$u_t \sim \mathcal{N}(0, \sigma_u^2).$$

It is helpful to define a shorthand for the first-differenced variables,

$$X_t = x_t - x_{t-1},$$

$$Y_t = y_t - y_{t-1}.$$

**Parameter estimation.** Lafond et al. (2018) estimate two parameters using simple OLS on Eq. 7,

$$\hat{\omega} = \frac{\sum_{i=2}^{m+1} X_i Y_i}{\sum_{i=2}^{m+1} X_i^2}, \tag{8}$$

$$\hat{\sigma}_\eta^2 = \frac{1}{m-1}\sum_{i=2}^{m+1}(Y_i - \hat{\omega}X_i)^2. \tag{9}$$

The parameter $-1 < \rho < 1$ is estimated separately, and Lafond et al. (2018) suggest to use the universal value

$$\rho = 0.19.$$

**Forecasts.** To make predictions, we assume that we have a series of non-stochastic future values for cumulative production or its the growth rates $\{X_{T+1} \dots X_{T+\tau}\}$. The predictions then are

$$\hat{y}_{T+\tau} \sim \mathcal{N}\Big(E[\hat{y}_{T+\tau}], \mathrm{Var}[\hat{y}_{T+\tau}]\Big), \tag{10}$$

with

$$E[\hat{y}_{T+\tau}] = y_T + \hat{\omega}\sum_{s=1}^{\tau} X_{T+s} = y_T + \hat{\omega}(x_{T+\tau} - x_T), \tag{11}$$

and

$$V(y_{T+\tau}) = \frac{\hat{\sigma}_\eta^2\Big[\rho^2 H_2^2 + \sum_{j=2}^{T-1}(H_j + \rho H_{j+1})^2 + (\rho + H_T)^2 + (\tau-1)(1+\rho)^2 + 1\Big]}{1+\rho^2}, \tag{12}$$

where

$$H_j = -\frac{\sum_{i=T+1}^{T+\tau} X_i}{\sum_{i=2}^{T} X_i^2} X_j.$$

**Comments.** Note that if $\rho = 0$,

$$\mathrm{Var}[\hat{y}_{T+\tau}] = \hat{\sigma}_\eta^2\left(\tau + \frac{\left(\sum_{i=T+1}^{T+\tau} X_i\right)^2}{\sum_{i=2}^{T} X_i^2}\right). \tag{13}$$

Note that if $X_i = r \; \forall i \in \{2 \dots T \dots T+\tau\}$, Wright's law is observationally equivalent to Moore's law and we recover the same formulas.

# 2 Implementation

The code is implemented in R (R Core Team 2017).

There is one function for each model (`MoorePredParam` and `WrightPredParam`), to produce the parameters of the forecasts, that is, the mean and variance of the distribution. A third function, `GetPreds`, takes the output of either `MoorePredParam` or `WrightPredParam`, and transforms the parameters into point forecasts and prediction intervals over the desired forecast horizon.

## 2.1 Moore's law

The function `MoorePredParam(DF, TAU, theta=0.63)` has the following inputs and output.

#### Inputs

1. A numeric vector, the historical values of the time series to be predicted.
2. A numeric scalar, the prediction horizon.
3. Optionally, a numeric scalar, the value of $\theta$ to override the default $\theta = 0.63$.

**Output.** A data.frame, with one line per future period, and two columns:

- Column 1: Values of the point forecasts for the log variable (Equation 5)
- Column 2: Values of the variance of the point forecasts (Equation 6)

## 2.2 Wright's law

The function `WrightPredParam(DF, e.future, rho=0.19)` has the following inputs and output:

#### Inputs

1. A data.frame with two columns, each with numeric values:
   - Column 1: Values of dependent variable (e.g. cost).
   - Column 2: Values of independent variable (e.g. experience).
2. A numeric vector, the future values of the independent variable.

**Output.** A data.frame, with one line per future period, and two columns:

- Column 1: Values of the point forecasts for the log variable (Equation 11).
- Column 2: Values of the variance of the point forecasts (Equation 12).

## 2.3 Transforming forecasts parameters into point forecasts and prediction intervals

**Inputs.**

1. The output of either `MoorePredParam`, or `WrightPredParam`.
2. A probability for the prediction interval, e.g. 0.95 for 95% prediction intervals.

**Output.** Point forecasts ($\equiv$ median $\equiv$ mean of the log), lower bound, and upper bound. These are based on the Normal distribution, see Eqs. 4 and 10.

# 3 Example

We reproduce here Fig. 13 in Lafond et al. (2018), which includes forecasts from both Moore's law and Wright's law. The parameters used in the paper are $\tau = 9$, $\theta = 0.23$, $\rho = 0.19$. The prediction intervals are 95%.[1]

We load a dataframe called `DF` with costs in the first column, and experience in the second column.

**Moore's law.** The first function retrieves the mean and variance of the forecasts

```
pred_param_Moore<-MoorePredParam(DF[,1],tau=9,theta=0.23)
```

The output looks like

| | E(y) | Var(y) |
|---|---|---|
| 1 | -0.76 | 0.02 |
| 2 | -0.88 | 0.06 |
| ... | ... | ... |
| 9 | -1.73 | 0.36 |

The results are then passed to the second function

```
pred_pi_Moore<-GetPreds(pred_param_Moore,0.95)
```

The output looks like

| | point_forecast | lower_bound | upper_bound |
|---|---|---|---|
| 1 | 0.47 | 0.35 | 0.64 |
| 2 | 0.42 | 0.26 | 0.67 |
| ... | ... | ... | ... |
| 9 | 0.18 | 0.05 | 0.58 |

**Wright's law.** The procedure is the same as for Moore's law, except that one first needs to produce a series of future values of cumulative production. Fig. 13 in Lafond et al. (2018) is based on the assumption that experience will keep growing at the historical rate $r$, which is

```
r<-mean(diff(log(DF[,2])))
E.FUTURE<-DF[TT,2]*exp(r*(1:tau))
```

[1] Lafond et al. (2018) use $\pm 2$ standard deviations, but the chart below uses the more precise values $\pm 1.959\ldots$, based on exact quantiles of the normal distribution.

The two functions are then used as for Moore's law

```
pred_param_Wright<-WrightPredParam(DF,E.FUTURE,rho=0.19)
pred_pi_Wright<-GetPreds(pred_param_Wright,0.95)
```

Fig. 1 shows the results[2].

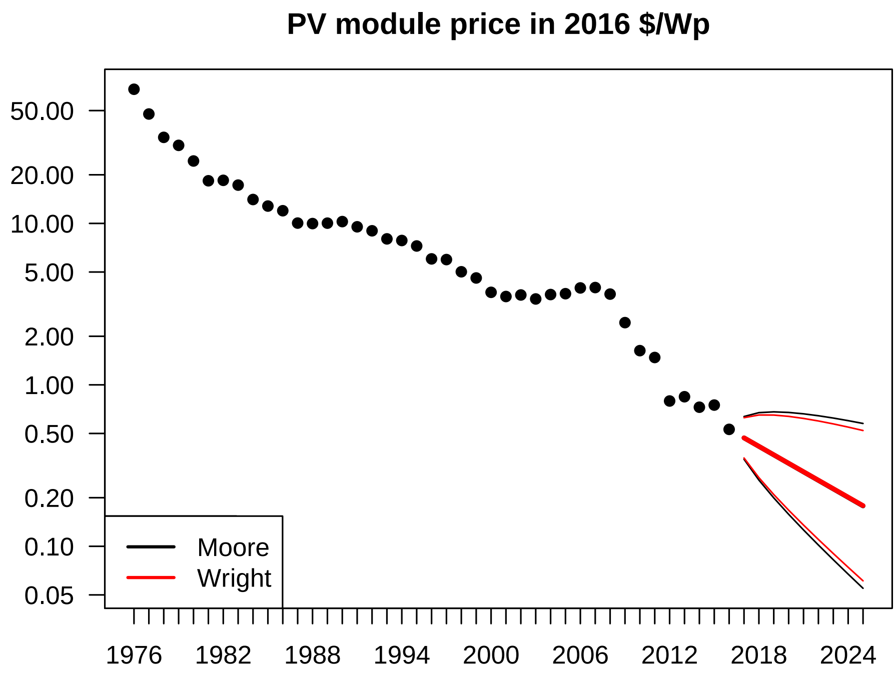


Figure 1: Reproducing Fig. 13 from Lafond et al. (2018).

# References

Farmer, J. D. & Lafond, F. (2016), 'How predictable is technological progress?', *Research Policy* **45**(3), 647–665.

Lafond, F., Bailey, A. G., Bakker, J. D., Rebois, D., Zadourian, R., McSharry, P. & Farmer, J. D. (2018), 'How well do experience curves predict technological progress? a method for making distributional forecasts', *Technological Forecasting and Social Change* **128**, 104–117.

R Core Team (2017), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
**URL:** *https://www.R-project.org/*

[2]Fig. 13 in Lafond et al. (2018) also includes an approximation to Eq.12, which is omitted here.