ESL-TR-93/05-03

Properties of $CV\,[NAC]$ for Linear Energy Models

David K. Ruch
Sam Houston State University
&
Energy Systems Laboratory

May 1993

1. Introduction and Notation

The stability of the NAC estimate for the $PRISM$ model has been shown empirically to be greater than the stability of the other parameter estimates for the $PRISM$ model [1]. The stability of the NAC parameter estimate, as measured by $CV\,[NAC]$, is shown in this report to be superior to the slope and intercept estimates' stability for many linear energy models of the form $E = a + bT$ and a large range of temperature data sets.

Define T_L, as in Fels [1], to be the long-term average temperature (30 years is recommended). Then recall that the parameters a, b, and NAC for the model $E = a + bT$ are related by

$$NAC = 365\,(a + bT_L)\,. \tag{1.1}$$

Let $\bar{T}$ denote the mean of the data set $\{T_i\}_{i=1}^{n}$. We shall need the following change of variable below.

$$T' = T - T_L.$$

Let

$$s_T = \sqrt{\frac{\sum\left(T_i - \bar{T}\right)^2}{n}}$$

and

$$S_{TT} = \sum \left(T_i - \bar{T}\right)^2.$$

We observe that

$$\bar{T}' = \bar{T} - T_L,\ S_{TT} = S_{T'T'},\ \text{and}\ \left(s_T\right)^2 = \frac{S_{TT}}{n}. \tag{1.2}$$

2. *NAC* Stability for Two Parameter Linear Cooling Models

Before stating and proving the stability theorem for NAC, in terms of $CV\,[NAC]$, a few comments on the assumptions are in order. The first assumption in (2.1) says that the typical daily energy consumption, $a + bT_L$, is greater than a, the energy consumption when $T = 0$. The second assumption is that $\bar{T}$, the mean of the temperature data, is closer to the long term average temperature T_L than to zero. These conditions will be satisfied by any cooling model and most data sets, with the exception of winter periods in cold climates. If these conditions do hold, then the NAC estimate will be more reliable than the intercept estimate.

Theorem 2.1. *Suppose that the model $E = a + bT$ is fit to the data set $\{T_i, E_i\}$. Then*

$$|CV\,(a)| > |CV\,(NAC)|$$

if

$$|a| < |a + bT_L|\ \text{ and }\ \left|\bar{T} - 0\right| > \left|\bar{T} - T_L\right|. \tag{2.1}$$

Proof. Using the standard definition of variance (see any introductory statistics book), we see that if the second part of (2.1) holds, then

$$var\,(a) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{T}^2}{S_{TT}}\right] > \sigma^2 \left[\frac{1}{n} + \frac{\left(\bar{T} - T_L\right)^2}{S_{TT}}\right] = var\,(a + bT_L) \tag{2.2}$$

since

$$S_{TT} = \sum \left(T_i - \bar{T}\right)^2 = \sum \left(T_i' - \bar{T}'\right)^2 = S_{T'T'}.$$

Now by (2.2) and the first part of (2.1), we have

$$|CV\,(a)| = \frac{\sqrt{var\,(a)}}{|a|} > \frac{\sqrt{var\,(a + bT_L)}}{|a|} > \frac{\sqrt{var\,(NAC)}}{365\,|a + bT_L|} = |CV\,(NAC)|. \blacksquare$$

3. NAC Stability for a General Two Parameter Linear Model.

The next result says that the parameter estimate for NAC is more reliable than the slope estimate b under very moderate conditions on the temperature data. Consider the assumption (3.1). Note that T_0 is the temperature at which the model is no longer physically meaningful, for it is the temperature at which the energy use is zero. Thus for most data sets used to model energy use, $|T_L - T_0|$ will be quite large in comparison to $\left|T_L - \bar{T}\right|$. The condition "$|T_L - T_0| > \sqrt{2} \cdot s_T$" is not a strong assumption, for this means that the spread of the data, as measured by s_T, a (slightly biased) estimator of the standard deviation of the data, is small in comparison to $|T_L - T_0|$. Most data sets used to model energy use will easily satisfy both of these conditions.

Theorem 3.1. *Suppose that the model $E = a + bT$ is fit to the data set $\{T_i, E_i\}$. Define T_0 by $a + bT_0 = 0$. Then*

$$|CV\,(b)| > |CV\,(NAC)|$$

if

$$|T_L - T_0| > \sqrt{2} \cdot \max\left(\left|T_L - \bar{T}\right|, s_T\right). \tag{3.1}$$

Proof. If condition (3.1) holds, then

$$|T_L - T_0| > \sqrt{(s_T)^2 + \left(\bar{T} - T_L\right)^2},$$

so

$$|T_L - T_0| = \left|T_L + \frac{a}{b}\right| = |b|^{-1} \cdot |a + bT_L|$$

yields

$$|a + bT_L| > |b| \cdot \sqrt{(s_T)^2 + \left(\bar{T} - T_L\right)^2}. \tag{3.2}$$

Recalling (1.2), we observe that (3.2) is equivalent to

$$|a + bT_L| > |b| \cdot \sqrt{\frac{S_{TT}}{n} + \left(\bar{T}'\right)^2}.$$

Thus

$$\frac{|a + bT_L|}{\sqrt{\frac{1}{n} + \frac{\left(\bar{T}'\right)^2}{S_{T'T'}}}} > \frac{|b|}{\frac{1}{\sqrt{S_{T'T'}}}}.$$

Now the model $E = a + bT$ can be written as $E = (a + bT_L) + bT'$, so

$$\frac{|a + bT_L|}{\sqrt{var\,(a + bT_L)}} > \frac{|b|}{var\,(b)},$$

or

$$\frac{|NAC|}{s.e.\,(NAC)} > \frac{|b|}{s.e.\,(b)}. \tag{3.3}$$

Hence (3.3) is equivalent to

$$|CV\,(b)| > |CV\,(NAC)|\,. \blacksquare$$

References

[1] M. Fels, Special issue devoted to measuring energy savings: The Scorekeeping Approach, *Energy and Buildings*, 9, 1986.