A Non-parametric Approach to Pricing and Hedging Derivative Securities: with an Application to LIFFE Data

by

J. A. Barria[1] and S. G. Hall[2]

Abstract

It is important for financial institutions to develop methods to predict their exposure and keep their risk under control. Portfolio managers can insure themselves against the value (of a diversified stock portfolio) dropping below a certain level, by holding in conjunction with the stock portfolio, an index option derivative security. The work reported in this paper is concerned with the study of non-parametric methods for estimating the pricing formula of option derivative securities. Two non-parametric approaches, the Projection pursuit method (PPR) and the Local polynomial approach (LOESS), are studied and compared to a benchmark parametric Black-Scholes (B-S) approach.

The practical relevance of these approaches is tested, when applied to pricing and hedging of real-world LIFFE FTSE 100 index options from April 97 to November 97. We compare the two methods by means of constructing a riskless portfolio of stocks, bonds and option derivatives securities. The portfolio is then delta-hedged on a daily basis using a dynamic trading strategy in stocks and bonds during the lifetime of the option instrument. The tests carried out show that both methods generate similar responses, although each method can outperform the others depending on market conditions, such as, time to maturity of the option instrument.

[1]J. A. Barria is with The Electrical and Electronic Engineering Department, Imperial College of Science Technology and Medicine.

[2]S. G. Hall is with The Management School, Imperial College of Science Technology and Medicine.

1. Introduction

Derivative instruments are securities whose value is obtained from the value of an underlying security or basket of securities. The most common derivative instruments are options, forwards and futures. Much of the success of the market for options (and other derivative) securities may be traced back to the seminal work done by F. Black and M. Scholes [BLA73] and R. Merton [MER73], in which closed-form options pricing formulas were obtained through a dynamic hedging argument and a no-arbitrage condition. These formulas have since been generalised and applied to a vast range of securities and contexts. The derivation of these pricing formulas, either analytically or numerically, depends on the particular parametric form of the underlying asset's price dynamics.

In this paper non-parametric models for pricing and hedging derivative securities, in which the data, e.g., asset price, strike price and time to maturity, is allowed to determine both the dynamics of the asset's price, and its relation to the prices of derivative securities with minimal assumptions on the asset's price and the derivative pricing model (See, e.g., [BAR97], [BAR91], [HAR95], [HAR90], [MUL88], and [FAN96]). Non-parametric models would estimate a conditional expected value $E[y_t|x_t]$ without imposing any functional form upon this relation,

$$y_t = m(x_t) + u_t \tag{1}$$

where $m(x_t)$ is some unknown function of x_t, and u_t is assumed to be noise. In this paper y_t is regarded as the price of a call option derivative security.

The price of a call option can be derived using parametric approaches. For example, the Black-Scholes (B-S) formula, links the price of the derivative asset, S, the strike or exercise price, X, the instantaneous standard deviation (or volatility), σ, the risk-free interest rate, r, and the maturity (or time to expiration), T, of the derivative security. Unfortunately, the B-S formula is known to break down for out-of-the money (when the discount exercise price is greater that the asset price) and short-maturity options, where the degree of nonlinearity is high. Moreover, implicit volatility, computed from options prices, change over time and seems to be a function of the ratio between the price of a derivative asset, S, and the strike price X [PAG96].

The basic idea behind non-parametric estimation of $m(x_t)$ is to combining a set of basis functions $\psi_j(x)$, and to then estimate the correspondong set of parameters β_j, leading to $\hat{m}\ (x_t) = \sum_{j=1}^{M} \hat{\beta}_j \psi_j(x)$. Hutchinson et. al. [HUT94] studied the options model above using neural networks. Kearns [KEA93] simulated data from an option pricing model with stochastic volatility and non-parametrically estimated the non-linear function implied by a particular pricing model using the flexible Fourier form of Gallant [GAL81]. Both of these methods have approximated the unknown function globally and then predicted what value it

would have at given points $x_t = x$.
An alternative non-parametric approach is to estimate $\alpha = m(x_t)$ by using only those observations whose value is close to the target value x. An appropriate estimate can be found (if m(x) is assumed constant) by choosing α to minimise $\sum^{T}_{t=1}(y_t - \alpha)^2 K((x_t - x) / h)$, where K(.) is a kernel or weighting function that gives low weight to observations for which x_t is far from x. The window-width parameter, h, determines exactly how far away the observation can be in order to be included in the computation. The principal disadvantages of such local analysis are that there are well known biases in the estimator of m(x) and, when the $\dim(x_t)$ is large, it is likely that very few observations will be used to determine each point. In recent years an attempt has been made to combine the two approaches by assuming that the function can be parametrically approximated around the point x by, e.g., a linear polynomial, so that the optimisation problem becomes one of choosing α and β such that $\sum_{t=1}^{T}(y_t - \alpha - \beta(x_t - x))^2 K((x_t - x)/h)$ is minimised. In [FAN96], it is shown that these locally parametric methods can produce big improvements in the properties of the local estimator. Recently, Bossaerts and Hillion [BOS97] have proposed using other functions rather than polynomial functions to price options.

The underlying aim of the work reported in this paper is to extend the Hutchinson et. al. non-parametric estimation tests by including the multivariate local polynomial regression approach (LOESS) and contrasting it with the projection pursuit approach (PPR). The set of data used is obtained from LIFFE FTSE-100 Index options from April -97 to November 97.

Even though, some theoretical advantages of non-parametric models over the parametric models have been presented in [HUT94], such approaches would not be appropriate for rarely traded derivatives or newly created derivatives. Furthermore, if the fundamental asset's price dynamics is well understood and an analytical expression for the derivative's price is available under these dynamics, then the parametric formula will almost always dominate the non-parametric approach in pricing and hedging accuracy.

Section 2 gives a concise background review of non-parametric approaches. Section 3 investigates and evaluates the range of applicability of non-parametric approaches. The Projection pursuit method (PPR) and the Local polynomial approach (LOESS) are studied and compared to a benchmark parametric Black-Scholes (B-S) approach. The comparison is done in terms of an out-of-sample delta-hedging strategy using real-world data obtained from London International Financial Future Exchange (LIFFE). The paper concludes in Section 4.

2. Non-parametric Regression methods

The well known B-S model [BLA90 and HUL97] and its extension assume that the probability distribution of the stock price at any given time is lognormal. If this assumption is incorrect, there are liable to be biases in the prices produced by the model. To this effect, the flexibility of non-parametric approaches is extremely helpful in a preliminary and exploratory statistical analysis of a data set [HAR90]. For example, if no a priori model information about the regression curve is available, the non-parametric analysis could help in suggesting simple parametric formulations of the regression relationship.

Linear regression is one of the most widely used non-parametric techniques. For data in which the graphical evidence clearly indicates a linear relationship between the predictors and the response, a linear model provides a simple and easy description of the data. While this approach has been widely used, it suffers from a few drawbacks [FAN96]. Another intuitive estimator for the conditional mean function m(x) is the running local average or its improved version, the locally weighted average [FAN96]. Since bias increases and variance decreases with increasing bandwidth h (which is a non-negative number controlling the size of the local neighbourhood of the region around x), selection of h is a compromise between bias and variance in order to achieve small mean squared error [FAN96].

2.1. Local polynomial regression and derivative estimation

From the function approximation point of view, some well used estimators (e.g. the Nadaraya-Watson estimator) uses a locally constant approximation. This approximation may suffers from large bias, particularly in regions where m'(x) is large. The bias of the Nadaraya-Watson estimator is also large at the boundary of the sample points. One way to repair these drawbacks is to use a higher-order approximation [HAR95]. Suppose that the regression function m(x) can be approximated locally to x, using a Taylor's expansion, by a polynomial of order p,

$$\mathrm{m}(\mathrm{X}_i) \approx \sum_{\mathrm{j=0}}^{\mathrm{p}} \frac{\mathrm{m}^{(\mathrm{p})}(\mathrm{x})}{\mathrm{j!}} (\mathrm{X}_i \text{ - } \mathrm{x})^{\mathrm{j}} \equiv \sum_{\mathrm{j=0}}^{\mathrm{p}} \beta_{\mathrm{j}}(\mathrm{X}_i \text{ -}\mathrm{x})^{\mathrm{j}} \tag{2}$$

for x in a neighbourhood of x_0. This suggests fitting a local polynomial regression,

$$\min \sum_{\mathrm{i=1}}^{\mathrm{n}} \{ \mathrm{Y_i} \text{ - } \sum_{\mathrm{j=0}}^{\mathrm{p}} \beta_{\mathrm{j}}(\mathrm{X_i}\text{-}\mathrm{x})^{\mathrm{j}}\}^2 \mathrm{K_h}(\mathrm{X_i} \text{ - } \mathrm{x}) \tag{3}$$

where K(.) denotes a kernel function, and h is a bandwidth. Note that when p = 0, the above estimator is the Nadaraya-Watson approach. An in depth treatment of the optimal bandwidth, $\mathrm{h}_{\mathrm{opt}}$, and the values of the bias and variance of this regression method can be found in [FAN96].

Let $\hat{\beta}_0$(x), ... , $\hat{\beta}_{\mathrm{p}}$(x) be the minimiser of equation (3). The local polynomial regression estimator of the regression function is $\hat{m}(\mathrm{x}_0) = \hat{\beta}_0(\mathrm{x}_0)$. The whole curve $\hat{m}(.)$ is obtained by running the above local polynomial regression with x_0 varying in an appropriate domain of interest. Furthermore, local polynomial fitting can easily be applied to derivative estimations. Suppose that we fit the local polynomial of order p, then we can obtain an estimator for $\mathrm{m}^{(\nu)}$by

$$\hat{m}^{(\nu)}(\mathrm{x}) = \nu!\ \hat{\beta}_\nu(\mathrm{x}) \tag{4}$$

Usually, the order of the polynomial p is taken to be p = ν+1 or occasionally p = ν+3, based on the consideration of the efficiency of the estimator and cost of computation.

2.2. Multivariate regression approaches

In multivariate regression problems, one of the tasks is to study the structural relationship between the response variable Y and the vector of covariates X = $(X_1, X_2, \dots, X_d)^T$ via m(x) = E(Y | X = x), where x = $(x_1, \dots, x_d)^T$ and m(x) = $m(x_1, \dots, x_d)$. The most flexible models do not make any assumption about the form of the d-variate function m(x). Although generalisation of most of the univariate smoothing techniques to a multivariate surface is feasible, surface smoothing is characterised by a serious problem known as the curse of dimensionality. This problem refers to the fact that a local neighbourhood in higher dimensions is no longer local, that is, a neighbourhood with a fixed percentage of data points can be very big and far from what is understood by the term local.

An additive model extends the notion of a linear model by allowing some or all linear functions of the predictors to be replaced by arbitrary smooth function of the predictors (See. e.g. [BRE85], [FAN96]). In aditive models, a significant limitation on the type of surfaces considered is that the effects enter the model additively, without interactions. A model with component wise additive main terms, and pair wise interactions term is presented in [FAN96]. However, the number of parameters in the model could be high, and therefore, the curse of dimensionality is still present.

The two approaches reported in this paper overcome most of the limitations of previously mentioned models. The multivariate local polynomial regression estimator (LOESS) is a local linear regression estimator procedure proposed by Cleveland and Devlin [CLE88]. In this approach the smoothing is done along the coordinate axes of the covariate, and the amount of smoothing is the same in each direction (See also Appendix A.1). In the other approach, the projection pursuit regression (PPR), the basic idea is to take advantage of the fact that a regression surface may be of a simple additive struture. Further more, instead of using constant functions of projections along the coordinate axes, the regression surface is approximated by a sum of empirically determined univariate functions of a particular projection (See also Appendix A.2).

3. Results and Comments

In this section we investigate the range of applicability of two non-parametric approaches when used to price and hedge call options securities. The projection pursuit method (PPR) and the local polynomial approach (LOESS), are compared to a parametric Black-Scholes (B-S) approach. The techniques are compared in terms of an out-of-sample delta-hedging strategy using real-world data obtained from London International Financial Future Exchange (LIFFE). Although recent theoretical development suggest that there are significant connections between many of the non-parametric methods (see e.g., references in [HUT94]), a rigorous comparison of these methods is not our primary goal.

To study the different approaches we use the same simplification as in [HUT94]. Hutchinson's presentation of results relies on Merton's proposition that assumes that the statistical distribution of the underlying asset's return is independent of the level of the stock price S, and therefore, the option pricing formula is homogeneous of degree one in both S and X. Hence, Y = c/X and X_1 = S/X , X_2 = T-t; where c is the price of a call option derivative security, S is the price of the derivative asset, X is the strike or exercise price, and T-t is the maturity (or time to expiration). If this is the case, we need only to estimate m(S/X, 1 , T-t).

Since we need to construct a delta-hedging strategy, the first derivatives of c need to be evaluated. Delta can be estimated using the LOESS approach (by means of expression (4)). For PPR, however, the use of a smoother for estimating the non-linear function Y, forces a numerical approximation of Delta. The numerical approximation of Delta is accomplished in two steps: (i) a first-order finite difference (DIFF) for the range of the stock price S, (ii) and then applying PPR to the datum point obtained from DIFF. Note that (ii) can also be estimated using LOESS and hence we have also included these results. The non-parametric model complexity was chosen following the findings reported in [HUT94].

Examples of the estimates and errors for the different procedures studied are presented in the sequence of Figures 1 to Figure 5. Figure 1.a shows a typical out-of-sample data set, and Figure 1.b shows the Delta estimates using the first order finite difference (DIFF) approach. Figure 2 shows the two non-parametric regressions here studied (LOESS and PPR) on $Y = \frac{\hat{C}}{X}$ (call option price/strike price) of the out-of-sample data set, and their corresponding residual error plots. Recall that X_1 = S/X (stock price/strike price) and X_2 = T-t (time to expiration).

The results presented in the sequence of Figure 3 to Figure 5 are examples of the following procedures to estimate the Delta of the call option.

LOESS + LOESS: This procedure obtains the $Y = \frac{\hat{C}}{X}$ and the value of Delta using the LOESS approach (expression (4)). The option Delta estimation,and the residual errors of this procedure are shown in Figure 3.

DIFF + LOESS: This procedure estimated first the $Y = \frac{\hat{C}}{X}$ surface using a first order finite difference method (DIFF). To estimate (interpolate) a specific value of Delta uses the LOESS approach (expression (4)).The option Delta estimation and the residual errors of this procedure are shown in Figure 4.

DIFF + PPR: This procedure estimated the $Y = \frac{\hat{C}}{X}$ surface using a first order finite difference method (DIFF). To estimate (interpolate) a specific value of Delta uses the PPR approach. The option Delta estimation,and the residual errors of this procedure are shown in Figure 5.

3.1. Performance measures

We use the same measures provided in [HUT94]hence, just the main points are repeated here. A measure of performance for a given option pricing formula is the difference between the terminal value of the call and the terminal combined value of the stock and bond positions. More formally, denote V(t) as the monetary value of the replicate portfolio at date t and let,

$$V(t) = V_S(t) + V_B(t) + V_C(t) \tag{5}$$

where $V_S(t)$ is the monetary value of stocks, $V_B(t)$ is the monetary value of bonds, and $V_C(t)$ is the monetary value of call options held in the portfolio at date t. The initial comparison of this portfolio at date 0 is assumed to be,

$$V_S(0) = S(0)\Delta_{F_{np}}(0), \qquad \Delta_{F_{np}}(0) \equiv \frac{\partial F_{np}(0)}{\partial S} \tag{6}$$

$$V_C(0) = -\ F_{LIFFE}(0) \tag{7}$$

$$V_B(0) = -\ (V_S(0) + V_C(0)) \tag{8}$$

where $F_{LIFFE}(.)$ is the actual value of the call option and $F_{np}(.)$ is its non-parametric approximation. Since the stock purchase is wholly financed by the combination of riskless borrowing and proceeds from the sale of the call option, the initial value of the replicating portfolio is zero (i.e. $V(0) = 0$).

Prior to expiration, and at discrete and regular intervals of length τ (which we take to be approximately one day in our tests), the stock and bond positions in the replicating portfolio will be rebalanced so as to satisfy the following relations:

$$V_S(t) = S(0)\Delta_{F_{np}}(t), \qquad \Delta_{F_{np}}(t) \equiv \frac{\partial F_{np}(t)}{\partial S} \tag{9}$$

$$V_B(t) = e^{r\tau}V_B(t - \tau) - S(0)(\Delta_{F_{np}}(t)\text{-}\Delta_{F_{np}}(t - \tau)) \tag{10}$$

where $t = k\tau \leq T$ for some integer k. The tracking error of the replicating portfolio is then defined to be the value of the replicating portfolio V(T) at expiration date T. In [HUT94] the following "tracking error" performance measure is suggested,

$$\xi = e^{-rT}E[\ |\ V(T)\ |\ ] \tag{11}$$

The quantity ξ is simply the present value of the expected absolute tracking error of the replicating portfolio.

Another measure of performance may be defined by combining the information contained in the expected tracking error with the variance of the tracking error. In [HUT94] the "prediction error" η is defined as:

$$\eta = e^{-rT}\sqrt{E^2[V(T)] + Var[V(T)]} \tag{12}$$

which is the present value of the square root of the sum of the squared expected tracking error and its variance. Note that the expected tracking error of a delta-hedging strategy might be zero, but the strategy is a poor one if the variance of the tracking error were large.

3.2. Testing non-parametric approaches

The implementation of the non-parametric models and the computation of the graphics presented in this section were carried out using the software package S-Plus [VEN94] an extension of the statistical language S [BEC88]. The Delta hedging test scenario was implemented using Microsoft Excel spreadsheet and the test paths were constructed using [HUL97]. Despite the fact that the B-S model is generally not used in its original form in practice, we used it here because it is still a widely used benchmark model, and because it serves as an example of a parametric model whose assumptions are questionable in the context of LIFFE's real-world data. The LIFFE contract months are March, June, September, December plus the three nearby months. The last trading day is the third Friday of the expiry month (or the last trading business day preceding the third Friday). The data of our empirical analysis are daily closing prices of FTSE 100 LIFFE options for the period from January-1997 to November-1997. We divide the LIFFE data into five non-overlapping three-month periods. For the LIFFE data, the number of future call options per subperiod ranged from 1680 to 2868, with an average of 2163. This data set differs from a synthetic data due to the presence of noise in the real-world option prices and the irregular trading activity of the options, especially for near-term out-of-money options. To limit the effects of non-stationarities we test the regressions only on the data from the immediately following month.

The B-S parameters r (risk free rate) and volatility, σ, were estimated using a window of the most recent data. Specifically, we estimate the B-S volatility, σ, for a given future contract using the immediately preceding subperiod data, and a weighted average of the type suggested by Latane [CHI78]. We estimate the risk free rate, r, for each option as the yield of the 7 days interbank loan average of the preceding month to the initial activity of the option. The approximation to the cumulative normal distribution function to evaluate the B-S model to up to six decimal points suggested by [HUL97] and [BLA90] is used in the test paths.

3.3. Out-of-sample pricing and hedging

To see how the performance varies, we divide the input space into two regimes: short and medium term regimes for the time-to-expiration input. For the short term regime we mean less than one month, and medium term refers to up to 3 months. The other regimes are in-, near- and out-of-the money for the stock-price/strike-price (S/X) input, that is, S/X = {0.97, 1.00, 1.03}. In each one of the tests presented here, the performance of the non-parametric Delta-hedging strategies are compared to the performance of a Delta-hedging strategy using the B-S formula. The values of the following tables is the aggregate result of averages on the paths tests.

3.3.1. Relative tracking error comparison

The results here presented use the ratio between the non-parametric approach to the benchmark B-S solution, that is, $\xi_{\text{Rel.}} = \xi_{\text{np}}/\ \xi_{\text{B-S}}$. Our results are summarized in Table 1 for the short term regime, and in Table 2 for the medium term regime. The results show that some of the non-parametric approaches exhibit less tracking error than B-S in the in-the-money and near-the-money regimes but not in the out-the-money regimes.

$\xi_{\text{Rel.}} = \xi_{\text{np}}/\ \xi_{\text{B-S}}$	BL-SC	LOESS+LOESS	DIFF+LOESS	DIFF+PPR
In_money	1.00	1.28	0.94	0.85
out_money	1.00	3.42	3.86	4.25
near_money	1.00	0.76	0.95	1.10

Table 1: Relative tracking error $\xi_{\text{Rel.}} = \xi_{\text{np}}/\ \xi_{\text{B-S}}$ comparison for short term regime.

$\xi_{\text{Rel.}} = \xi_{\text{np}}/\ \xi_{\text{B-S}}$	BL-SC	LOESS+LOESS	DIFF+LOESS	DIFF+PPR
In_money	1.00	0.99	0.73	0.36
out_money	1.00	1.12	1.27	1.60
near_money	1.00	0.88	0.99	1.19

Table 2: Relative tracking error $\xi_{\text{Rel.}} = \xi_{\text{np}}/\ \xi_{\text{B-S}}$ comparison for medium term regime.

3.3.2. Relative predictor error comparison

The prediction error combines the expectation and variance of the absolute tracking error, hence the estimated prediction error is calculated with the sample mean and sample variance of $|V(T)|$ taken over the set of test paths.

$\eta_{\text{Rel.}} = \eta_{\text{np}}/\ \eta_{\text{B-S}}$	BL-SC	LOESS+LOESS	DIFF+LOESS	DIFF+PPR
In_money	1.00	1.62	1.41	1.07
out_money	1.00	1.46	1.55	1.80
near_money	1.00	0.80	0.85	0.99

Table 3: Relative predictor error $\eta_{\text{Rel.}} = \eta_{\text{np}}/\eta_{\text{B-S}}$ comparison for short term regime.

$\eta_{\text{Rel.}} = \eta_{\text{np}}/\ \eta_{\text{B-S}}$	BL-SC	LOESS+LOESS	DIFF+LOESS	DIFF+PPR
In_money	1.00	1.07	0.86	0.58
out_money	1.00	1.41	1.49	1.87
near_money	1.00	0.90	0.95	1.07

Table 4: Relative predictor error $\eta_{\text{Rel.}} = \eta_{\text{np}}/\eta_{\text{B-S}}$ comparison for medium term regime.

These results show that some of the non-parametric approaches exhibit less tracking error than B-S In and Near-the-money regimes for all maturities (LOESS +LOESS and DIFF+PPR); but not in the out-the-money regime.

We can say that all approaches generate similar responses surfaces and that large errors tend to occur at the kink-point for options at the money, at expiration and also along the boundary of the sample points. Without being to conclusive about these preliminary findings we can observe that B-S approach was consistently better out-the-moeny for all regimes. This seems to suggest that since it is known that the B-S formula break down for out-the-money regimes; the non-parametric approaches are not being very useful in the set of data here analysed. These findings are generally in contrast with the results of [HUT94, p. 884] where all non-parametric approaches are better, except for the near-term in-the-money options.

4. Final Remarks

The underlying aim of the work reported in this paper was to evaluate the usefulness of two non-parametric approaches when used to price and hedge call options securities. The Projection pursuit method (PPR) and the Local polynomial approach (LOESS) formed the core non-parametric algorithms with which the Black-Scholes parametric approach was compared. The comparison was made in terms of an out-of-sample Delta-hedging dynamic trading strategy using real-world data obtained from LIFFE.

The tests reported here aim at investigating, the ability to replicate the option through a dynamic hedging strategy. The results reported in this paper contrast Hutchinson's et al findings. In Hutchinson et al [HUT94] the out-of-sample tests show some evidence that non-parametric approaches outperform the Black-Scholes model on real-world data. In our case, Black-Scholes outperforms the non-parametric approaches when the options price is out-of-the money and/or close to the boundaries of the sample data points. These results can be partially explained by the fact that the non-parametric methods were not "fine tuned" in relation to, e.g., optimal number of data points and size of bandwidth. This highlight the high sensitivity of non-parametric techniques to the "tuning" of their specifications. The choice of model specifications could easily have been biased due to the easiness of manipulation and flexibility of the

LOESS model.

In relation to the choice of local polynomial regression, if the global function is likely to have a specific shape, then it would make sense to use a function which properly represents that shape. For example in [BOS97], the Black-Scholes formula, was used to obtain option prices. An alternative venue for research is to attempt hybrid solutions. In [BOS95] a parametric-driven modelling approach, but a non-parametric conditional volatility method is applied in the study of currency exchange rates.

Another significant sources of error in financial analysis lies in the fact that the small number of observations in a given data set make it difficult to discriminate statistically between alternative hypotheses. Sampling errors also exist when, e.g., data comes from a single set of observations and may therefore not be very representative. Also, the lower the number of observations, the lesser the dependability of the model. On the other hand, big samples have a decisive impact on the accuracy of the estimation. These issues were not explicitly addressed in this study and therefore, the relation between the sample size and approximation error could have influenced our findings. As it has been pointed out in e.g. [HUT94] and [CHO96] this is an important aspect that deserves further investigation.

Finally, the statistical properties of market dynamics are very different from those assumed by derivatives pricing models, which are based on low-frequency data. The classical approach of using low frequency financial data fails to reveal that sampling time is not independent of the pricing process itself [CHO96]. For example, transactions are more likely to occur when there is new information, affecting the variance of the transaction price series, making the whole process market behaviour dependent. In this context non-parametric approaches may become a viable search for new option pricing strategies that could incorporate the full complexity of market behaviour, which would allow us to study and implement more sophisticated discrete hedging trading policies.

References

[BAR97] J. A. Barria von Bischhoffshausen "A General Approach to Pricing and Hedging Derivative Securities: with and Application to LIFFE Data", MBA Thesis, The Management School, Imperial College of Science Technology and Medicine, U. of London, 1997.
[BAR91] W.A. Barnett, J. Powell and G. Tauchen Ed. "Non parametric and semiparametric methods in econometrics, and statistics", Cambridge University Press, 1991.
[BEC88] R. A. Becker, J. M. Chambers and A. R. Wilks, "The new S language: A programming environment for data analysis and graphics", Wadsworth & Brooks, Pacific Grove, California 1988.
[BLA73] F. Black and M. Scholes, "The pricing of options and corporate liabil-

ities", Journal of Political Economy, V 81, pp. 637-654, May-June 1973
[BLA90] D. Blake, "Financial market analysis", McGraw-Hill, London, 1990.
[BOS95] P. Bossaerts and W. Hardle, "Foreign exchange rates have surprising volatility", First international conference in high-frequency data in finance, Zurich, March 1995.
[BOS97] P. Bossaerts and P. Hillion, "Local parametric analysis of hedging in discrete time", Journal of Econometrics, V 81, pp. 243-272, 1997.
[BRE85] L. Breiman and J. H. Friedman, "Estimating optimal transformations for multiple regression and correlation", J. Amer. Statist. Assoc., 80, pp. 580-619, 1985.
[CHO96] D. N. Chorafas, "How to understand and use mathematics for derivatives", Vol. 1 and Vol. 2, Euromoney Publ. plc, London, 1996
[CLE88] W. S. Cleveland and S. J. Devlin, " Locally-weighted regression: an approach to regression analysis by local fitting", J. Amer. Statist. Assoc., 83, pp. 597-610, 1988.
[CHI78] D. Chiras and S. Manaster, "The Information content of stock prices and test of market efficiency", Journal of Financial economics, Vol 6, pp. 213-234, September 1978
[FAN96] J. Fan. and I. Gijbels, "Local Polynomial Modelling and its applications", Chapman-Hall, London, 1996.
[FRI81] J. H. Friedman and W. Stuetzle, "Projection pursuit regression", J. Amer. Statist. Assoc., 76, pp. 817-823, 1981
[GAL81] R. Gallant, "On the bias in flexible functional forms and an essentially unbiased form: The Fourier flexible form", Journal of Econometrics, 15, pp 211-244, 1981.
[HUL97] J. Hull, "Options, futures, and other derivative securities", Prentice-Hall, International, 1997.
[HUT94] J. M. Hutchinson, A. W. Lo and T. Poggio, "A non-parametric approach to pricing and hedging derivative securities via learning networks, Journal of Finance, pp. 851-889, Vol. XLIX, July 1994
[HAR95] W. Hardle, S. Klinke and B.A. Turlach, "XploRe: An interactive statistical computing environment", Springer-Verlag, New-York, 1995.
[HAR90] W. Hardle, "Applied non-parametric regression", Cambridge Univ. Press, New-York, 1990.
[KEA93] P. Kearns, "Volatility and the pricing of interest rate derivative claims", Ph. D. Thesis, U. of Rochester, 1993.
[MER73] R. C. Merton, "Theory of Rational option pricing", Bell Journal of Economics and management Science, V 4, pp. 141-183, Spring 1973.
[MUL88] H-G. Muller, "Non-parametric regression analysis of longitudinal data", Lecture notes in Statistics,Springer-Verlag, Berlin, 1988
[PAG96] A. Pagan, "The econometrics of financial markets", Journal of Empirical Finance, Vol 3, pp. 15-102, 1996
[VEN94] W.N. Venables and B.D. Ripley, "Modern applied statistics with S-Plus", Springer-Verlag, New-York, 1994.

Appendix A:

A.1. The multivariate local polynomial regression estimator: LOESS

The local linear fit is presented (i.e. p = 1). For simplicity of presentation we focus only on estimation of the regression function m(x) ($\nu = 0$) [FAN96]. Lets assume that K is a multivariate probability density function, such that $\int K(u)du = 1$ and $\int uK(u)du = 0$. Further we assume that the mean of the density function K(.) is zero and the covariate matrix of K is $\mu_2(K)\ I_d$, with I_d the $d \times d$ identity matrix. Define

$$K_B(u) = \frac{1}{|B|} K(B^{-1}u),$$

where B is a non-singular $d \times d$ matrix, the bandwidth matrix, and |B| denotes its determinant. The observations are given by $\{(X_i^T,\ Y_i)\text{: } i = 1, \ldots ,n\}$, with $X_i = (X_{i1}, \ldots , X_{id}\ )^T$. Let $x^T = (x_1, \ldots , x_d)$ be a point in $\mathbb{R}^d$. Then, the multivariate version of (3) is given by the following minimisation problem

$$\sum_{i=1}^{n}\{Y_i - \beta_0 - \sum_{j=1}^{d} \beta_j (X_{ij}-x_j)\}^2 K_B(X_i-x), \tag{13}$$

with respect to $\beta = (\beta_0, \ldots , \beta_d)^T$, where now

$$\beta_0 = m(x) \qquad \text{and} \qquad \beta_j = \frac{\partial m(x)}{\partial x_j}, \qquad j = 1, \ldots ,d \tag{14}$$

A robust version of the above multivariate local linear regression estimation procedure, called LOESS, was proposed by Cleveland and Devlin [CLE88]. In LOESS, the bandwidth matrix B is taken to be of the form $B = hI_d$. This means that smoothing is done along the coordinate axes of the covariates, and that the amount of smoothing is the same in each direction. The multivariate kernel function is taken to be of the form,

$$k(u) = k(u_1, \ldots , u_d) = W\ \{\ (\ \sum_{j=1}^{d} u_j^2)^{1/2}\ \} \tag{15}$$

where W(.) is a univariate kernel function. Cleveland and Devlin take W(.) to be the tricube kernel. Furthermore, in LOESS the following Euclidean distance is used as a distance between points in the d-dimensional space $\mathbb{R}^d$

$$\rho(u,v) = \ \{\ \sum_{j=1}^{d}(u_j-v_j)^2\ \}^{1/2} \tag{16}$$

Finally LOESS use a nearest neighbour type of bandwidth, i.e., the neighbourhood of a particular observation X_k is determined by its associated bandwidth h_k which is the r-th smallest number among $\rho(X_k,X_j)$, for j=1, ... ,n. Therefore, when estimating m(.) at the observation X_k, a weight

$$K\{\frac{1}{h_k}\ (X_i - X_k)\} = W\ [\ \frac{1}{h_k}\ \{\ \sum_{j=1}^{d}(X_{ij} - X_{kj})^2\}^{1/2}] \tag{17}$$

is assigned to each observation X_i. The bias and variance of this approach can be found in [FAN96].

A.2. Projection Pursuit Regression: PPR

The basic idea of additive models is to take advantage of the fact that a regression surface may be of a simple, additive structure. In the proposal of Friedman and Stuetzle [FRI81], the idea is that instead of using constant functions of projections along the coordinate axis, the regression surface is approximated by a sum of empirically determined univariate ridge functions $\{g_j\}$ of projection $\beta^T x$,

$$m(x) = \sum_{j=1}^{p} g_j\{\beta_j^T x\} \tag{18}$$

In this class of exploratory projection technique, the idea is to describe "interesting" projections by maximising an objective function or projection pursuit index. The classical projection pursuit tries to find non normal projections of the data, searching for information not revealed by the covariate structure.

Let $\beta_1, \beta_2, \ldots, \beta_p$ denote p-dimensional unit vectors, as "directions" vectors, the projection pursuit regression (PPR) algorithm [FRI81] finds M_0, direction vectors $\beta_1, \beta_2, \ldots, \beta_{M_0}$ and non-linear transformations $g_1, g_2, \ldots, g_{M_0}$ such that

$$E[Y|x_1, x_2, \ldots, x_p] = \mu_y + \sum_{m=1}^{M_0} a_m g_m\{\beta_m^T x\} \tag{19}$$

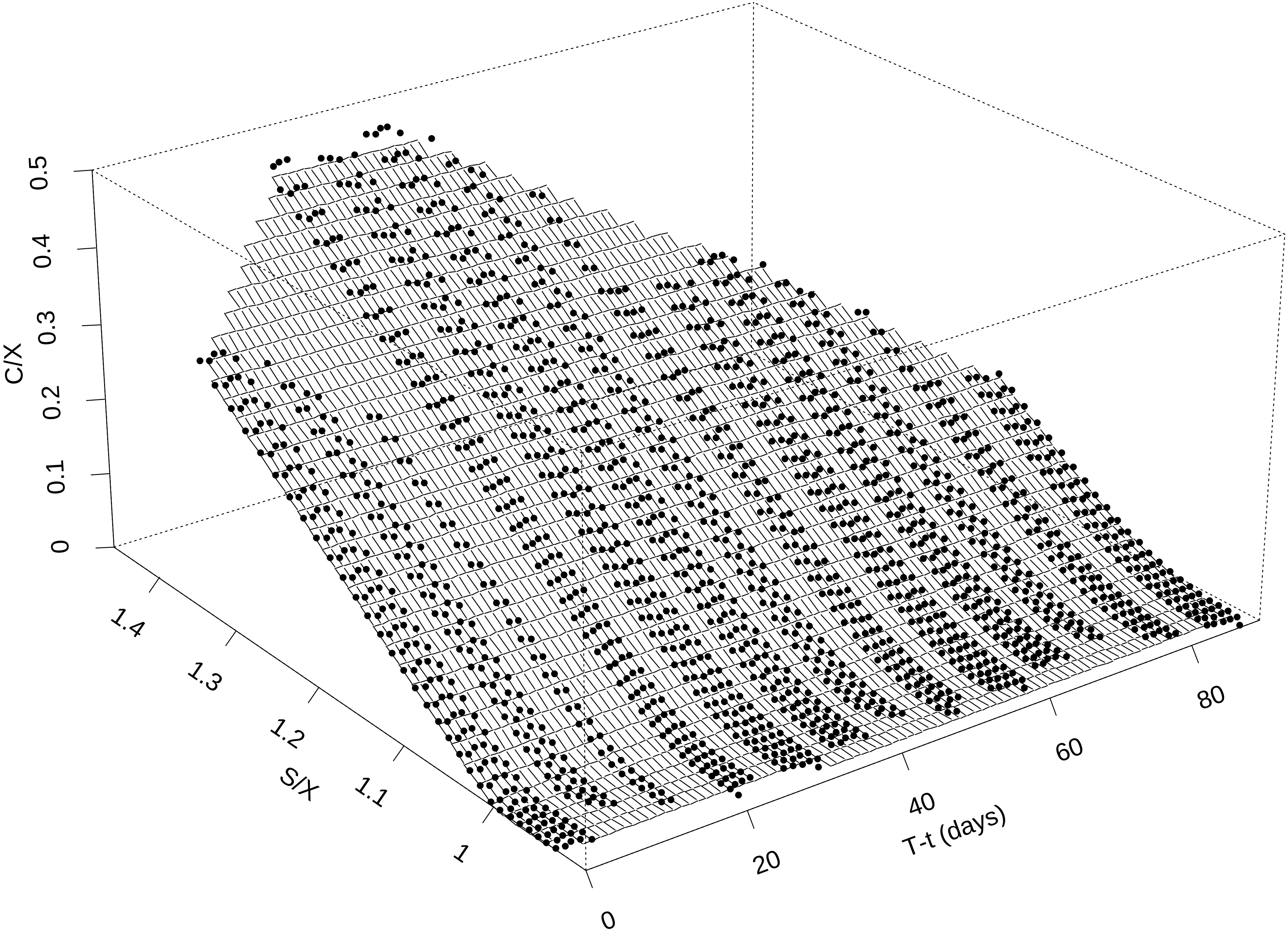
where $\mu_y = E(Y)$, and the g_m have been standardised to have mean zero and unity variance:

$$E g_m\{\beta_m^T x\} = 0, E g_m^2\{\beta_m^T x\} = 1, m = 1, \ldots, M_0$$

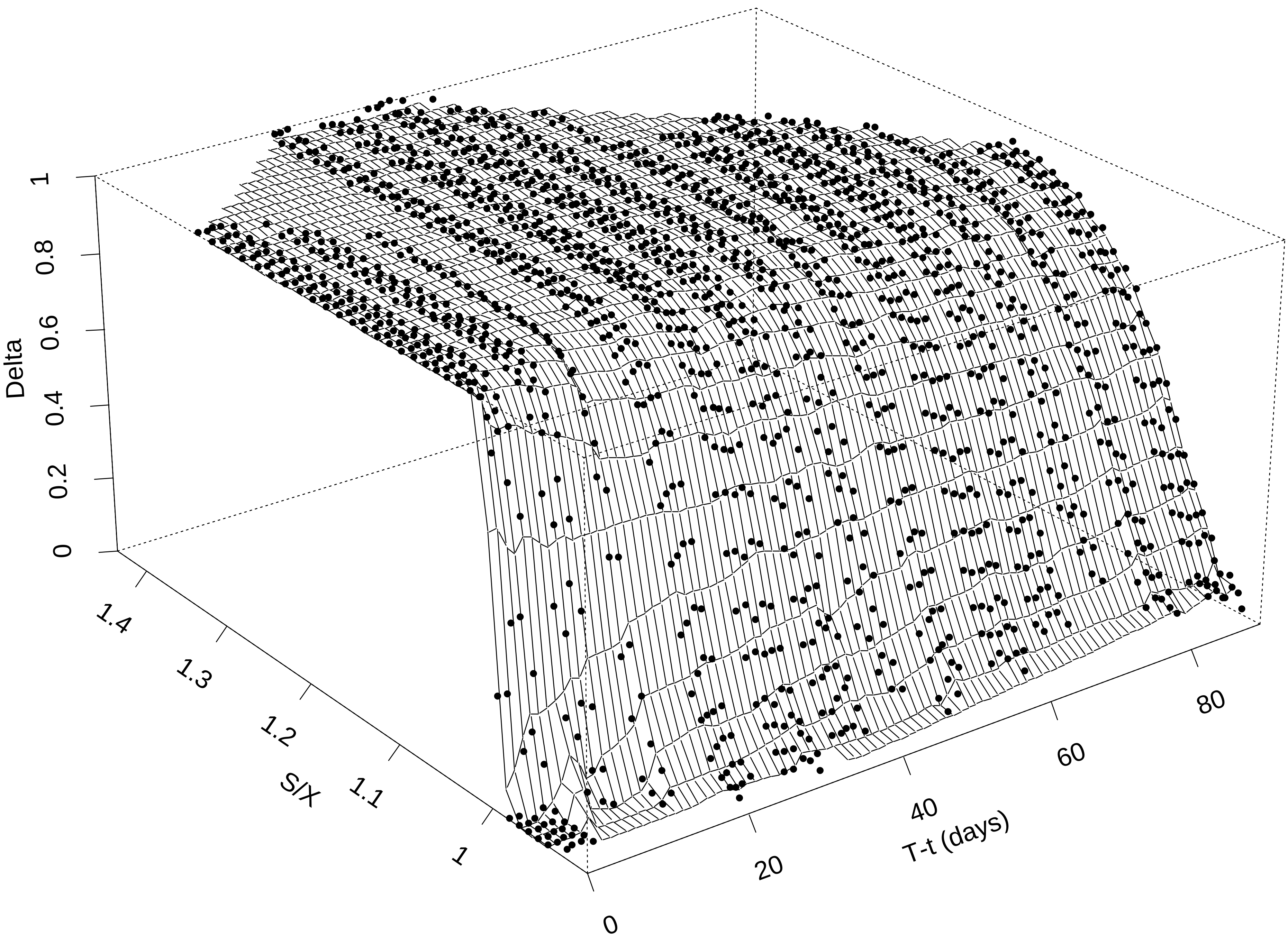
The observations Y_i, $X_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^T$, $i = 1, 2, \ldots, n$, are assumed to be independent and identically distributed random variables. The model parameters a_m, g_m, β_m, m= 1, ..., M_0 in equation (18) minimise the mean squared error

$$E[y - \mu_y - \sum_{m=1}^{M_0} a_m g_m\{\beta_m^T x\}]^2 \tag{20}$$

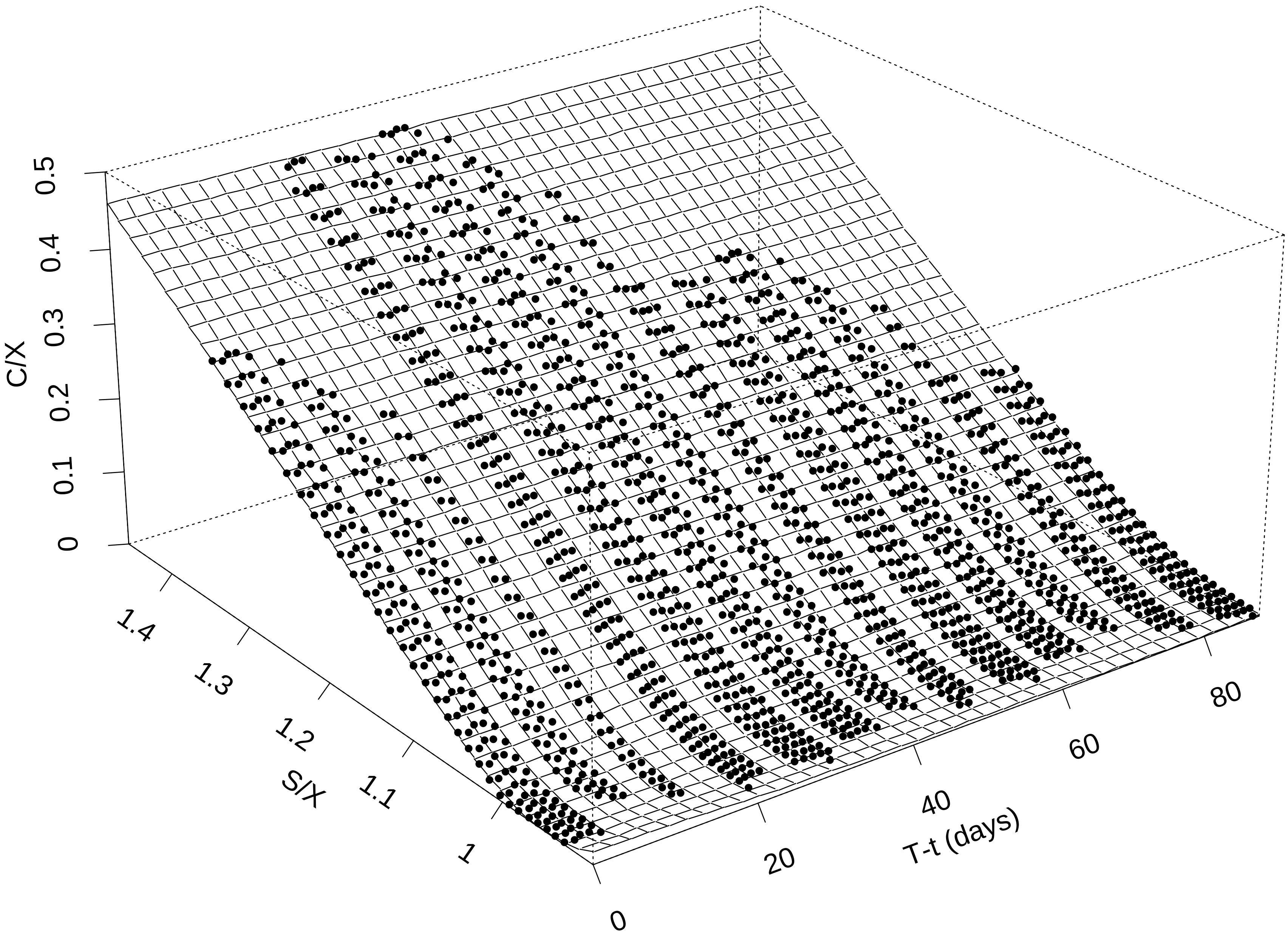
over all possible a_m, g_m, β_m.



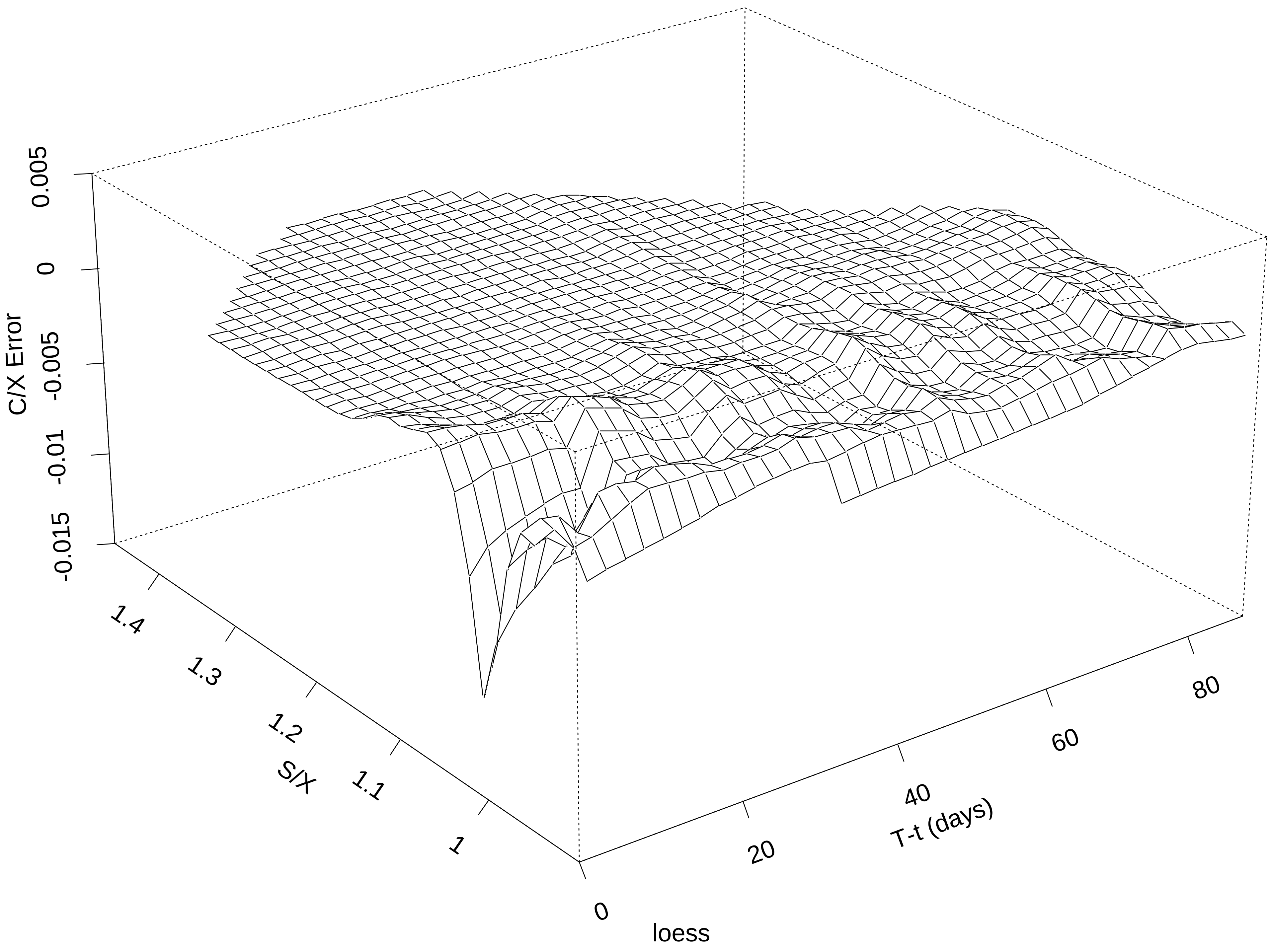
C/X
0
0.1
0.2
0.3
0.4
0.5
S/X
1
1.1
1.2
1.3
1.4
T-t (days)
0
20
40
60
80



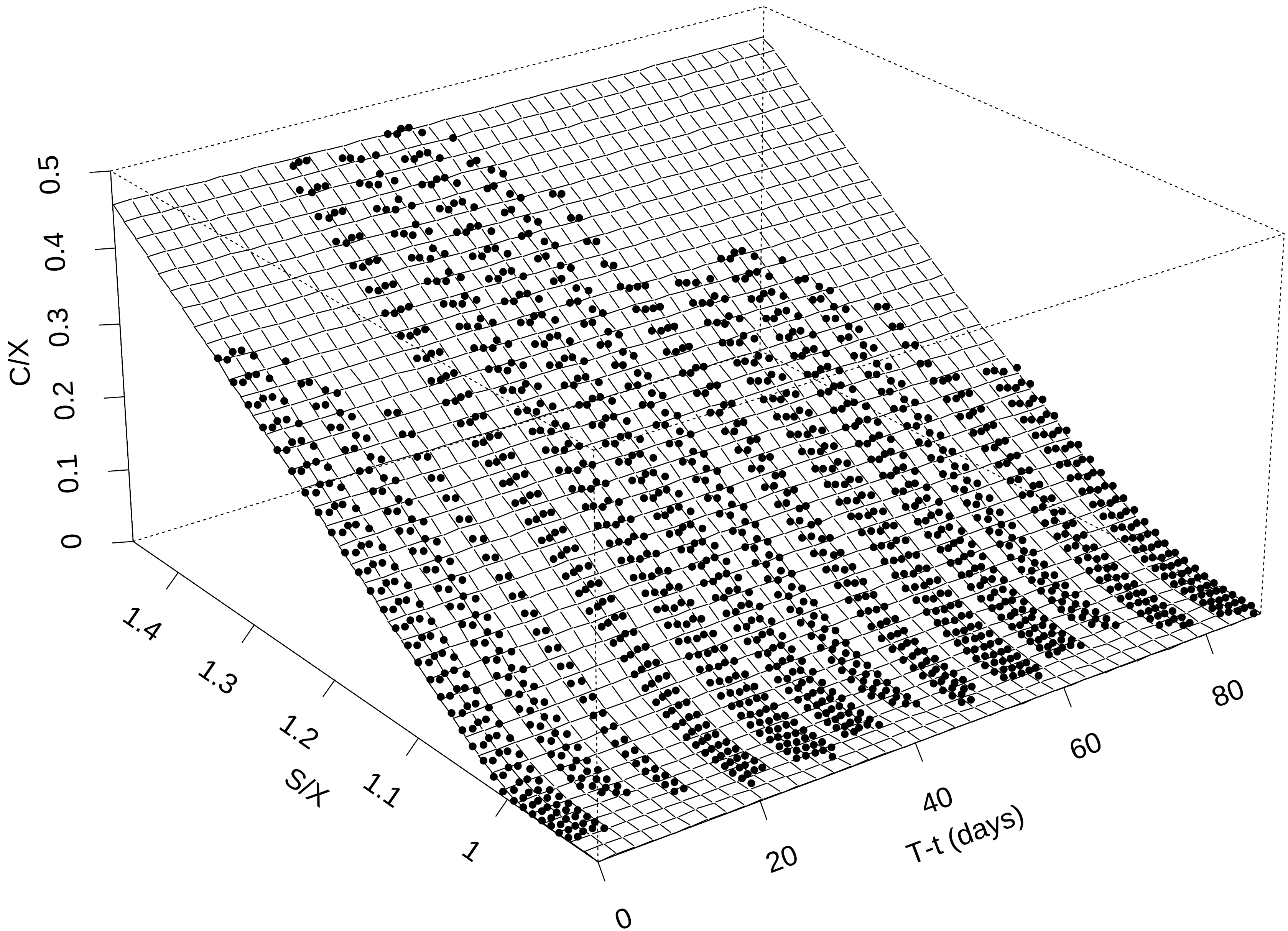
Delta
0
0.2
0.4
0.6
0.8
1
S/X
1
1.1
1.2
1.3
1.4
T-t (days)
0
20
40
60
80



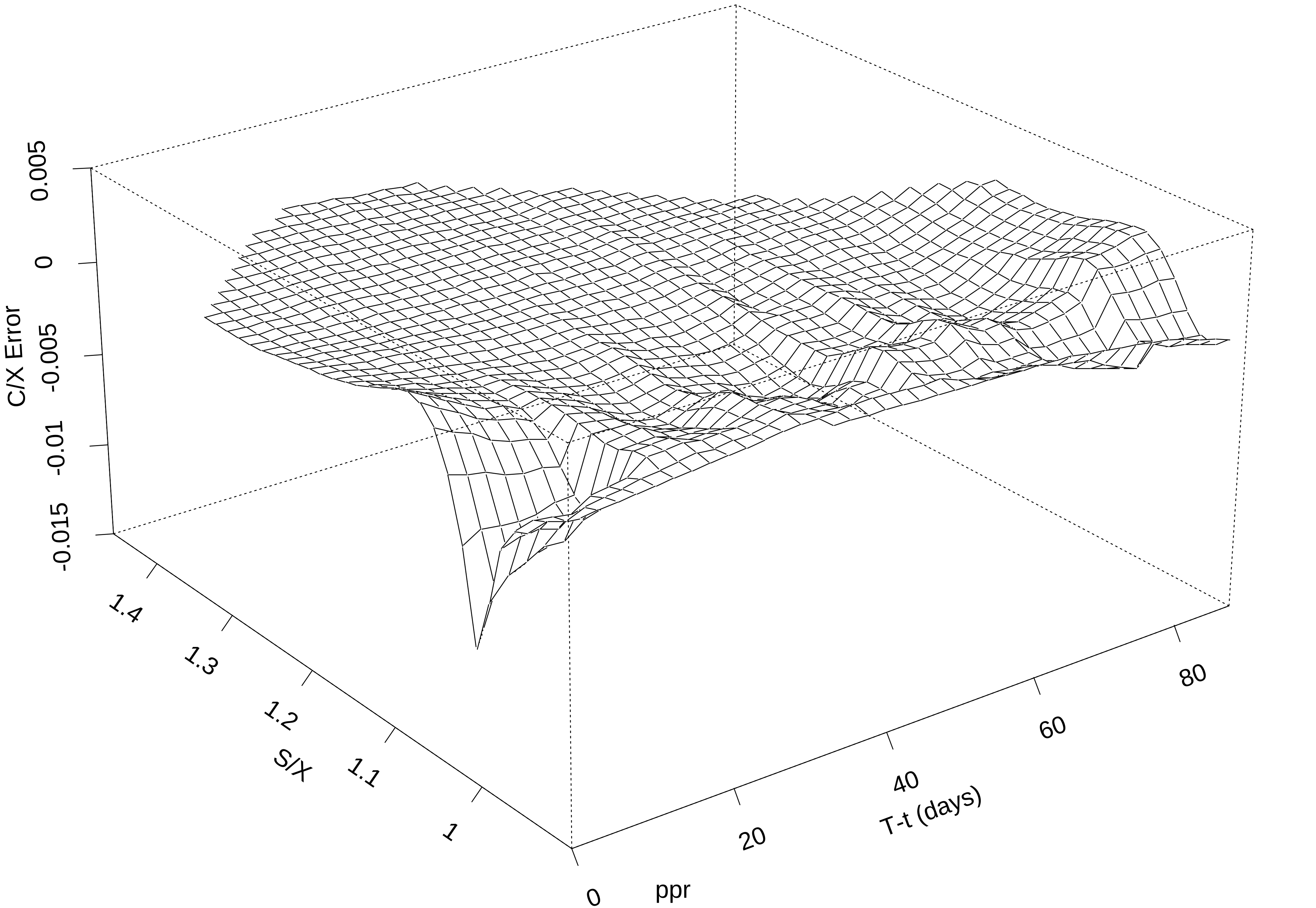
loess



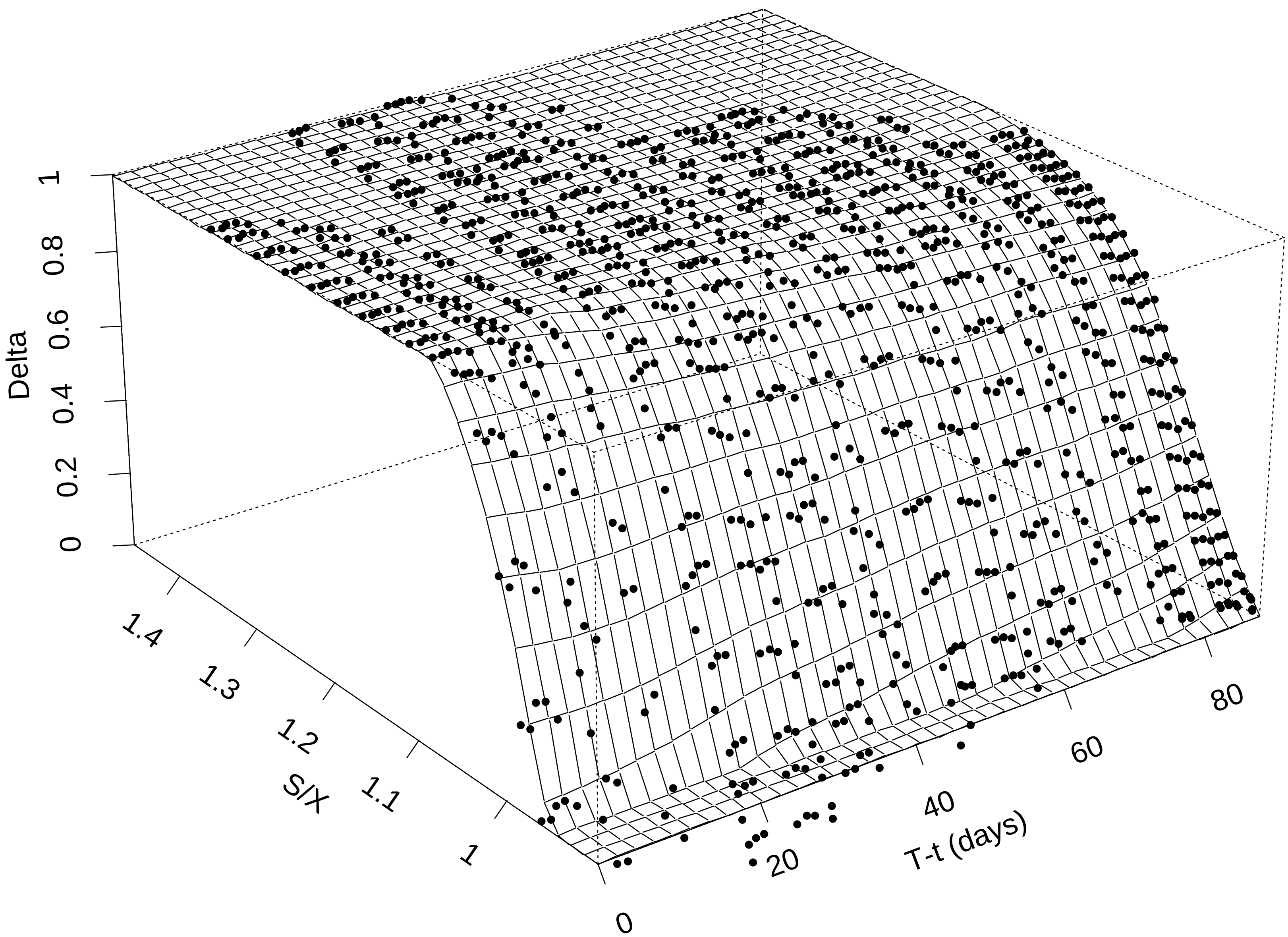
C/X Error
0.005
0
-0.005
-0.01
-0.015
S/X
1.4
1.3
1.2
1.1
1
T-t (days)
0
20
40
60
80
loess



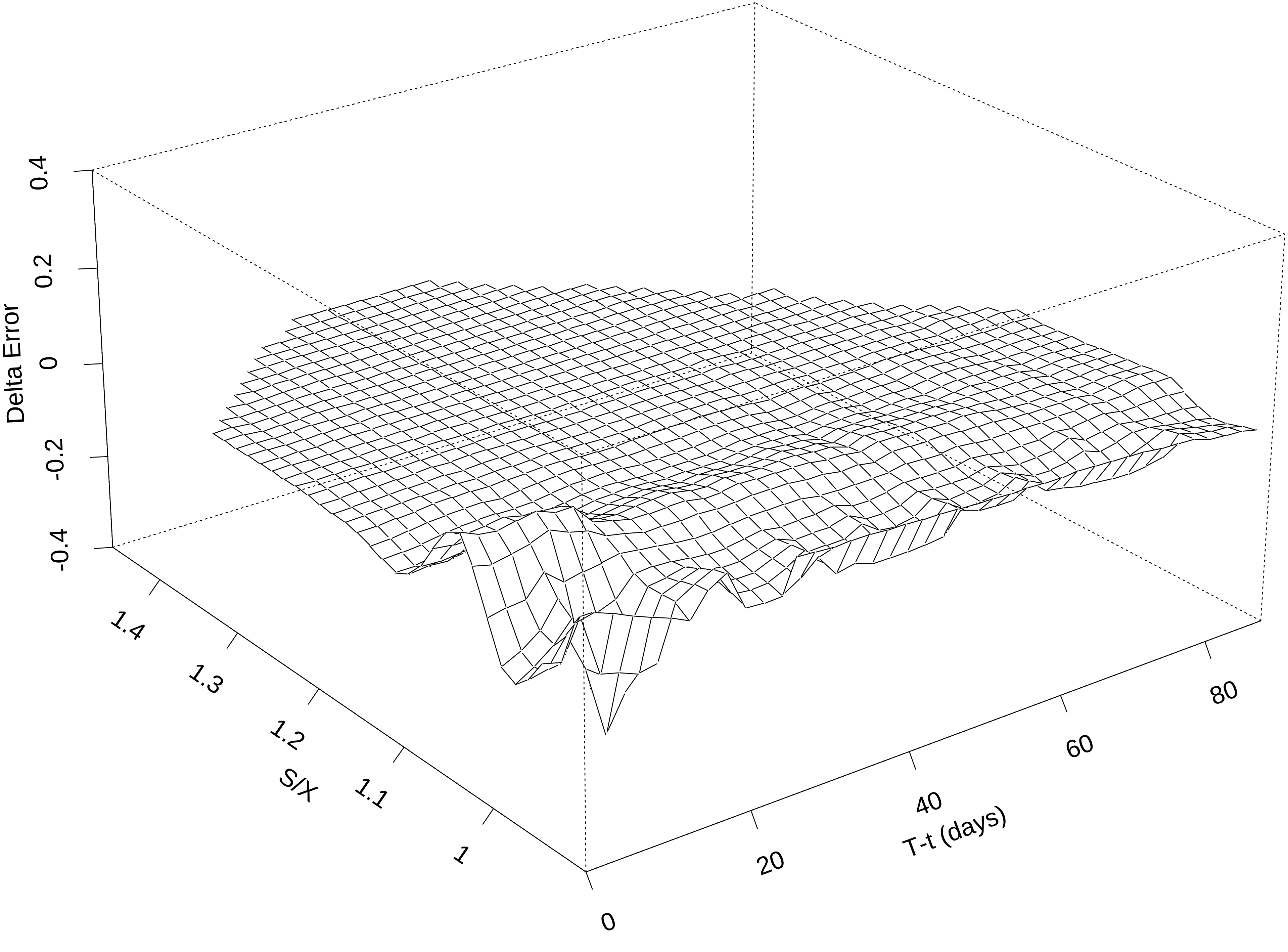
ppr



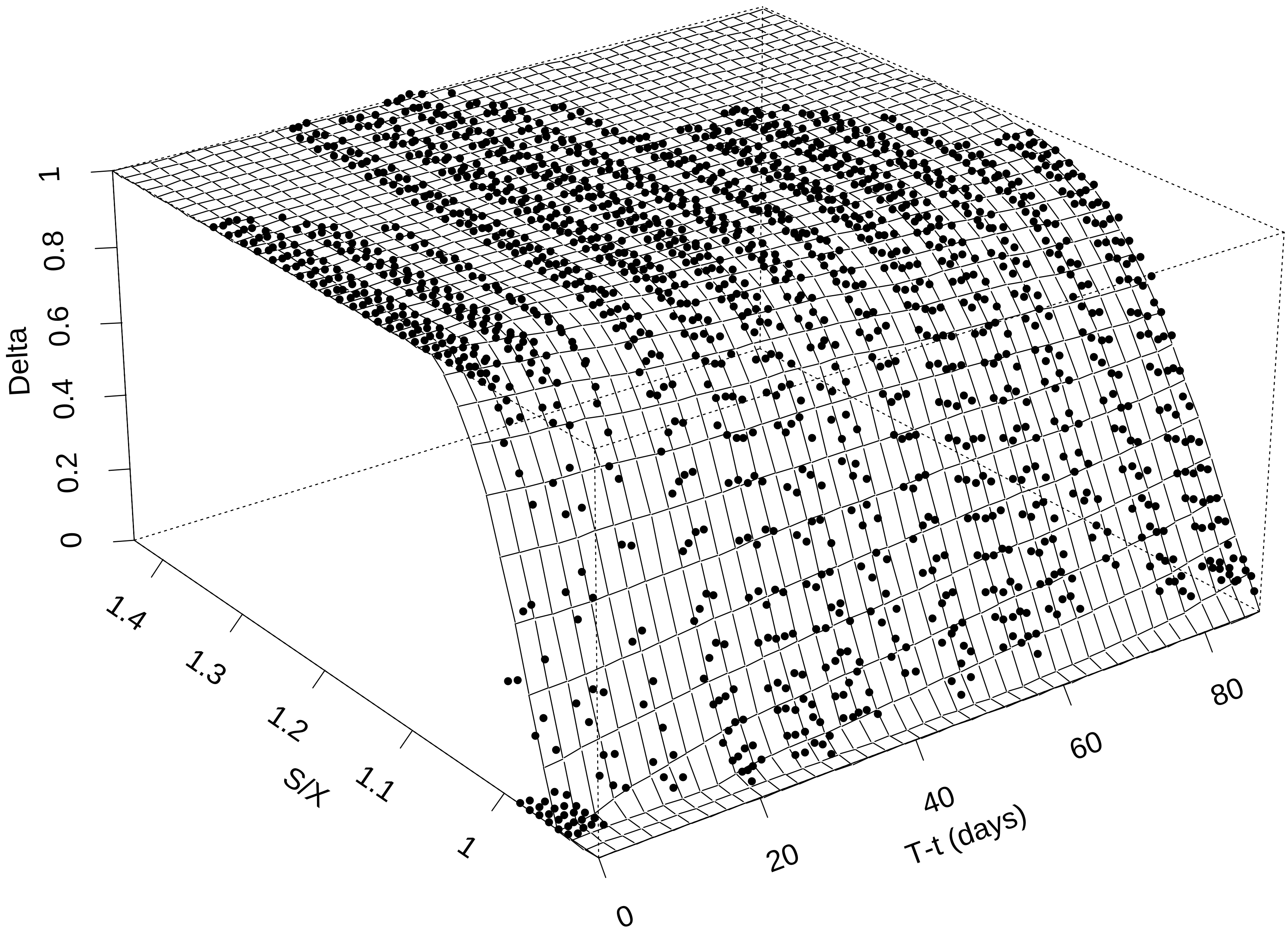
C/X Error
0.005
0
-0.005
-0.01
-0.015
S/X
1.4
1.3
1.2
1.1
1
T-t (days)
0
20
40
60
80
ppr



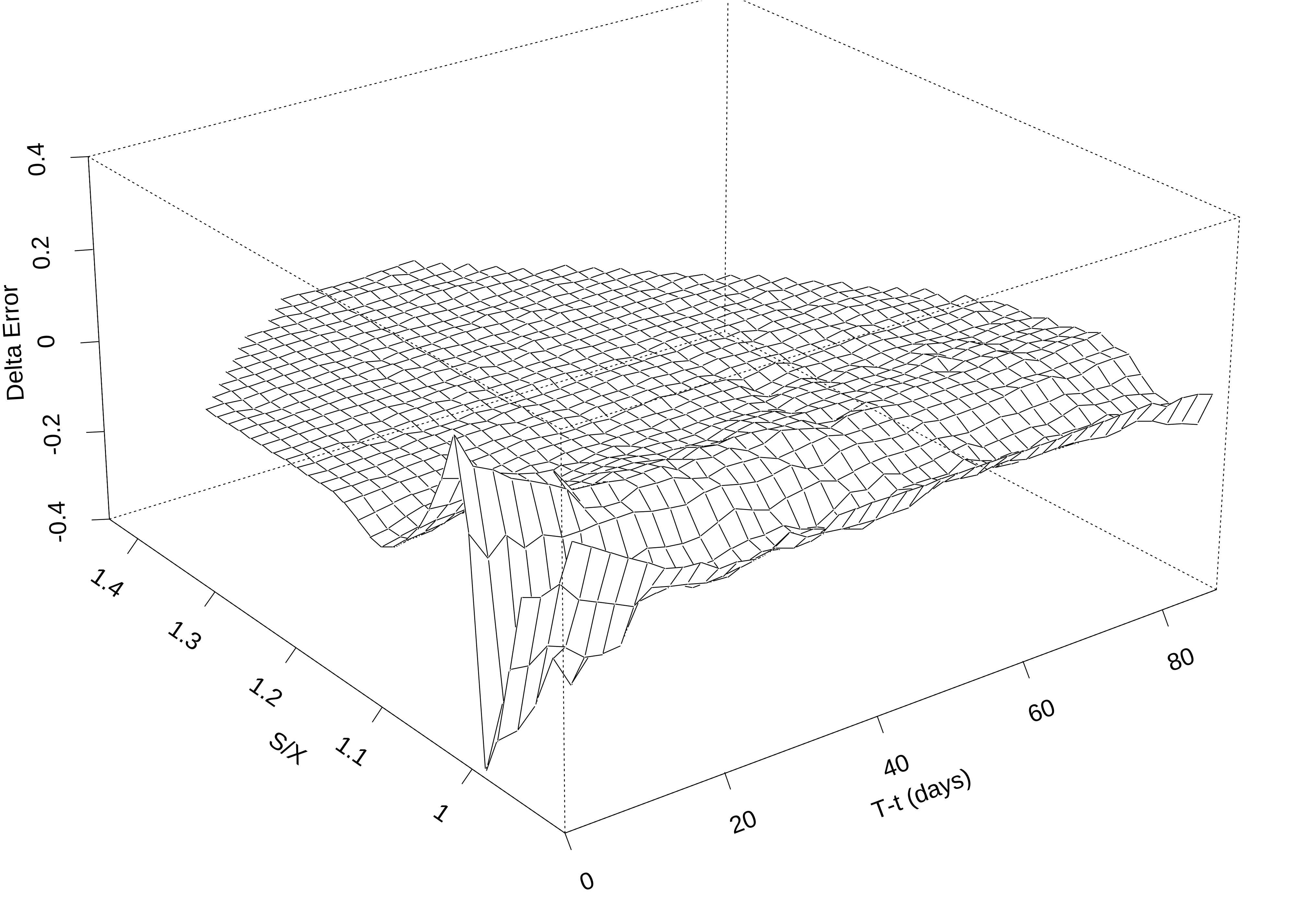
loess+loess



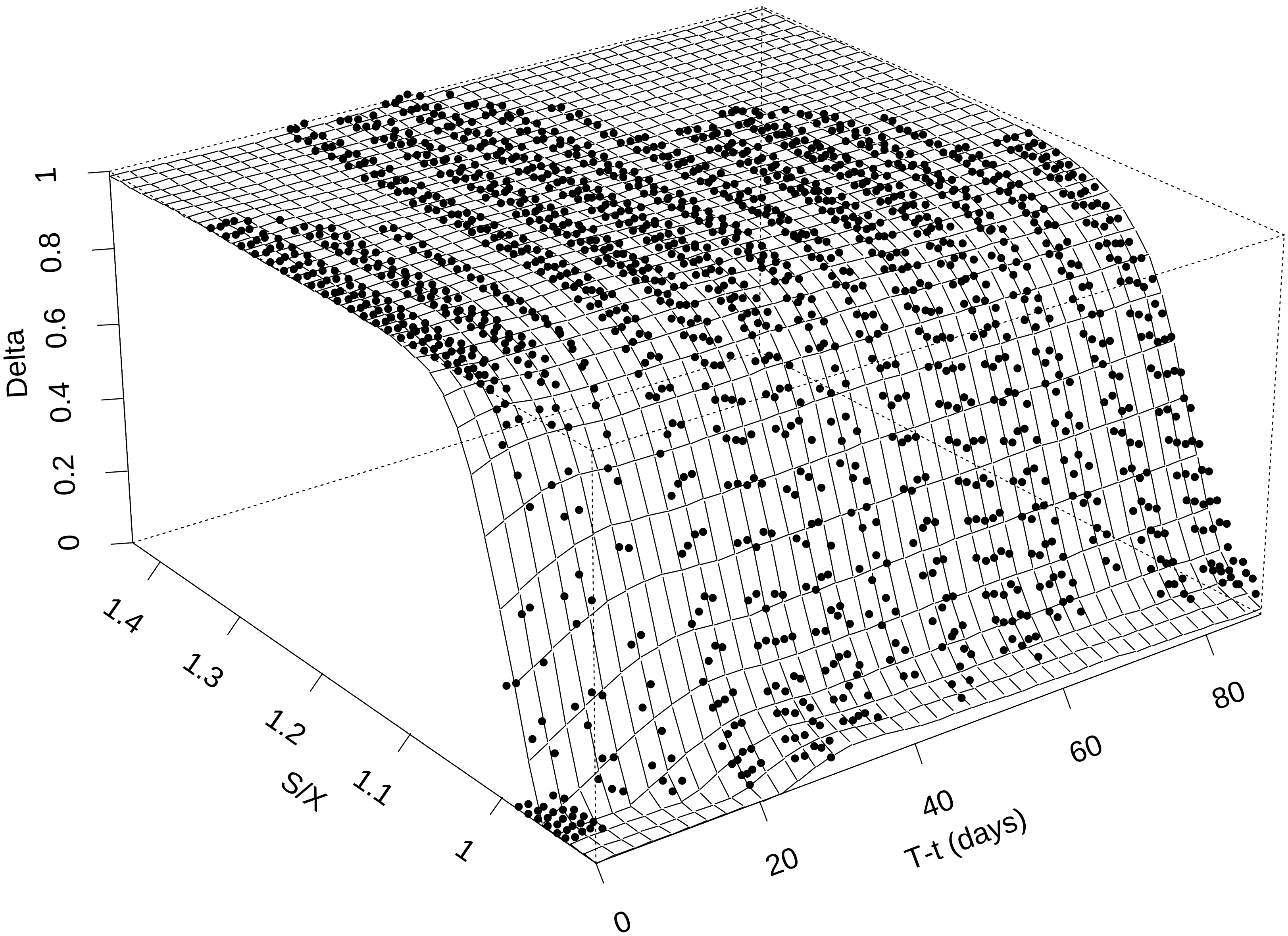
loess+loess



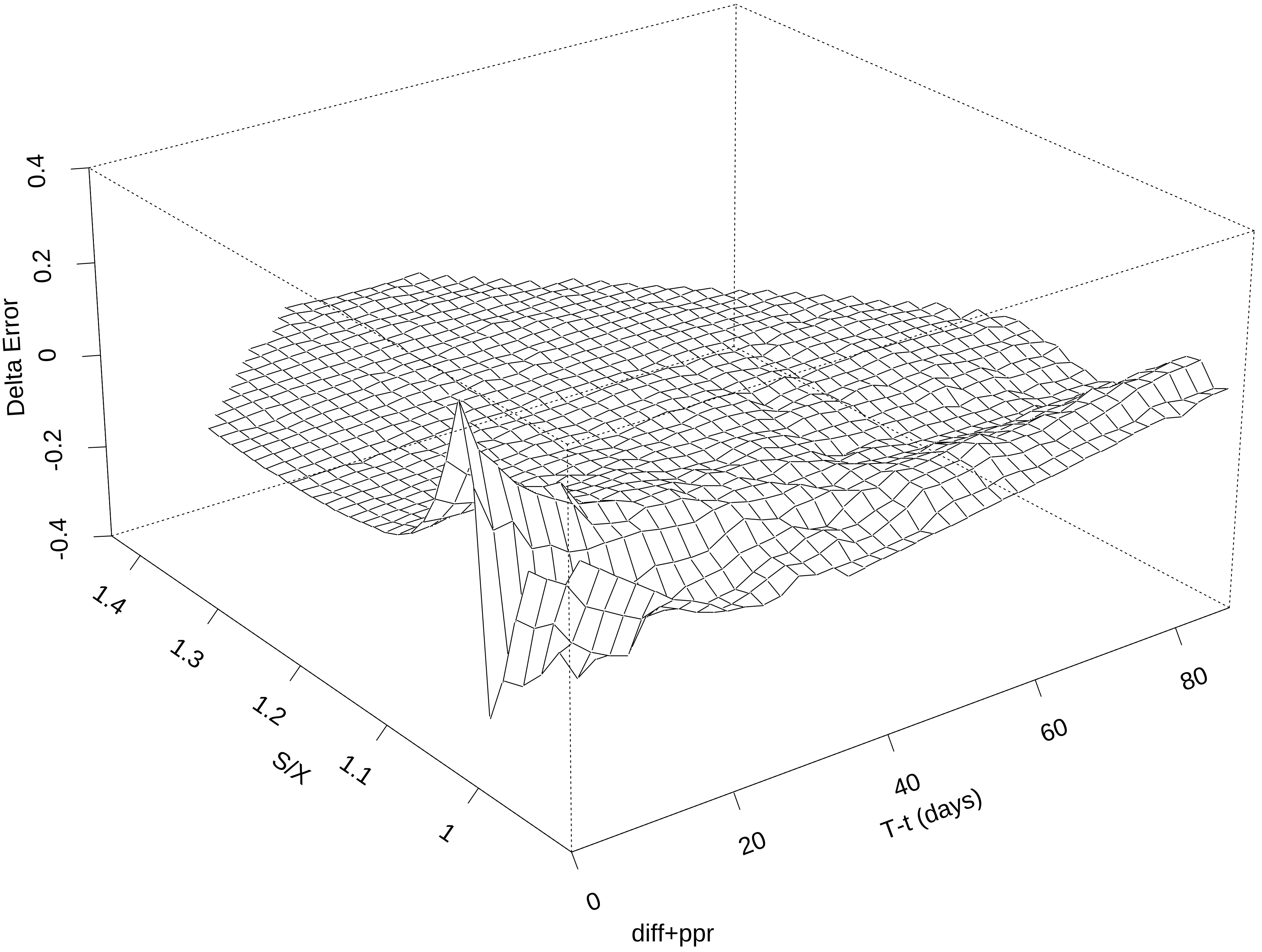
diff+loess



diff+loess



diff+ppr



0.4
0.2
0
-0.2
-0.4
Delta Error
1.4
1.3
1.2
1.1
1
S/X
0
20
40
60
80
T-t (days)
diff+ppr