Bidirectional Estimation and Confidence Regions for TES Processes

Invited paper, presented to honor Edward G. Coffman, Jr. on his 60th Birthday

David L. Jagerman
Benjamin Melamed

NEC USA, Inc.
C&C Research Laboratories
4 Independence Way
Princeton, NJ 08540

Abstract

TES (Transform-Expand-Sample) is a versatile class of stationary stochastic processes which can model arbitrary marginals, a wide variety of autocorrelation functions, and a broad range of sample path behaviors. The TES modeling methodology aims to simultaneously capture the empirical marginal distribution (histogram) and autocorrelation function of empirical time series, assuming only that they are from a stationary probability law.

In this paper we utilize the known transition structure of TES processes to calculate bidirectional point estimates for these processes as conditional expectations of the process and its time-reversed version, given the current value. We also show how to construct symmetric confidence regions about these point estimates. We demonstrate our results with an example, using the software environment, TEStool, which supports TES modeling.

1 Introduction

TES (Transform-Expand-Sample) is a versatile class of stochastic sequences defined via an autoregressive scheme with modulo-1 reduction and additional transformations [9, 4, 5, 6]. The scope of TES models encompasses a wide variety of sample path behaviors, which in turn give rise to autocorrelation functions with diverse functional forms — monotone, oscillatory, alternating and others. TES sequences are readily generated on a computer, and their autocorrelation functions can be numerically computed from accurate analytical formulas at a modest computational cost.

The TES modeling methodology strives to fit a model to an empirical sample path (field measurements) [10, 11]. It generates a TES model whose marginal distribution is always guaranteed to match its empirical counterpart (histogram); simultaneously, the model's autocorrelation function can approximate the leading empirical autocorrelations to considerable lags. Furthermore, the sample paths generated by the model are required to "resemble" their empirical counterparts.

TES modeling requires software support. A visual interactive software environment, called TEStool, allows users to efficiently pursue TES modeling activities [3]. This environment facilitates heuristic searches for TES models and relieves the tedium of repetitive search by casting it into an intuitive procedure, similar to a video game. A modeling algorithm has also been incorporated into TEStool, which largely automates the modeling process [7].

This paper treats bidirectional forecasting of TES processes, i.e., both forward (postgrade) and backward (retrograde) predictions, the latter being based on time reversal. The known transition structure of TES processes is utilized to calculate bidirectional point estimates as conditional expectations of the process and its time-reversed version, given the current value, as well as symmetric confidence regions about them. Following some preliminaries in Section 2, postgrade forecasting is treated in Sections 3 – 4, and retrograde forecasting in Section 5. Section 6 concludes the paper with an example using TEStool.

2 Preliminaries

This section contains preliminaries pertaining to TES processes. The reader is referred to [9, 4, 5, 6] for additional details.

Throughout this paper we adopt the following notation. The indicator function of a set A is written $1_A(x)$. The Laplace Transform of a function $h(x)$ is consistently denoted by appending a tilde to the function symbol, i.e.,

$$\tilde{h}(s) \;=\; \int_{-\infty}^{\infty} e^{-sx}\, h(x)\, dx. \tag{2.1}$$

Let $\{V_n\}_{n=1}^{\infty}$ be a sequence of iid (independent identically distributed) random variables, called the *innovation sequence*. Let U_0 be uniform on $[0,1)$ (denoted by $U_0 \sim$ Uniform(0,1)), and independent of the innovations $\{V_n\}$. We define two classes of TES sequences, called TES$^+$ and TES$^-$ *background processes*, each consisting of autocorrelated, identically distributed Uniform(0,1) sequences. The corresponding generic processes are denoted by $\{U_n^+\}_{n=0}^{\infty}$

and $\{U_n^-\}_{n=0}^{\infty}$, and given, respectively, by

$$U_n^+ = \begin{cases} U_0, & n = 0 \\ \langle U_{n-1}^+ + V_n \rangle, & n > 0 \end{cases} \quad (2.2)$$

and

$$U_n^- = \begin{cases} U_n^+, & n \text{ even} \\ 1 - U_n^+, & n \text{ odd}. \end{cases} \quad (2.3)$$

The angular brackets in (2.2) denote the modulo-1 (fractional part) operator $\langle x \rangle = x - \max\{\text{integer } n : n \le x\}$. The superscript notation in (2.2)-(2.3) is motivated by the fact that $\{U_n^+\}$ and $\{U_n^-\}$ can generate lag-1 autocorrelations in the range [0,1] and [-1,0], respectively. In the sequel, we shall always append plus or minus superscripts to other mathematical objects associated with $\{U_n^+\}$ and $\{U_n^-\}$ in a natural way; however, the subscript is omitted when the distinction is immaterial. Accordingly, from [4], both $\{U_n^+\}$ and $\{U_n^-\}$ are Markov with corresponding τ-step transition densities, $g_\tau^+(y|x)$ and $g_\tau^-(y|x)$, $\tau = 0, 1, \ldots$, given by

$$g_\tau^+(y|x) = \begin{cases} \displaystyle\sum_{\nu=-\infty}^{\infty} \tilde{f}_V^\tau(i2\pi\nu)\, e^{i2\pi\nu(y-x)}, & 0 \le y, x < 1 \\ 0, & \text{otherwise} \end{cases} \quad (2.4)$$

and

$$g_\tau^-(y|x) = \begin{cases} g_\tau^+(y|x), & 0 \le y, x < 1, \ \tau \text{ even} \\ \displaystyle\sum_{\nu=-\infty}^{\infty} \tilde{f}_V^\tau(i2\pi\nu)\, e^{-i2\pi\nu(y+x)}, & 0 \le y, x < 1, \ \tau \text{ odd} \\ 0, & \text{otherwise} \end{cases} \quad (2.5)$$

where $\tilde{f}_V$ is the Laplace transform (2.1) of the innovation density, f_V. Similarly, $\bar{\tilde{f}}_V$ will denote the complex conjugate of a Laplace transform, the bar notation indicating conjugation.

A Lebesgue measurable transformation D from $[0, 1)$ to the reals is called a *distortion*. It is used to transform a sequence $\{U_n\}$ with uniform marginals on [0,1) to a sequence $\{X_n\}$, such that $X_n = D(U_n)$. In particular, when $D = F^{-1}$ for any cumulative distribution function (cdf), F, then D corresponds to the *inversion method* [2]; each X_n then has the common marginal distribution F. The stochastic sequences, $\{X_n^+\}_{n=0}^{\infty}$ and $\{X_n^-\}_{n=0}^{\infty}$, called TES$^+$ and TES$^-$ *foreground processes*, respectively, are obtained as

$$X_n^+ = D(U_n^+), \quad X_n^- = D(U_n^-). \quad (2.6)$$

Lemma 1 in [4] provides the theoretical basis for the TES modeling methodology. It states that if $U \sim$ Uniform$(0, 1)$ and V is arbitrarily distributed but independent of U, then $\langle U + V \rangle \sim$ Uniform$(0, 1)$. Consequently, both $\{U_n^+\}$ and $\{U_n^-\}$ have Uniform$(0, 1)$ marginals, so $\{X_n^+\}$ and $\{X_n^-\}$ have general marginals.

In order to "smooth" TES sample paths, we routinely use the so-called family of *stitching transformations* S_ξ, $0 \le \xi \le 1$, defined by

$$S_\xi(y) = \begin{cases} \dfrac{y}{\xi}, & \text{if } 0 \le y < \xi \\ \dfrac{1-y}{1-\xi}, & \text{if } \xi \le y < 1 \end{cases} \quad (2.7)$$

Note that for $0 < \xi < 1$, the "smoothing" effect of the S_ξ is due to their continuity on the unit circle. It is quite straightforward to show that S_ξ preserves uniformity for for all $0 \le \xi \le 1$, i.e., if $U \sim$ Uniform$(0, 1)$ then $S_\xi(U) \sim$ Uniform$(0, 1)$ [9]. The composite distortion D_ξ, defined by

$$D_\xi(x) = D(S_\xi(x)), \quad x \in [0, 1), \quad (2.8)$$

is more general than D since $D_1 = D$ as a special case. In particular, any empirical density (histogram) can be expressed mathematically as a probabilistic mixture of uniform random variables, and the associated cdf, H, is just a piecewise-linear function. It is straightforward to construct the inverse distribution function $D = H^{-1}$, and consequently, any prescribed empirical distribution can be easily fitted.

3 Point Estimators for TES Processes

In this section we treat point estimators for TES processes τ time units away ($\tau = \pm 1, \pm 2, \ldots$), conditioned on the present. More specifically, we derive the (conditional) point estimates

$$m_X^+(u, \tau) = E[X_{n+\tau}^+ | U_n^+ = u] \quad (3.9)$$

and

$$m_X^-(u, \tau) = E[X_{n+\tau}^- | U_n^- = u] \quad (3.10)$$

and then proceed to costruct symmetric confidence regions about them, using the transition densities (2.4) and (2.5). Note that negative τ corresponds to retrograde forecasting.

We begin with postgrade estimation for TES$^+$ processes. From the construction of $\{X_n^+\}$ in (2.8),

$$m_X^+(u, \tau) = \int_0^1 D(S_\xi(v)\, g_\tau^+(v|u)\, dv. \quad (3.11)$$

Using (2.7), one may dissect the integral into two parts, followed by a change of variables, yielding

$$\begin{aligned} m_X^+(u, \tau) &= \int_0^\xi D\left(\frac{v}{\xi}\right) g_\tau^+(v|u)\, dv \\ &+ \int_\xi^1 D\left(\frac{1-v}{1-\xi}\right) g_\tau^+(v|u)\, dv \\ &= \xi \int_0^1 D(v)\, g_\tau^+(\xi v|u)\, dv \\ &+ (1-\xi) \int_0^1 D(v)\, g_\tau^+(1-(1-\xi)v|u)\, dv. \end{aligned}$$

Finally, we substitute (2.4) into the above. Assuming that $g_\tau^+(v|u)$ is of bounded variation, and since D is summable, the Parseval formula [1] justifies the interchange of integration and summation, yielding

$$m_X^+(u,\tau) \;=\; \sum_{\nu=-\infty}^{\infty} e^{-i2\pi\nu u}\,\tilde{f}_V(i2\pi\nu)^\tau\,\xi\int_0^1 e^{i2\pi\nu\xi v}\,D(v)\,dv + \sum_{\nu=-\infty}^{\infty} e^{-i2\pi\nu u}\,\tilde{f}_V(i2\pi\nu)^\tau\,(1-\xi)\int_0^1 e^{-i2\pi\nu(1-\xi)v}\,D(v)\,dv. \tag{3.12}$$

To simplify (3.12), we introduce the Fourier coefficients

$$q_\nu(\xi) \;=\; \xi\int_0^1 e^{-i2\pi\nu\xi v}\,D(v)\,dv, \tag{3.13}$$

$$r_\nu(\xi) \;=\; (1-\xi)\int_0^1 e^{i2\pi\nu(1-\xi)v}\,D(v)\,dv, \tag{3.14}$$

$$s_\nu(\xi) \;=\; q_\nu(\xi) \;+\; r_\nu(\xi). \tag{3.15}$$

Note that the $s_\nu(\xi)$ constitute the Fourier coefficients of the function D_ξ of (2.8). It is also of interest to note the relations,

$$r_\nu(\xi) \;=\; \bar{q}_\nu(1-\xi) \tag{3.16}$$

and

$$s_0(\xi) \;=\; \int_0^1 D(v)\,dv \;=\; E[X_n^+], \tag{3.17}$$

the latter showing that $s_0(\xi)$ is, in fact, independent of ξ and equals the common mean of the X_n^+. Finally, $m_X^+(u,\tau)$ can be simply written in terms of the Fourier coefficients above as

$$m_X^+(u,\tau) \;=\; \sum_{\nu=-\infty}^{\infty} \bar{\tilde{f}}_V(i2\pi\nu)^\tau\,s_\nu(\xi)\,e^{i2\pi\nu v}. \tag{3.18}$$

Postgrade estimation for TES$^-$ processes is similar. An analogous derivation of $m_X^-(u,\tau)$ yields

$$m_X^-(u,\tau) \;=\; \int_0^1 D(S_\xi(v)\,g_\tau^-(v|u)\,dv, \tag{3.19}$$

and the TES$^-$ analog of (3.18) is

$$m_X^-(u,\tau) = \begin{cases} \displaystyle\sum_{\nu=-\infty}^{\infty} \bar{\tilde{f}}_V(i2\pi\nu)^\tau\,s_\nu(\xi)\,e^{i2\pi\nu v}, & \tau \text{ even} \\ \displaystyle\sum_{\nu=-\infty}^{\infty} \bar{\tilde{f}}_V(i2\pi\nu)^\tau\,\bar{s}_\nu(\xi)\,e^{i2\pi\nu v}, & \tau \text{ odd} \end{cases} \tag{3.20}$$

4 Confidence Regions for TES Processes

In this section we derive formulas for TES confidence regions of the form $(m_X(u,\tau) \pm \gamma)$, for $\gamma > 0$, with corresponding probablities, $1 - \alpha_X(u,\tau)$. These regions are obtained by first taking symmetric intervals on the unit circle, which straddle the point estimate for the background process, $\{U_n\}$; since intervals on the circle can consist of one or two intervals on the real line [10], one gets in fact confidence regions for $\{U_n\}$. The requisite confidence regions for the associated foreground process, $\{X_n\}$, are then obtained via the appropriate distortion, and their complementary probability is

$$\alpha_X(u,\tau) \;=\; P\{|X_{n+\tau} - m_X(u,\tau)| > \gamma | U_n = u\}. \tag{4.21}$$

Thus, $\alpha_X(u,\tau)$ is the probabilty that the confidence region does not contain the true value of the foreground process, τ lags away, given the current value of the background process. For a prescribed $\alpha = \alpha_X(u,\tau)$ and given u and τ, one wishes to calculate the corresponding $\gamma = \gamma(\alpha,u,\tau)$. To this end, we first obtain formulas which relate α to γ.

We begin with postgrade confidence regions of TES$^+$ processes. For $X_n^+ = F_X^{-1}(S_\xi(U_n^+))$,

$$\begin{aligned}
\alpha_X^+(u,\tau) \;&=\; \int_{\{v:\,|F_X^{-1}(S_\xi(v)) - m_X^+(u,\tau)| > \gamma\}} g_\tau^+(v|u)\,dv \\
&=\; \int_{\{v:\,S_\xi(v)) > F_X(m_X^+(u,\tau)+\gamma)\}} g_\tau^+(v|u)\,dv \\
&\quad+\; \int_{\{v:\,S_\xi(v)) < F_X(m_X^+(u,\tau)-\gamma)\}} g_\tau^+(v|u)\,dv \\
&=\; \int_0^{\xi F_X(m_X^+(u,\tau)-\gamma)} g_\tau^+(v|u)\,dv \\
&\quad+\; \int_{\xi F_X(m_X^+(u,\tau)+\gamma)}^{1-(1-\xi)F_X(m_X^+(u,\tau)+\gamma)} g_\tau^+(v|u)\,dv \\
&\quad+\; \int_{1-(1-\xi)F_X(m_X^+(u,\tau)-\gamma)}^{1} g_\tau^+(v|u)\,dv.
\end{aligned}$$

Substituting above the representation (2.4) for $g_\tau^+(v|u)$ yields after some manipulation,

$$\begin{aligned}
\alpha_X^+(u,\tau) \;=\;& 1 \;-\; F_X(m_X^+(u,\tau)+\gamma) \;+\; F_X(m_X^+(u,\tau)-\gamma) \\
&+ \sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} \frac{\bar{\tilde{f}}_V(i2\pi\nu)^\tau}{i2\pi\nu}\,e^{i2\pi\nu u} \;\times \\
&\Big[e^{-i2\pi\nu\xi F_X(m_X^+(u,\tau)+\gamma)} - e^{i2\pi\nu(1-\xi)F_X(m_X^+(u,\tau)+\gamma)} + \\
&e^{i2\pi\nu(1-\xi)F_X(m_X^+(u,\tau)-\gamma)} - e^{-i2\pi\nu\xi F_X(m_X^+(u,\tau)-\gamma)}\Big],
\end{aligned} \tag{4.22}$$

where $m_X^+(u,\tau)$ is given by (3.18).

Condider next postgrade confidence regions for TES$^-$ processes of the form $X_n^- = F_X^{-1}(S_\xi(U_n^-))$. For τ even,

$$\alpha_X^-(u,\tau) \;=\; \alpha_X^+(u,\tau), \tag{4.23}$$

and for τ odd,

$$\alpha_X^-(u,\tau) = \\ 1 - F_X(m_X^-(u,\tau)+\gamma) + F_X(m_X^-(u,\tau)-\gamma) \\ + \sum_{\substack{\nu=-\infty \\ \nu\neq 0}}^{\infty} \frac{\bar{\tilde{f}}_V(i2\pi\nu)^\tau}{i2\pi\nu} e^{i2\pi\nu u} \times \\ \Big[e^{-i2\pi\nu(1-\xi)F_X(m_X^-(u,\tau)+\gamma)} - e^{i2\pi\nu\xi F_X(m_X^-(u,\tau)+\gamma)} + \\ e^{i2\pi\nu\xi F_X(m_X^-(u,\tau)-\gamma)} - e^{-i2\pi\nu(1-\xi)F_X(m_X^-(u,\tau)-\gamma)}\Big], \tag{4.24}$$

where $m_X^-(u,\tau)$ is given by (3.20).

Finally, utilizing the computable relations, (4.22 – 4.24), for given u and τ and a prescribed α, one may calculate the corresponding $\gamma = \gamma(\alpha)$ by means of a suitable search procedure, such as interpolation ([8], Volume 2); a simple interpolation scheme, such as linear interpolation (false position) is quite effective.

5 Retrograde Forecasting for TES Processes

For retrograde forecasting of TES$^+$ processes, the innovation variates, V_n, are just replaced by $-V_n$ [4]; consequently, it only remains to replace the innovation density, f_V by f_{-V}, where

$$f_{-V}(x) = f_V(-x). \tag{5.25}$$

Furthermore, since $\tilde{f}_{-V}(s) = \tilde{f}_V(-s)$, it remains to replace $\tilde{f}_V(i2\pi\nu)$ in Eq. (4.22) by

$$\tilde{f}_{-V}(i2\pi\nu) = \tilde{f}_V(-i2\pi\nu) = \bar{\tilde{f}}_V(i2\pi\nu). \tag{5.26}$$

The retrograde forecasting of TES$^-$ processes involves a similar replacement in Eqs. (4.23) and (4.24).

At this juncture, it is worth explaining why retrograde forecasting is necassary for TES processes. In practice, the conditioning event, $\{U_n = u\}$, is unknown, and must be "deduced" from the observable event, $\{X_n = x\}$, where $X_n = F_X^{-1}(S_\xi(U_n))$. Since F_X^{-1} is monotone, one has $F_X(x) = S_\xi(u)$. However, for a given value, x, this equation has two solutions in u, as is evident from (2.7). The modeler then uses that solution for u which best fits the observable past, on the grounds that a better retrograde prediction of the past will also give rise to a better postgrade prediction of the future.

6 Conclusion

The point estimators and confidence regions described above have been incorporated into TEStool and tried on various empirical time series.

Figure 1 depicts a TEStool screen which displays the TES$^+$ model described in [11]. This model was constructed from an empirical record of a bitrate sequence of H.261 compressed video (see upper-left canvas), by matching its empirical distribution (histogram) (see upper-right canvas) and empirical autocorrelation function (see lower-left canvas); the innovation density of the model appears in the lower-right canvas. Next, focus on the upper-left canvas, which displays the final portion of the empirical time series (bullets) with the postgrade and retrograde point estimators (small squares) of Section 3, surrounded by 90% confidence regions of Section 4. The single bullet between observations 204 and 212 (which lacks a confidence region), is the foreground value, $x = D(u)$, corresponding to the "current" background value, u, on which all predictions are conditioned. Note that the confidence regions tend to increase in the absolute value of the lag, for both postgrade and retrograde predictions, reflecting the increasing variance (imprecision) of the estimators.

References

[1] Bary, N. K. *A Treatise on Trigonometric Series*, MacMillan, 1964.

[2] Bratley, P., Fox, B.L. and Schrage, L.E. *A Guide to Simulation*, Springer-Verlag, 1987.

[3] Geist, D. and Melamed, B. "TEStool: An Environment for Visual Intractive Modeling of Autocorrelated Traffic", *Proceedings of the 1992 International Conference on Communications*, Vol. 3, 1285-1289, Chicago, Illinois, 1992.

[4] Jagerman, D.L. and Melamed, B. "The Transition and Autocorrelation structure of TES Processes Part I: General Theory". *Stochastic Models*, Vol. 8(2), 193-219, 1992.

[5] Jagerman, D.L. and Melamed, B. "The Transition and Autocorrelation structure of TES Processes Part II: Special Cases". *Stochastic Models*, Vol. 8(3), 499-527, 1992.

[6] Jagerman, D.L. and Melamed, B. "The Spectral Structure of TES Processes", *Stochastic Models*, Vol. 10, No. 3, 599–618, 1994.

[7] Jelenković, P. and Melamed, B. "Algorithmic Modeling of TES Processes", *Preprint*, NEC USA Inc, Princeton, New Jersey, 1994.

[8] Knuth, D.E. *The Art of Computer Programming*, 2-nd edition, Addison-Wesley, 1981.

[9] Melamed, B. "TES : A Class of Methods for Generating Autocorrelated Uniform Variates". *ORSA J. on Computing*, Vol. 3(4), 317-329, 1991.

[10] Melamed, B. "An Overview of TES Processes and Modeling Methodology", in *Performance Evaluation of Computer and Communications Systems*, (L. Donatiello and R. Nelson, Eds.), 359–393, Lecture Notes in Computer Science, Springer-Verlag, 1993.

[11] Melamed, B. "The TES Methodology: Modeling Temporal Dependence in Empirical Time Series" *Proceedings of MASCOTS '93*, 11–16, La Jolla, California, 1993.

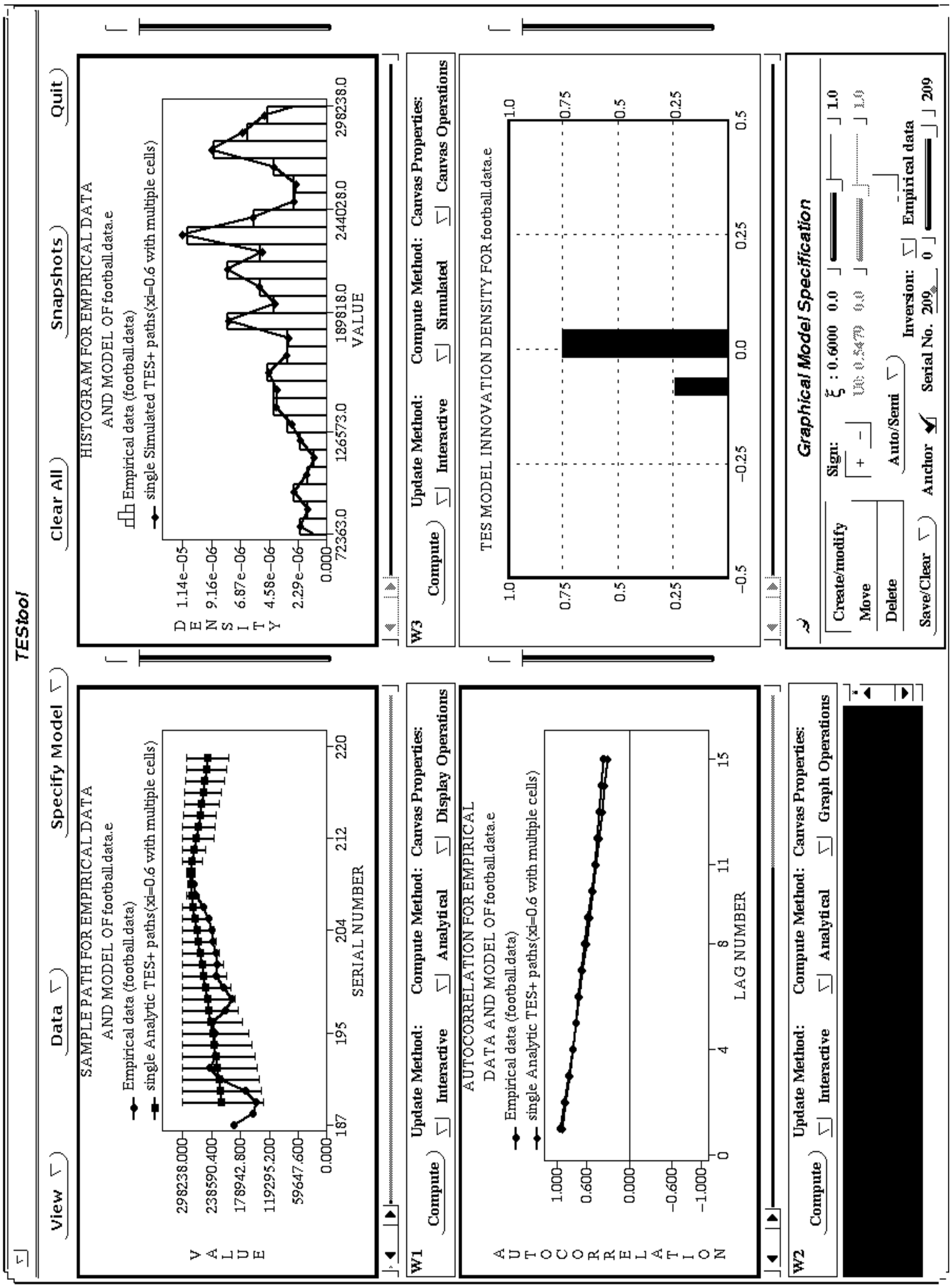

Figure 1: A TEStool screen showing biderctional forcasting for a TES model of compressed video bitrates