SMOOTHING WITH SPLIT LINEAR FITS*

John Alan McDonald
and
Art B. Owen

Department of Statistics, Stanford University
and
Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305

Abstract

We introduce a family of smoothing algorithms that can produce discontinuous output. Unlike most commonly used smoothers, that tend to blur discontinuities in the data, this smoother can be used for smoothing with edge detection. We cite examples of other approaches to (two-dimensional) smoothing with edge detection in image processing, and apply our one-dimensional smoother to sea surface temperature data where the discontinuities arise from changes in ocean currents.

(Submitted to Technometrics.)

1. Introduction

In recent years smoothers have become popular with statisticians and data analysts. For example smooth curves superimposed on scatterplots help one to understand the relationship between two variables (Mosteller and Tukey, 1977) and smooth curves estimated from data are used in nonparametric regression models such as projection pursuit regression (Friedman and Stuetzle, 1981). In this paper we introduce a smoother that produces piece-wise smooth curves with a small number of discontinuities in the function or its first derivative. This allows certain desirable features such as jumps, or instantaneous slope changes to be present in the smooth curve. We start with a general discussion of smoothing.

Commonly used smoothers include running averages, running medians (Mosteller and Tukey, 1977), smoothing splines (Wahba, 1984 and references cited there) and running robust regressions (Cleveland, 1977).

Whether we approach smoothing with the objective of summarizing a set of data, or with the idea that we are estimating an underlying function, our requirements are similar. We want a curve that is close to our data and looks 'smooth'.

Closeness to the data is often measured by squared error. Smoothness is more difficult to quantify; sometimes it is incorporated by adding to the sum of squared errors a penalty proportional to the integrated squared second derivative of the function. In this setup the optimal function is a smoothing spline, and the choice of a constant of proportionality governs the tradeoff between smoothness and goodness of fit.

If one is estimating a function in the presence of additive errors essentially the same tradeoff arises in terms of bias and variance. For example, if one estimates with running averages, the fewer points in the average the lower is the bias and the closer to the data is the smooth. Longer windows increase bias, but reduce the variance of the estimate, and produce smoother-looking output.

In some contexts the underlying function is a low frequency signal to which high frequency noise is added. The smoother is then thought of as a low-pass filter, and the design of the filter involves a tradeoff between noise filtered out and signal extracted.

The usual notions of smoothness involve the existence of some number of continuous derivatives, or bounds on norms of derivatives. These ways of quantifying smoothness usually rule out curves with discontinuities, or discontinuous lower derivatives. Thus they are too restrictive for the some smoothing problems. Curves with steps, abruptly changing derivatives or even cusps could easily be appropriate for some data, and such curves are also reasonable underlying functions in some situations. Consider for example: Sweazy's kinked demand curve (Lipsey, Sparks and Steiner, 1976) in microeconomics, the transition of air resistance from a quadratic function of velocity to a linear one at high velocities (Marion, 1970, Sec 2.4), the expected patterns in the fossil record under the punctuated equilibrium theory of evolution (Gould, 1980, Essay 17), or the New Jersey Pick-It Lottery Data (Becker and Chambers, 1984, section 1.2) in which the payoffs tended to be sharply higher for numbers less than 100 because people tended not to select lottery numbers with leading zeroes. In some applications, such as computer vision, the discontinuity (an edge, say) can be the most important part of the function. Most commonly used smoothers blur discontinuities in the function or its first derivative. (Even medians can blur discontinuities if the underlying function is not monotone.)

Our motivation came from looking at data with a continuous smooth superimposed and seeing that it would be improved by incorporating a discontinuity. We then sought smoothing techniques that would automatically put the desired discontinuities into the smooth without unduly sacrificing other aspects of the smooth. Our algorithm does this; the typical output contains a small number of discontinuity features (possibly zero) and is piece-wise smooth.

In the next section we present the split linear smoother, a one dimensional edge-detecting smoother. In section 3 we apply it to data generated by adding noise to a discontinuous function, and compare the results to some other smoothers. We also applied the smoother to daily readings of the sea surface temperature off the California coast. The temperature changes sharply when ocean currents change. This application is illustrated in section 4. Sections 5 through 7 contain comments on asymptotics, related work in image processing, and a summary with conclusions.

2. The Split Linear Smoother

We suppose that we have observations x_i, y_i, $i = 1, \ldots, n$ where $x_j \leq x_{j+1}$, and that we wish to smooth Y on X. That is we want to find a function of the x_i that is close to the y_i and is (piece-wise) smooth as described above.

A set of k successive x's is said to be a window of size k. A linear fit at point i over a window is the value at x_i of a line fit (typically by some kind of least squares) to the (X, Y) pairs in the window.

The split linear smoother begins by obtaining at each point i, a family of linear fits corresponding to a family of windows. Some of the windows are centered on point i (i.e. x_i is the median of the X's in the window), some of the windows have x_i as their left-most point, and some have it as their right-most point. For each of these orientations (left, right, and center) several (typically three to five) window sizes are used.

The smooth at point i is obtained as a weighted average of the linear fits there. The weights depend on a measure of the quality of the corresponding linear fits.

Finally, the above may be iterated. That is, the split linear smoother as described above is applied to its own output.

The reason that this algorithm can find isolated discontinuity features is that on either side of the feature, some of the lines fit well over their windows and the others are affected by extreme bias. The fits from the windows over which the data appear reasonably linear get most, or all, of the weight. This factor will also lead the smoother to put greater weight on the smaller windows in regions where the curvature is highest.

We discuss in turn: the technique for obtaining the fits, the goodness of fit criterion, the weighting function, the reason for iteration, and the design issues raised in choosing the windows.

We used the same set of window sizes for each window orientation and fit the lines to the windows by ordinary least squares. This way the typical fitted line provides three linear fits; a left-sided fit for the right-most point, a central fit, and a right-sided fit for the left-most point.

3

The use of ordinary least squares allows us to use updating formulae to fit all the lines for a given window size by 'passing the window over the data'. Using a (non-rectangular) kernel of weights would take a lot more computation, unless updating formulae could be found for the particular kernel used. The price we pay for this speed is that the ordinary least squares fits (as a function of x) are noisy, since points are added and dropped with unit weight at the ends of the windows where their influence is greatest. We raise this issue again when we discuss iteration.

The windows are shrunk as they approach the ends of the data, left-sided windows are not used for the left-most few points, and similarly at the right end of the data.

We based the smoother on linear fits because the usual running linear smoothers tend to provide better results than running averages (especially regarding end effects) and our experience with higher order local polynomials indicates that they behave erratically, probably because the extreme points get even greater influence than in a linear fit.

The fits were assessed on the basis of the mean squared residual about the line, taken over all points in the window except the target point. (We will use 'pmse' to refer to this pseudo-mse.) The target point is left out of the averaging to reduce the tendency of a linear fit to look good simply because it came close to the data point, and cause the resultant smooth to capture more of the noise at that point. (The point must be left in the fitting of the line, or else fits from the 'wrong' side of the target point will get high weights.) We tried cross-validated squared errors as a means of assessing the linear fits but found that the resulting smooths had a very jagged appearance. This could be because the cross-validated squared errors all depend so strongly on y_i. It may be that the use of one-sided windows exacerbates this problem beyond what it would be for combinations of central fits.

The split linear smooth value at point i is a weighted average of the linear fits there, with higher weights corresponding to better fits. If the pseudo-mse exceeds a cutoff value, then the associated linear fit gets zero weight. This way the split linear smoother can put all its weight on one side, and for example, exactly reproduce step functions. This scheme includes simply picking the fit from the best fitting line. We strongly recommend against that, because, where neighboring fits come from windows of differing size and/or orientation spurious bumps are added to the output. (A few large isolated discontinuities are acceptable, whereas a large number of small discontinuities are not.) More generally, the weight function should decline smoothly to zero as the pmse increases to the cutoff. Otherwise if a single window is just barely cut off in one point and just barely included in a neighboring point a spurious bump may result.

We chose to use weights proportional to the square of the difference between a window's pmse and the cutoff value for those pmse's below the cutoff. Other functions, such as powers higher than 2.0 would make a still smoother transition between zero weights and small non-zero weights at the expense of creating large weight differences for small pmse differences in some other range of pmses. A minor benefit of using an integral power is that it is fast to compute.

The cutoff pmse at a point was taken to be the average of the pmse's from all the fits at that point. Larger cutoffs would make the weights more nearly equal providing smoother looking output at the expense of blurring more discontinuity features. Similarly, smaller cutoffs trade off smoothness to find more features.

4

Because the rectangular windows we use each produce noisy fits, there is a tendency for the split linear smoother to produce output with a somewhat jagged appearance. To alleviate this we iterate the smoother; that is we apply the algorithm described above to its own output. We find that one such application tends to remove the noise. This is a small computational price to pay since we use the updating formulae to compute the linear fits. We also find that the iteration tends not to erode the discontinuities found in the first pass. Most other smoothers would reduce curvature and blur discontinuities at every iteration.

The curves produced by this algorithm tend to be piece-wise smooth. The size of the resulting pieces is governed by the sizes of the windows. The tendency is to produce pieces that are larger than all or most of the windows used. It is unlikely to produce pieces that are smaller than all the windows. The frequency with which piece-sizes between the smallest and largest window sizes will arise depends in part on the cutoff value used.

Using a large number of windows of slightly varying size (e.g. a large number of consecutive odd integral sizes) tends to produce smoother looking output, mimicking the effects of non-rectangular kernels.

Other ways of orienting the windows, such as putting one third of the data on one side of the target and two thirds on the other, were not used because it was thought that most of the relevant information would come from the central and extreme windows.

The split linear smoother is similar to the Supersmoother (Friedman and Stuetzle, 1982), except that the latter uses only centered windows and uses a somewhat different way of combining the basic fits.

3.   Simulated Examples

Figure 1 shows a sawtooth function with Gaussian noise added at $n = 256$ equispaced points between 0 and 1. The function consists of two line segments rising from 0 to 1. The standard deviation of the noise is one half that of the function. Superimposed on it is a central smooth (like the Supersmoother) based on only the central fits in the split linear smoother. (Three windows sizes, .2n, .3n, and .4n were used.) The curve is qualitatively smooth but, as is no surprise, blurs the discontinuity. Figure 2 shows the same data smoothed by running medians of 11 points. There the curve has no trouble finding the discontinuity, but appears very rough. The split linear smoother applied to this data is illustrated in Figure 3. It found a curve that has the discontinuity *and* is smooth.

The experiment described above was done 1000 times. Figure 4 shows some pointwise quantiles for the central smoother. The squares represent the true values (at every fifth point), and are drawn on the sawtooth curve. The outer envelope consists of the 5th and 95th percentiles of the 1000 smooths. The inner envelope is obtained from the quartiles and the central line is the pointwise median. Figures 5 and 6 present the same information for the median smoother and the split linear smoother respectively. From Figure 4 we can see that the ensemble of central smooths miss the discontinuity. From Figure 5 it appears that the ensemble of running medians does not miss the discontinuity, and neither does the ensemble of split linear smoothers.

We might also be interested in the width of the quantile envelopes. Those of the central smoother are generally the narrowest, and those of the running median are the widest. The

central smoother only uses three central windows (of differing size). The split linear smoother uses the same three windows, and six more windows (one of each size on each side). It is in that sense using more parameters than the central window, and so it is to be expected that the results are less concentrated. The running median smoother can be made to have much narrower quantile envelopes and a much smoother appearance by increasing the span, but then it would badly miss the discontinuity.

Another property of interest is the difference between the true function and the median of the smooths. This feature is a form of bias for the smoother, while the width of the intervals is related to the variance. The central smoother is severely biased near the discontinuity whereas the median smoother and the split linear smoother are mildly biased there.

While bias and variance are good optimality criteria in one sample location problems they do not tell the whole story in smoothing problems. A smoother could do well by both of these criteria, and yet never look much like the underlying function in qualitative terms. Other important criteria are: whether the smoother displays or blurs discontinuity features, whether the locations and magnitudes of such features are approximately right, and whether the smooth has roughly the correct curvature.

Although we don't show it in the figures, the central and split linear smooths both tended to be smooth (over the ensemble) and the running median smooth tended to be rough (over the ensemble). Thus the split linear smoother was the only one to get both features right, most of the time.

In addition to preserving jumps, the split linear smoother does well in regions of high curvature, that are still smooth in the analytic sense. (Most commonly used smoothers severely reduce such curvature.) To show this, the above experiment was reproduced for the function $sin^3(2\pi x^3)$ on $0 \leq x \leq 1$. As before, the standard deviation of the noise was half that of the function, and the same window sizes were used. Figures 7 through 12 show the same results for this function as Figures 1 through 6 do for the sawtooth.

The central smoother (Figures 7 and 10) provides smooth looking output, and as is typical, reduces the curvature, especially at the trough. It especially underestimates the depth of the trough, all the time.

On this data the median smoother (Figures 8 and 11) provides very rough looking output but in the ensemble, tracks the function with low bias but high variability, even to the extent of catching the inflection. The other smoothers do not catch the inflection because, unlike the median smoother, all their windows are large compared to the region over which the inflection occurs. (The variability and roughness can both be removed by increasing the window size, but then the ensemble ceases to track the function as well.)

The split linear smoother (Figures 9 and 12) produces smooth looking output that does not severely reduce the curvature. It has very good ensemble behaviour at the deep trough. It tends to slightly exaggerate the curvature at the 'first bend'. The quantile bands are narrower than those of the median smoother, but wider than those of the central smoother. An interesting feature of the ensemble of split linear smooths is the small 'goatee' in the lower quantiles near the right of Figure 12. This feature occurs at the boundary between points for which right sided windows were used and points for which there were deemed to be too few points to the right to fit a window. We have left this feature in to illustrate the importance of smooth transitions between zero and non-zero weights; weighting schemes that make abrupt

6

transitions can (and do) sprinkle such jagged features throughout the whole smooth. The transition can be smoothed out by applying a gradually increasing penalty to windows as their size decreases to the minimum.

We see from this experiment that the split linear smoother is able to find discontinuous features and curvature and produce smooth output in moderately large but noisy samples.

4. Smoothing the Sea Surface Temperature

The data are daily measurements of the temperature of the ocean off Granite Canyon, near Monterey California, over the period from 1 March 1971 to 28 February 1983. This and related data is discussed in Breaker, Lewis, and Orav (1983). The data appear noisy, but some features are discernable. There is a strong yearly cycle, so that summer and winter are easily identifiable. There are three el Niño years (including 1983), and they have hotter summers than the other years. Some years there is a cooling in the middle of the summer, some years not. The temperature peaks do not come at the same time each year. The feature most relevent to our discussion is the called the spring transition. Between February and April the ocean currents may change, causing an upwelling of cold water and a temperature drop of as much as 4 degrees Celsius in a few days. (It may otherwise take months for a change of this magnitude.) The transitions don't come every year, and in the years they do come, they don't always come at the same time, cause the same size of drop or take the same number of days to cool the surface. There are 6 spring transitions in the data, and also some autumn warming transitions. Such irregular seasonal behaviour is a serious problem for most time series techniques. If one wanted to take out the deterministic component, possibly as a preliminary to some standard time series analysis, it would be necessary to find a representation that preserved the gross features of the data such as its curvature, and the location and magnitude of the discontinuities.

The data are plotted as Figure 13, and a split linear smooth is superimposed on them. The smooth used window sizes 61, 91, and 121. While these window sizes are small as proportions of the data set they are large compared to the rate at which the temperature changes. The resulting smooth helps us to pick out the seasons, assess which years were the hot ones, and generally see where the data go.

Figure 14 is a close-up of a spring transition, with the smooth shown. Even though the smallest window size is 61 days, the smoothed temperature drops sharply. Figure 15 is a close-up of the third year in the data, the one in which summer seemingly never came. This year also exhibits a sharp temperature drop.

At such close range it becomes evident that what looks like noisy data on the scale of Figure 13, actually has a finer level of structure. The temperature goes through a steady alternation of temperature for most of this year. In Figure 15 the same year is shown with another split linear smooth. This one used window sizes of 11, 21, 31, 41, and 51. It fits especially well to the last few months shown. The earlier months appear to have some slightly finer structure. Similar finer structure persists over the 12 years. The choice of which smooth to use is not based on whether the smaller dips are 'really in the data' (they are), but instead on a determination of the scale on which we want to see the data.

In choosing window sizes one must remember that the resulting smooth will tend be

piece-wise smooth, and the pieces will tend to be larger than all or most of the window sizes.

5. Analytic Properties of S.L.S.

It would be desirable to have a smoother with the same capabilities as the split linear smoother but that is simpler conceptually and for which mathematical analysis would be more tractable. In our view the right theorems to prove are related to the probabilities of detecting the presence or absence of discontinuities and other features in finite samples. The experiment in the previous section suggests that the split linear smoother does well by these measures.

There may be some concern that one-sided estimates used as components in a smoother may be subject to certain pitfalls. We offer the following example to show that one-sided estimates can have asymptotic properties comparable to central ones. Suppose that we have pairs of reals $(x_i, y_i), i, \ldots, n$ and that at each x we estimate the conditional mean of Y given $X = x$ by a symmetric kernel that puts non-zero weight only on each of the first k_n points to the left of x and similarly on the right of x. Suppose that the i'th points to either side get the same weight. (Some modification is made to the left of the $k_n + $ 1st point and also at the right side of the sample. Typically k_n is a $o(n)$ as $n \to \infty$, so this modification is slight.) Finally suppose that all the kernel weights are non-negative and sum to unity at each x. In short it is a symmetric kernel running average. If this procedure is consistent in the sense of Stone (1977, p. 597), and we modify it at each x between x_{k_n+1} and x_{n-k_n} by doubling all the left-sided weights and setting all the right-sided weights to zero we get a left-sided kernel smoother. All of the weights used in this smoother are by construction between 0 and twice the corresponding weights in the symmetric kernel running average. It follows from Corollary 2 of Stone (1977) that this left-sided smoother is also consistent. (The consistency involved is in $\mathcal{L}^r$ for all $r > 1$ for which the r'th moment of Y is finite.) A similarly obtained right-sided kernel smoother would also be consistent. Even knowing the underlying mean function and chosing whichever fit (left or right) is worst is consistent. Conversely if the left and right sided smoothers are both consistent then so is their average, and hence so is the central smoother.

The situation is more complicated when linear fits are used, but we conjecture that under reasonable conditions on the regression of Y on X and the marginal distribution of X, that the split linear smoother will be consistent.

6. Related Work

The Supersmoother of Friedman and Stuetzle (1982) is similar to the split linear smoother, but uses only central smooths.

In the image processing literature similar algorithms have been developed for smoothing two-dimensional images. Scher, Velasco, and Rosenfeld (1980) consider the eight nearest neighbors of a point in a square grid. They use all eight different 'triangular wedges' consisting of the target point and three of its neighbors. They try various ways of combining averages of the points over the neighborhoods, iterating each procedure, and report on the relative merits of each method.

Nagao and Matsuyama (1979) also consider small square neighborhoods about each pixel.

They rotate an 'elongated bar mask' through each neighborhood, with one end fixed at the target pixel. Whichever position gives the minimum variance fit over the block is used to find the value for the target point. This procedure is iterated.

Haralick and Watson (1979) fit polynomials in the row and column variables over each of the K^2 blocks of dimension $K * K$ that contain the target point. They take the value from the block with least residual squared error, and iterate. They use $K = 3$ and linear or constant polynomials but give the least squares formulae for general K.

The window sizes used in image processing would seem small to most statisticians. Their advantage in image processing is that they make it easier to build specialized hardware for parallel implementation. They also lower the computational cost per iteration. Their drawback is that they do not provide output that is as smooth as one gets with larger windows. We think that two dimensional variants of the split linear smoother that operate on larger windows than those presently used could be useful in image processing.

7. Summary

We have described a smoother, based on running linear fits capable of producing curves with discontinuities, continuous curvature, and qualitative smoothness (between the discontinuities). The main idea is to take, at the i'th point in a sample, a weighted average of linear fits based on windows of various sizes and orientations; some of the windows are centered on i, others are entirely to the left of i, and still others are to the right of i.

The split linear smoother is shown to be superior to a similar central smoother and to running medians when it comes to reliably finding certain important qualitative features such as smoothness, discontinuities, and troughs in moderately large samples.

These results depend on the test function used. For example if there were no sharp features (e.g. the underlying function is constant) the split linear smoother in looking for them is essentially using more parameters than a central smoother and should therefore reproduce more noise or possibly create more artifacts. It should also be mentioned that medians with large spans would do very well on discontinuities if the underlying curve were monotone.

The smoother is illustrated on a daily record of sea surface temperature at Granite Canyon California. It produces a smooth version of the temperature without blurring the sudden temperature changes caused by changing ocean currents.

The split linear smoother provides an approach, in the one dimensional case, to the problem of smoothing without blurring boundaries (i.e. points at which the curve or its derivative are discontinuous). It could therefore be used to search for and quantify change points. With a small number of smooths, the discontinuities are easily seen in plots. If the process must be automated, the weights that determined the smooth could be used.

It is clear that other similar smoothers could be developed, but some features of the split linear smoother seem to the authors to be very important. The first such feature is the use of windows that are not centered. (Imagine running models over the data in which each form of non-linearity including curvature, cusps, and jumps could be assessed in every window.) The second point is that there must be a way of putting weight zero on all the windows for which a linear model is not appropriate. The third point is that weighting windows is better than

selecting windows, and more generally, the weights used should not change abruptly between neighboring points except at a discontinuity (otherwise one gets a large number of small steps in the output).

8. Acknowledgements

We thank Dr. P.A.W. Lewis for bringing to our attention the sea surface temperature data, and providing a copy of it.

9. References

Breaker, L.C., Lewis, P.A.W. and Orav, E.J. (1983) *Analysis of a 12-Year Record of Sea-Surface Temperatures Off Pt. Sur, California* Tech. Report NPS55-83-018 Naval Postgraduate School, Monterey CA

Cleveland, W.S. (1979) *Robust Locally Weighted Regression and Smoothing Scatterplots,* JASA 74

Friedman, J.H. and Stuetzle, W. (1981) *Projection Pursuit Regression,* JASA 76

Friedman, J.H. and Stuetzle, W. (1982) *Smoothing of Scatterplots,* Dept. of Statistics Tech. Rept. Orion 3, Stanford University.

Gould, S.J. (1982) *The Panda's Thumb,* W.W. Norton and Company Inc., New York

Haralick, R.M. and Watson, L. (1979) *A facet model for image data,* Proc. PRIP-79 *(Proc. of the IEEE Computer Society Conference on Pattern Recognition and Image Processing, 1979).*

Lipsey, R.G. , Sparks, G.R. and Steiner, P.O. (1976) *Economics, 2nd. Ed.,* Harper and Row, New York

Marion, J.B. (1970) *Classical Dynamics of Particles and Systems, 2nd. Ed.,* Academic Press, New York.

Mosteller, F. and Tukey, J.W. (1977) *Data Analysis and Regression,* Addison Wesley.

Nagao, M. and Matsuyama, T. (1979) *Edge preserving smoothing, Computer Graphics and Image Processing.*

Scher, A., Velasco, F.R.D., and Rosenfeld, A. (1980) *Some New Image Smoothing Techniques, IEEE Trans. on Systems, Man, and Cybernetics, SMC-10.*

Stone, C.J. (1977) *Consistent Nonparametric Regression* The Annals of Statistics, Vol.5,No.4, 595-645

Wahba, G. (1984) *Cross-Validated Spline Methods for the Estimation of Multivariate Functions from Data on Functionals* Statistics: An Appraisal, Proceedings 50th Anniversary Conference Iowa State Statistical Laboratory David, H.A. and David, H.T. , Editors The Iowa State University Press

Figure 1: A Central Smooth of a Sawtooth Function with Noise
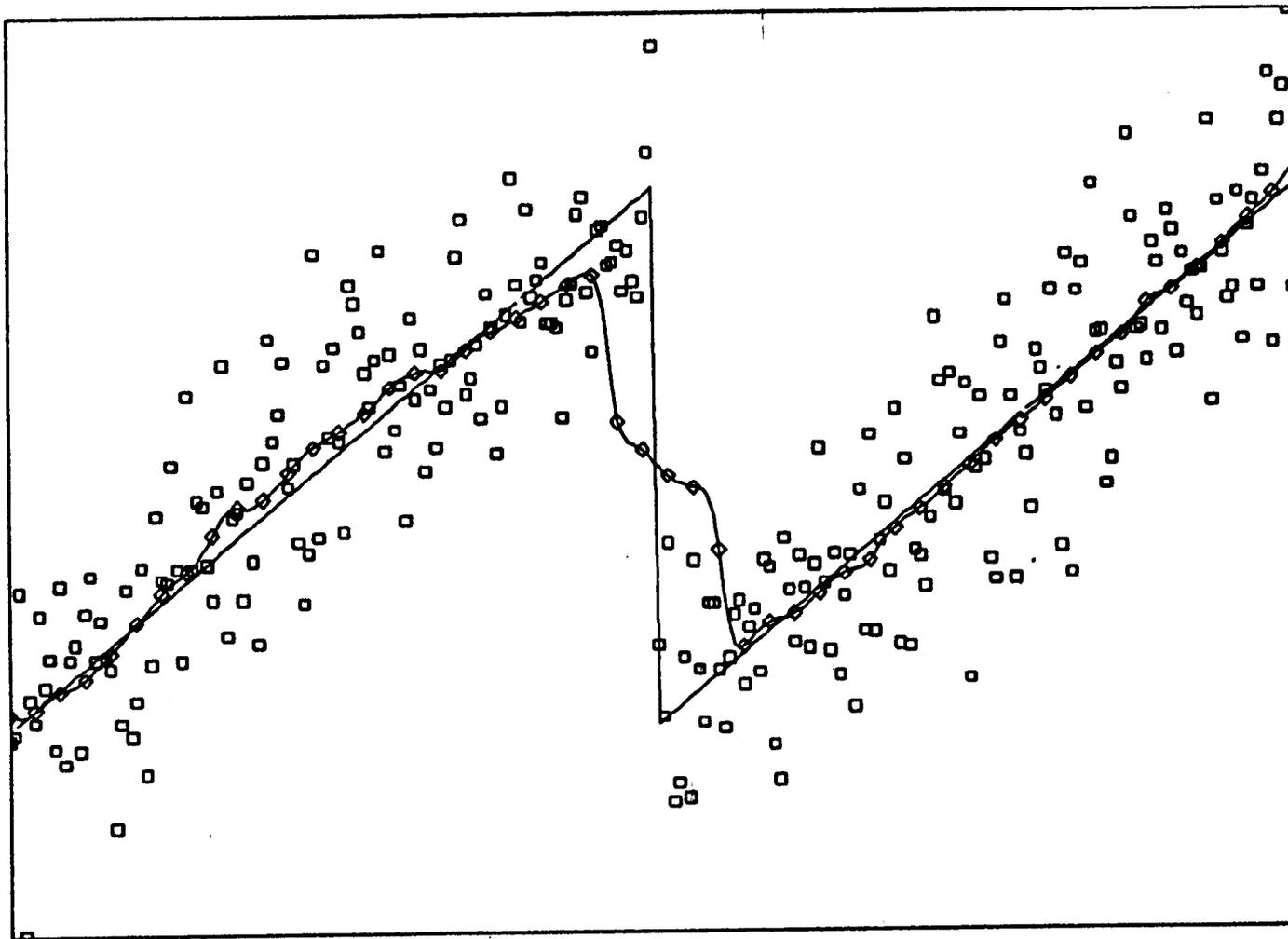
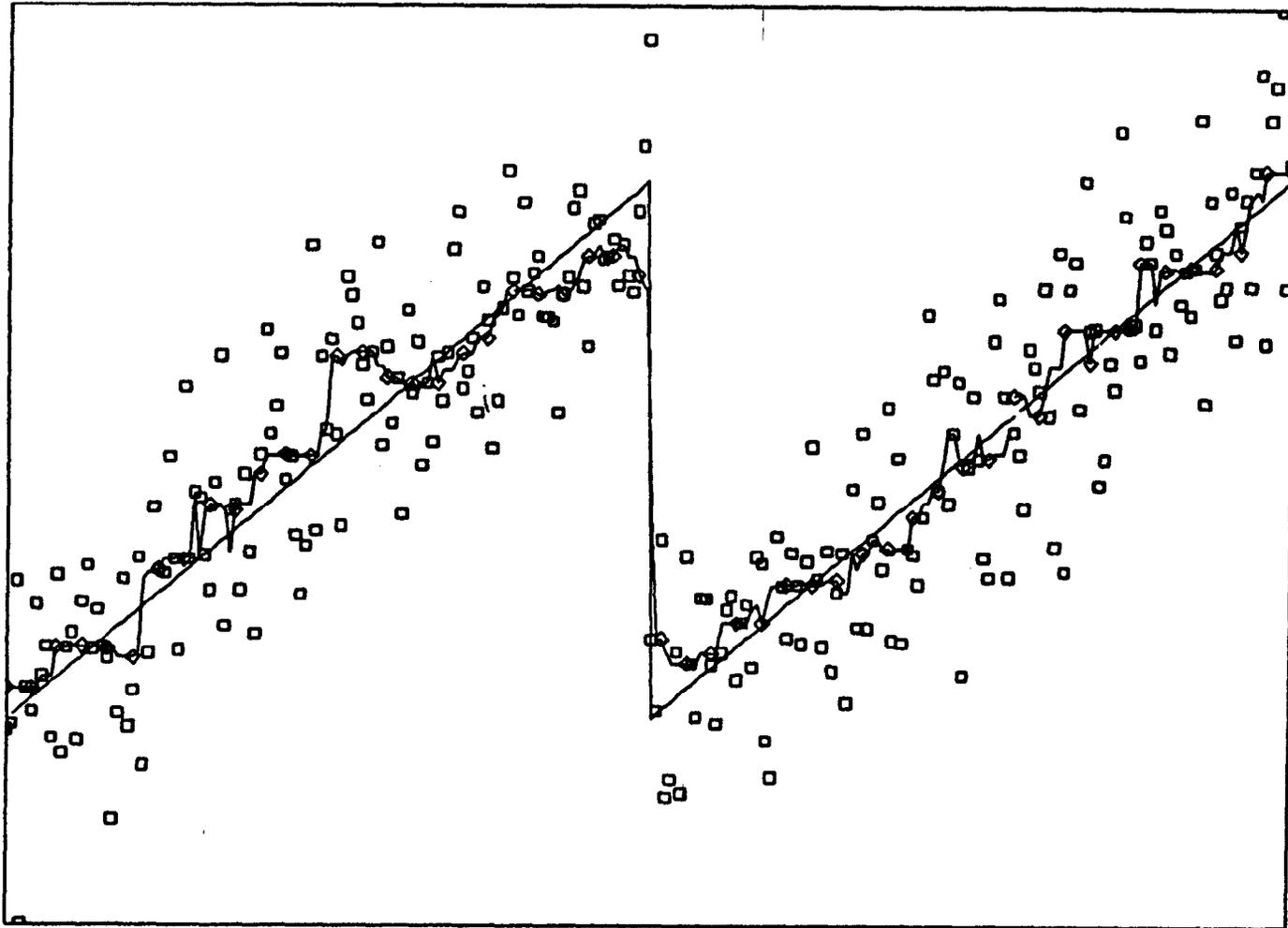Figure 2: A Median Smooth of a Sawtooth Function with Noise

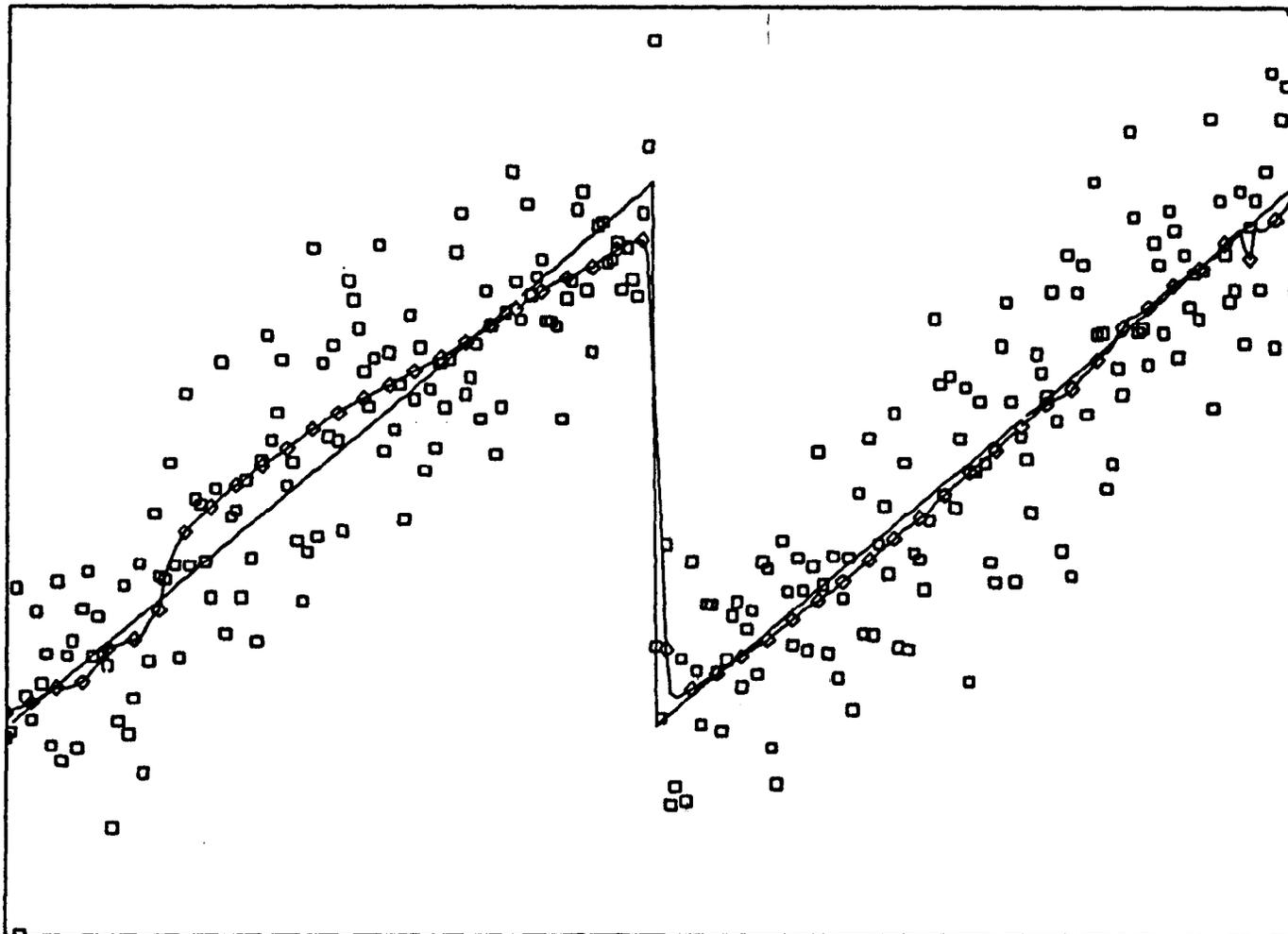Figure 3: A Split Linear Smooth of a Sawtooth Function with Noise

Figure 4: Ensemble of Central Smooths about the Sawtooth Function
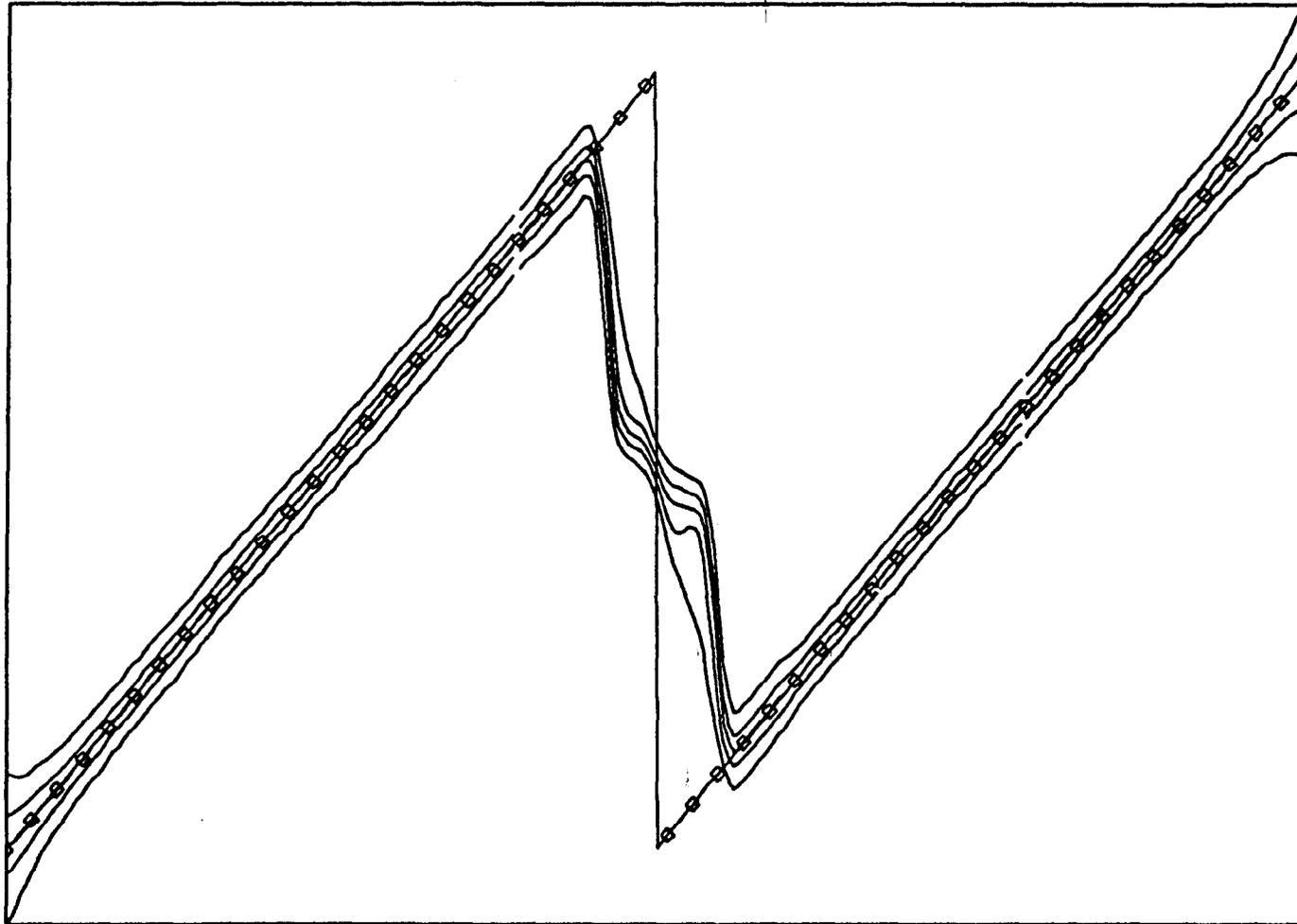
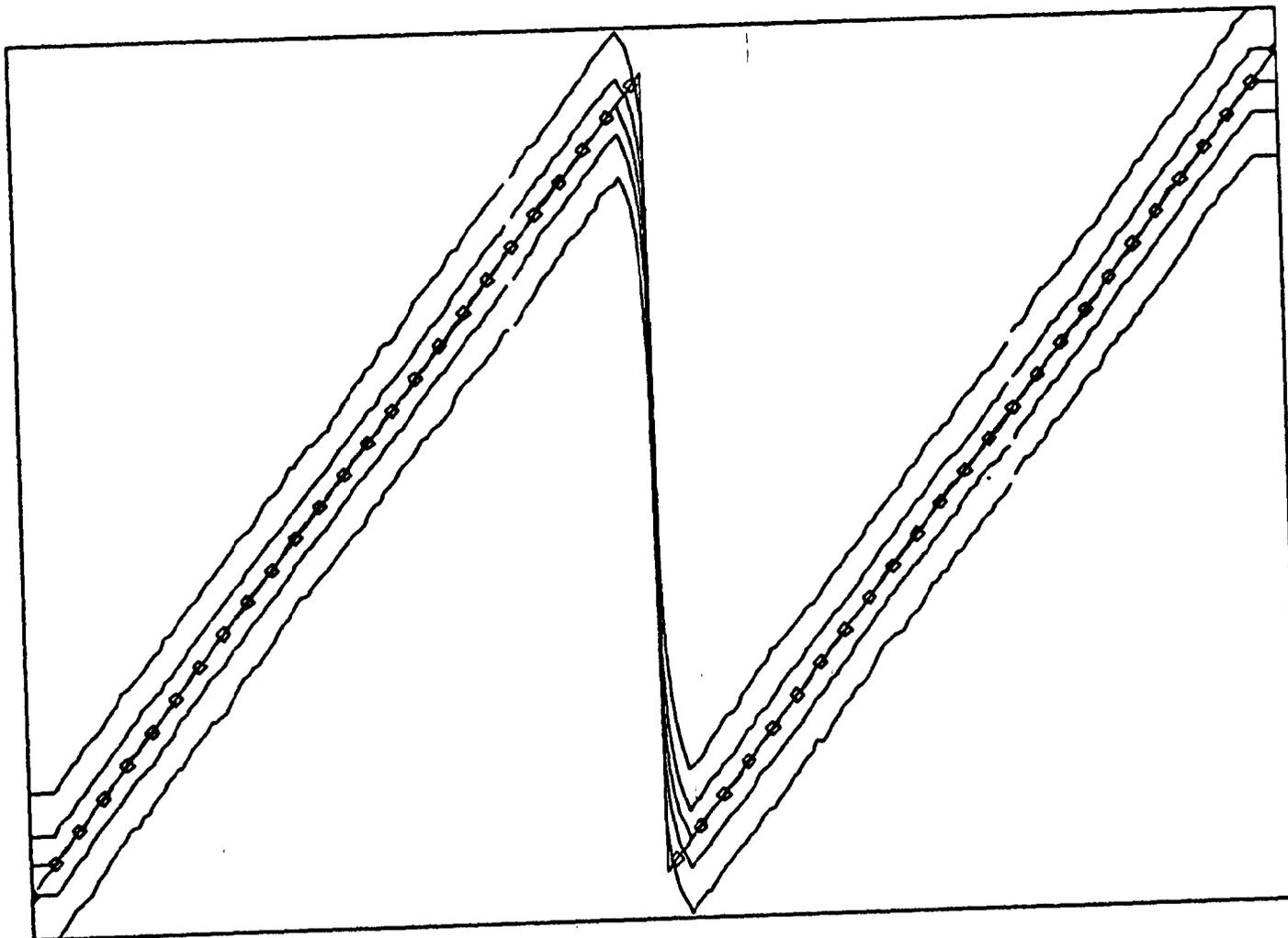Figure 5: Ensemble of Median Smooths about the Sawtooth Function

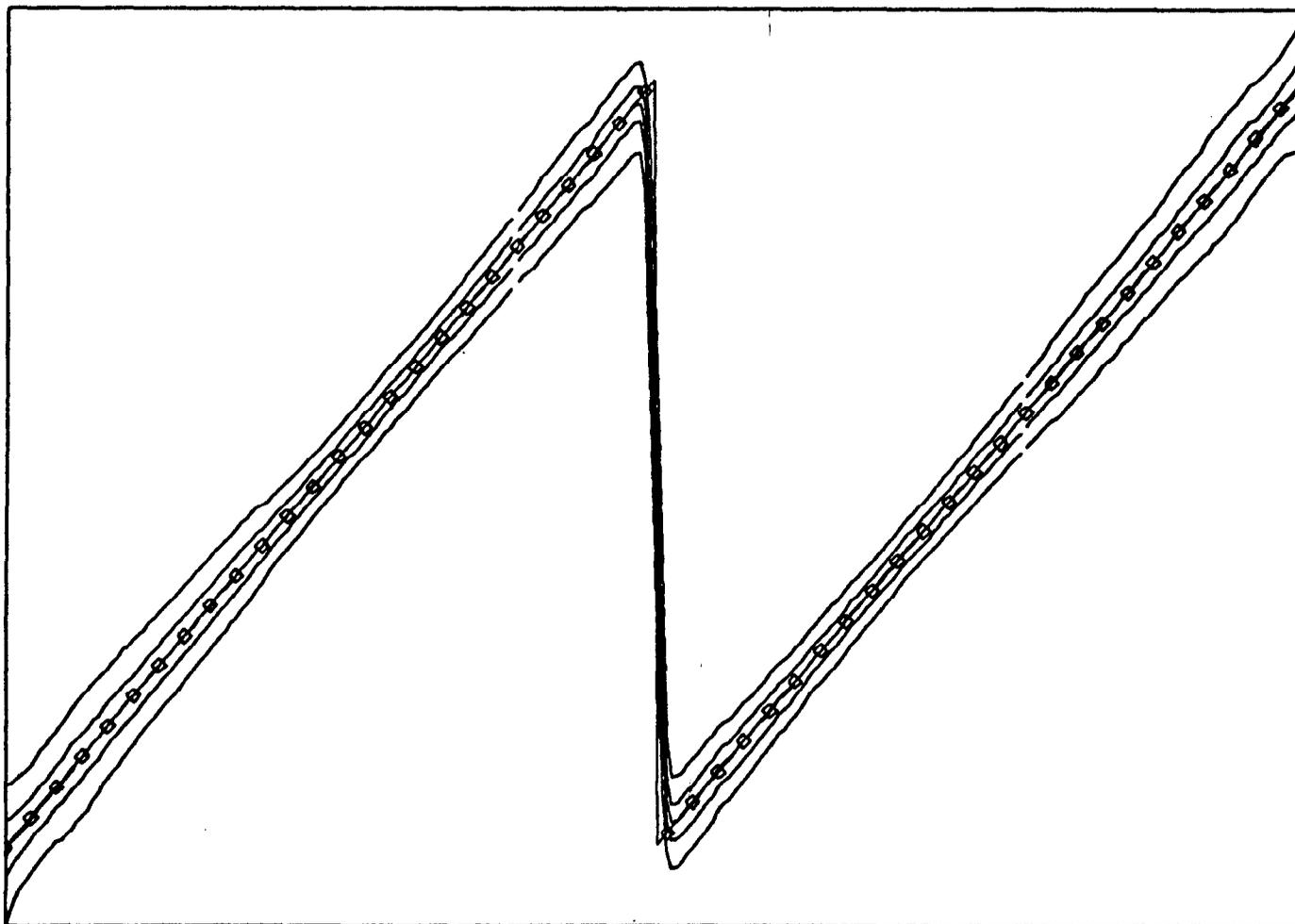Figure 6: Ensemble of Split Linear Smooths about the Sawtooth Function

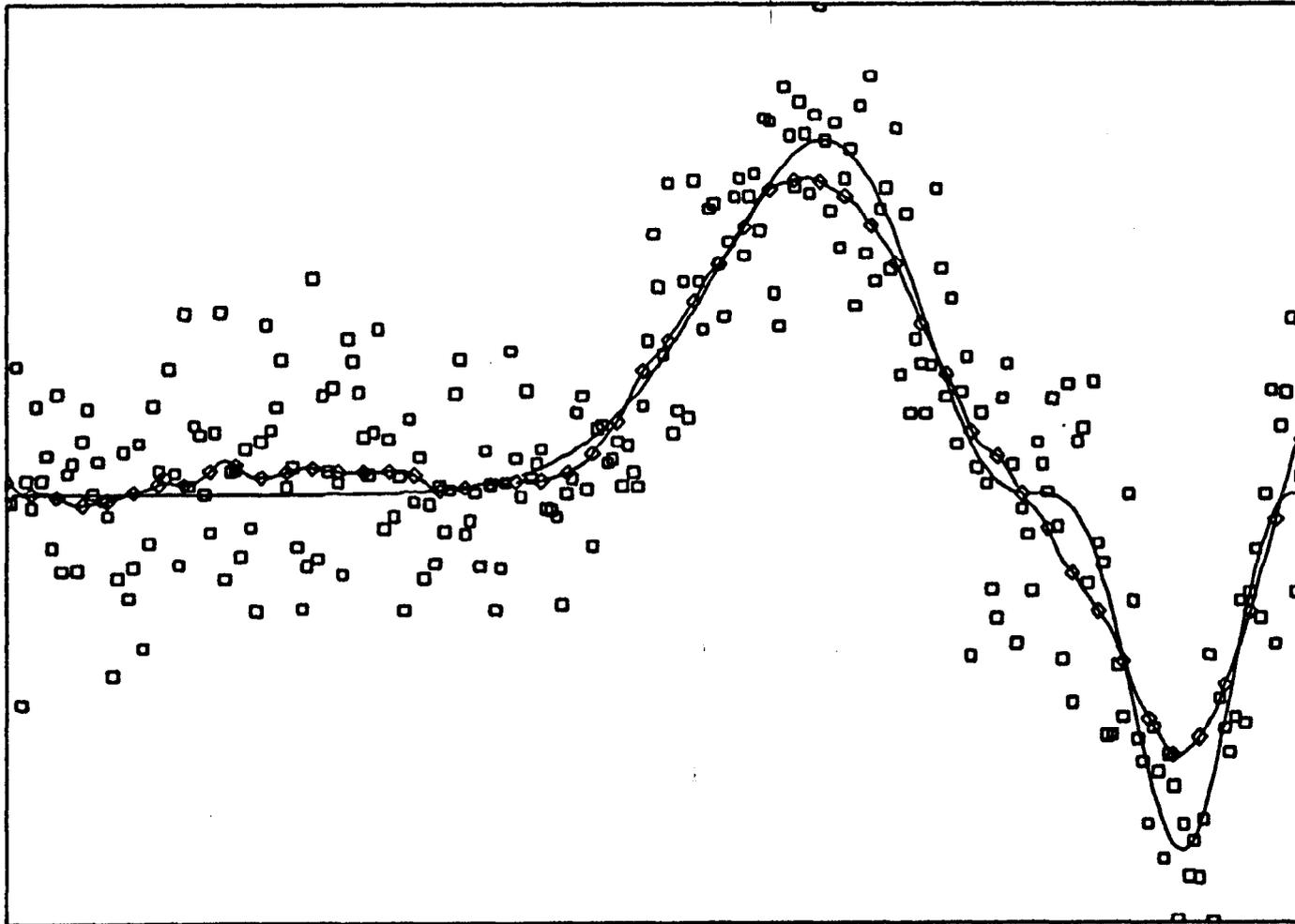Figure 7: A Central Smooth of $\sin^3(2\pi x^3)$ with Noise

Figure 8: A Median Smooth of $\sin^3(2\pi x^3)$ with Noise
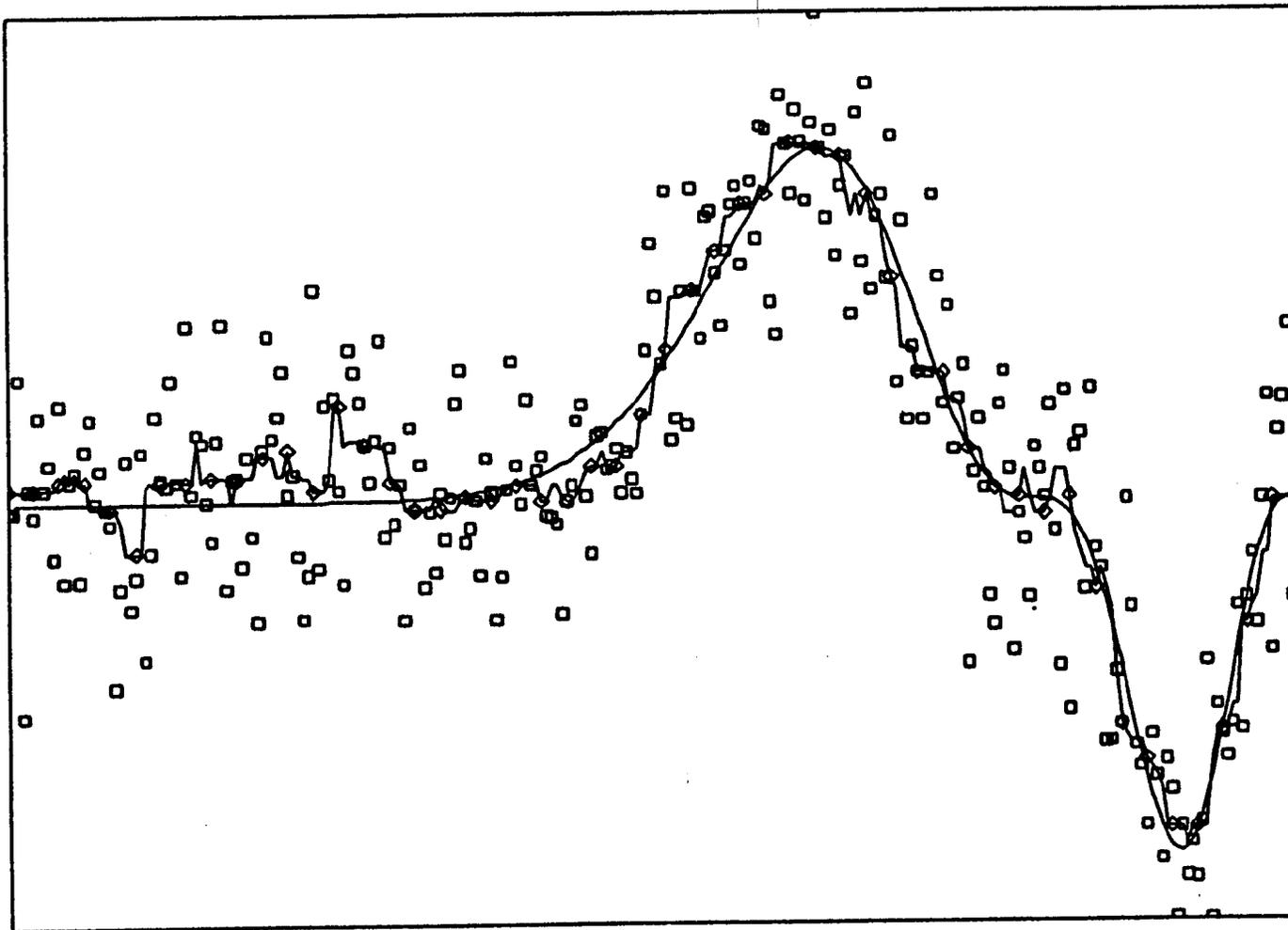
Figure 9: A Split Linear Smooth of $\sin^3(2\pi x^3)$ with Noise
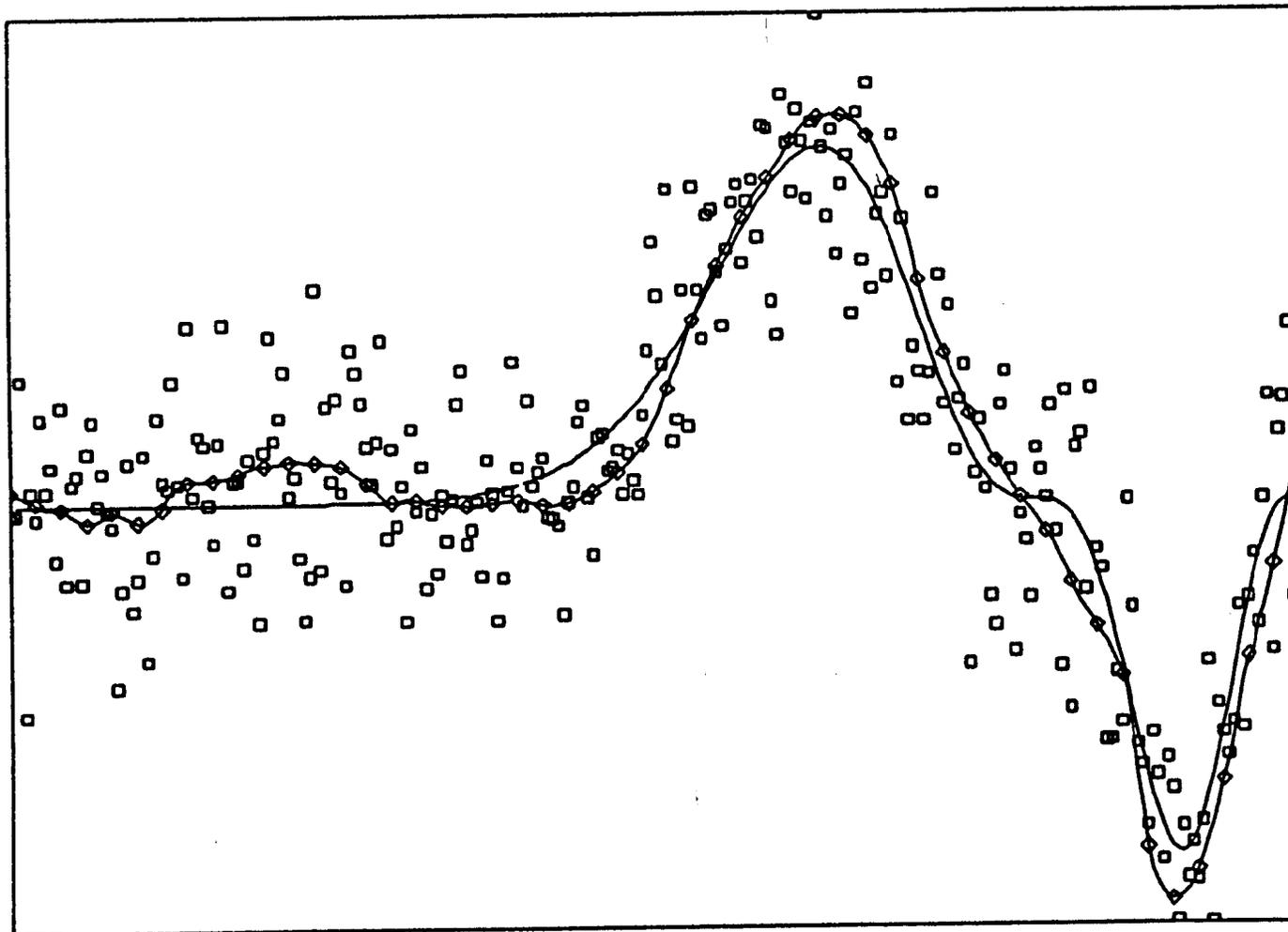
Figure 10: Ensemble of Central Smooths about $\sin^3(2\pi x^3)$
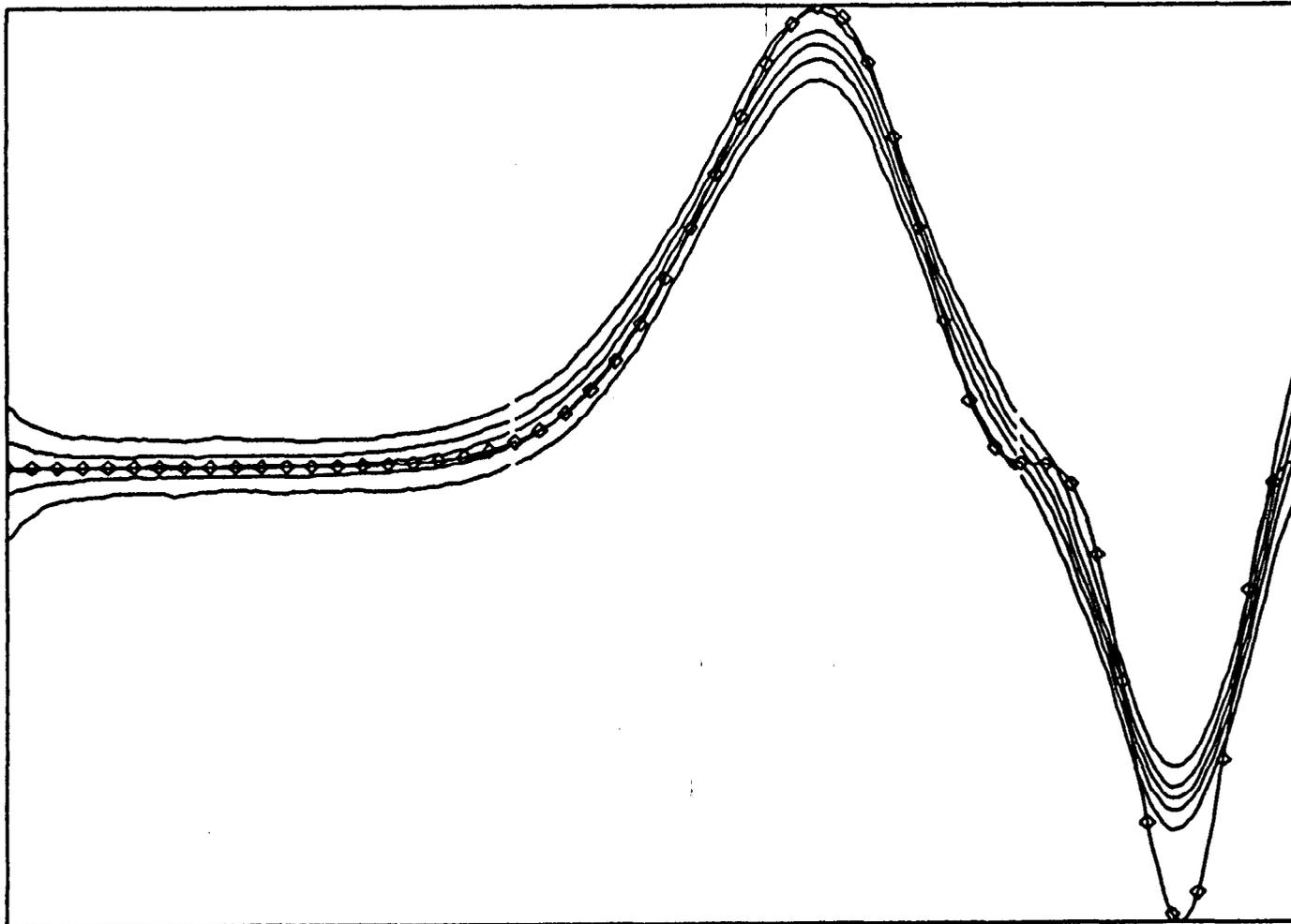
Figure 11: Ensemble of Median Smooths of $\sin^3(2\pi x^3)$
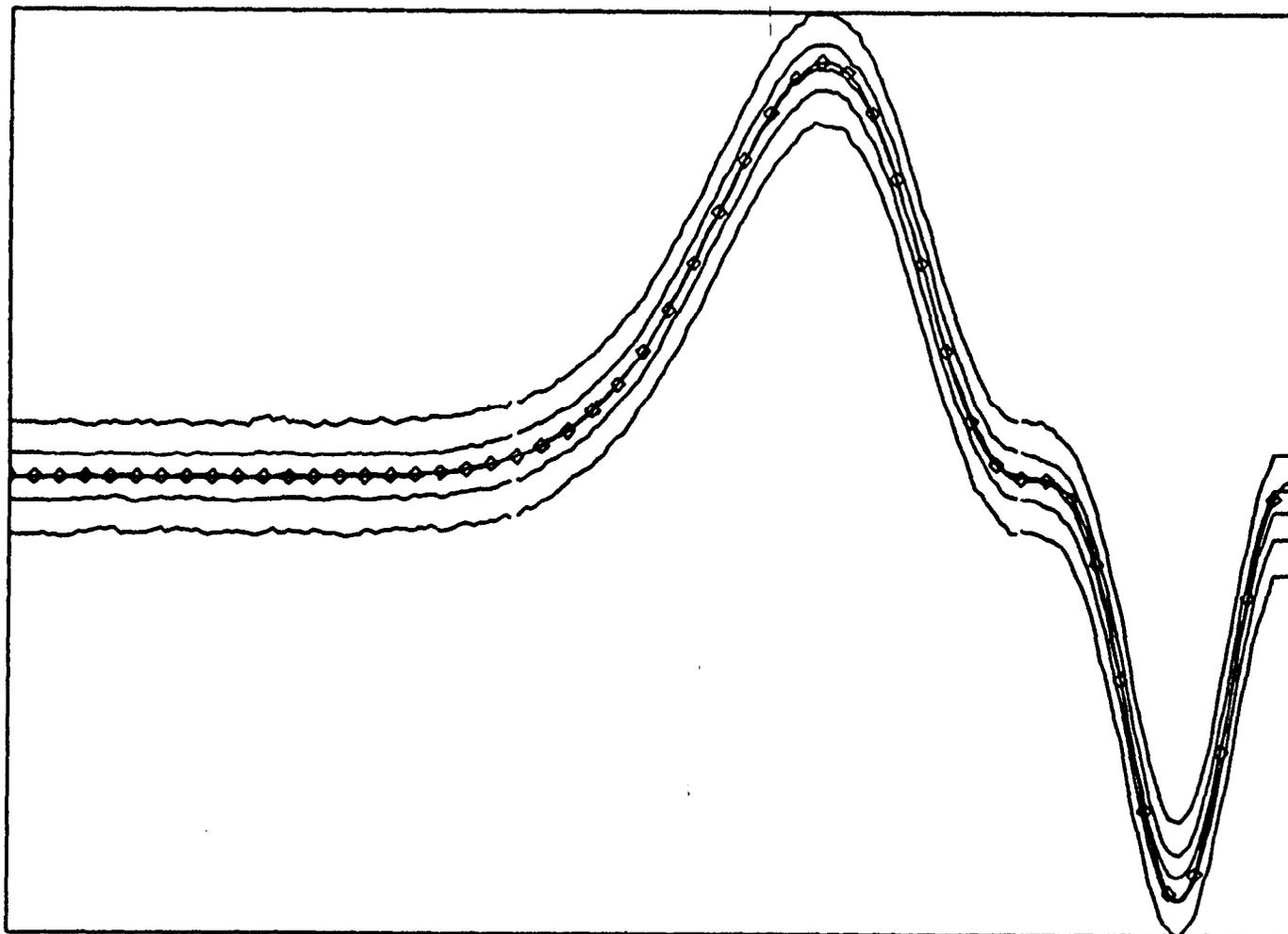
Figure 12: Ensemble of Split Linear Smooths of $\sin^3(2\pi x^3)$
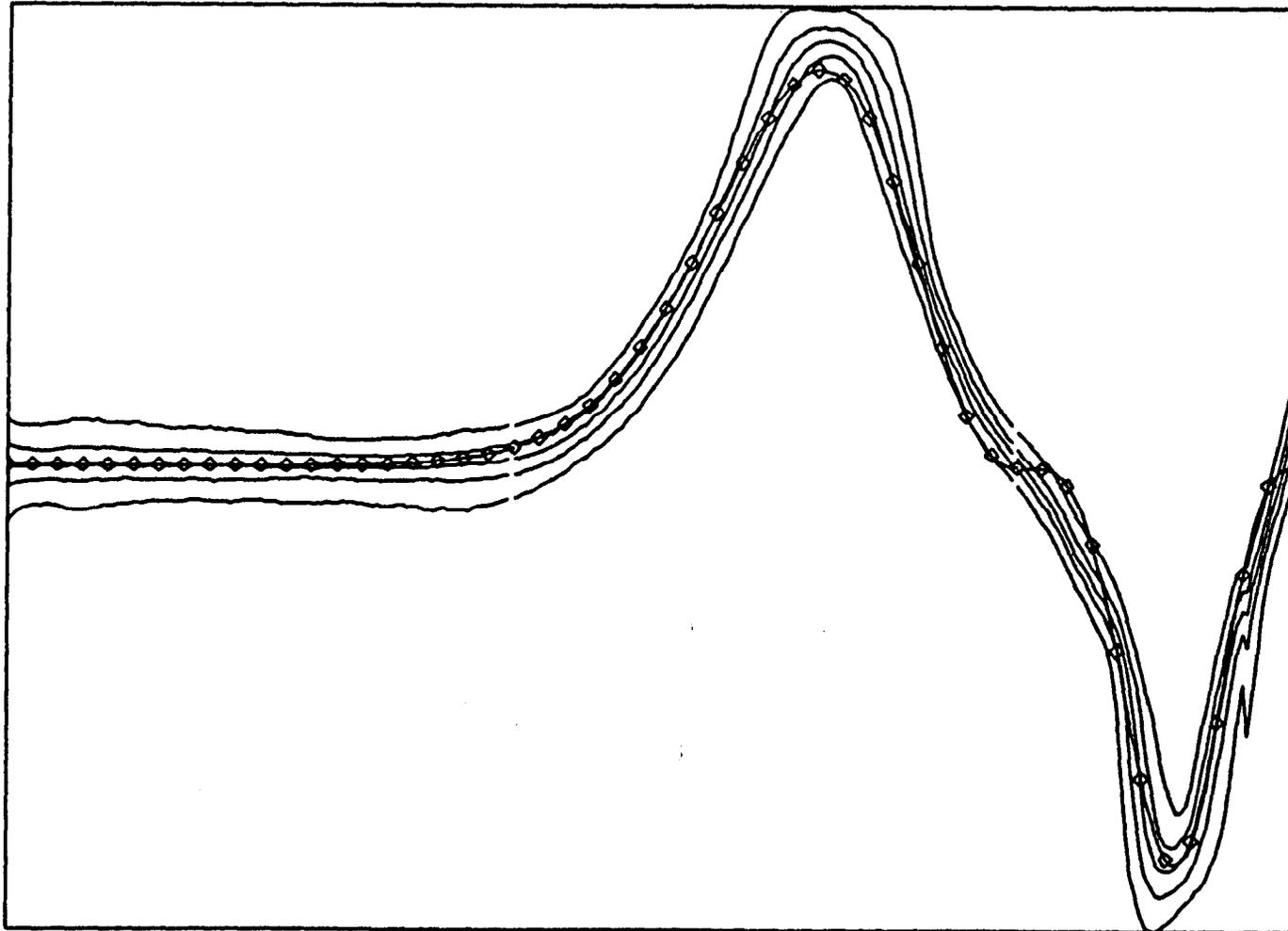
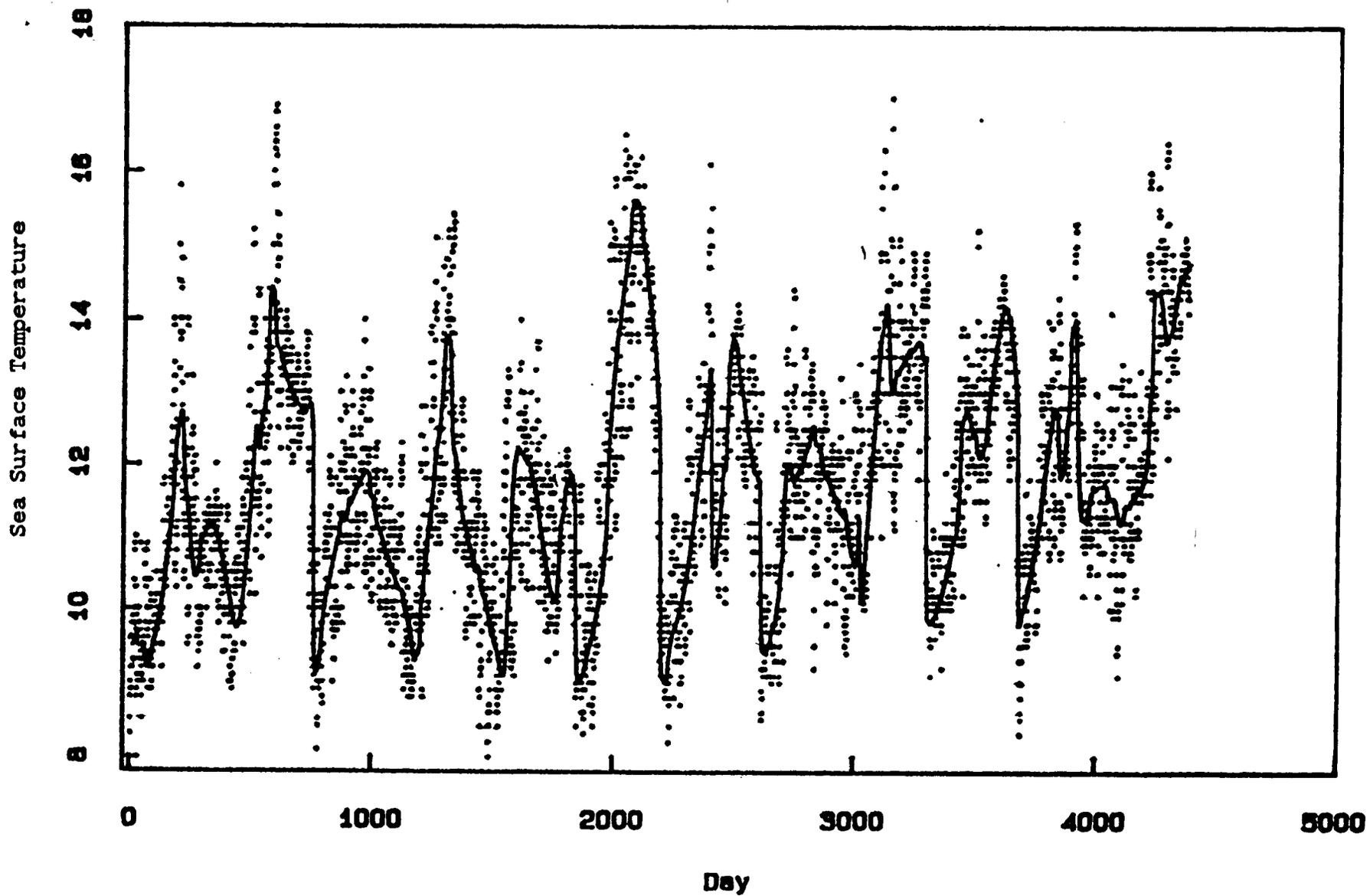Figure 13: 12 Years of Daily Sea Surface Temp. at Granite Cove California

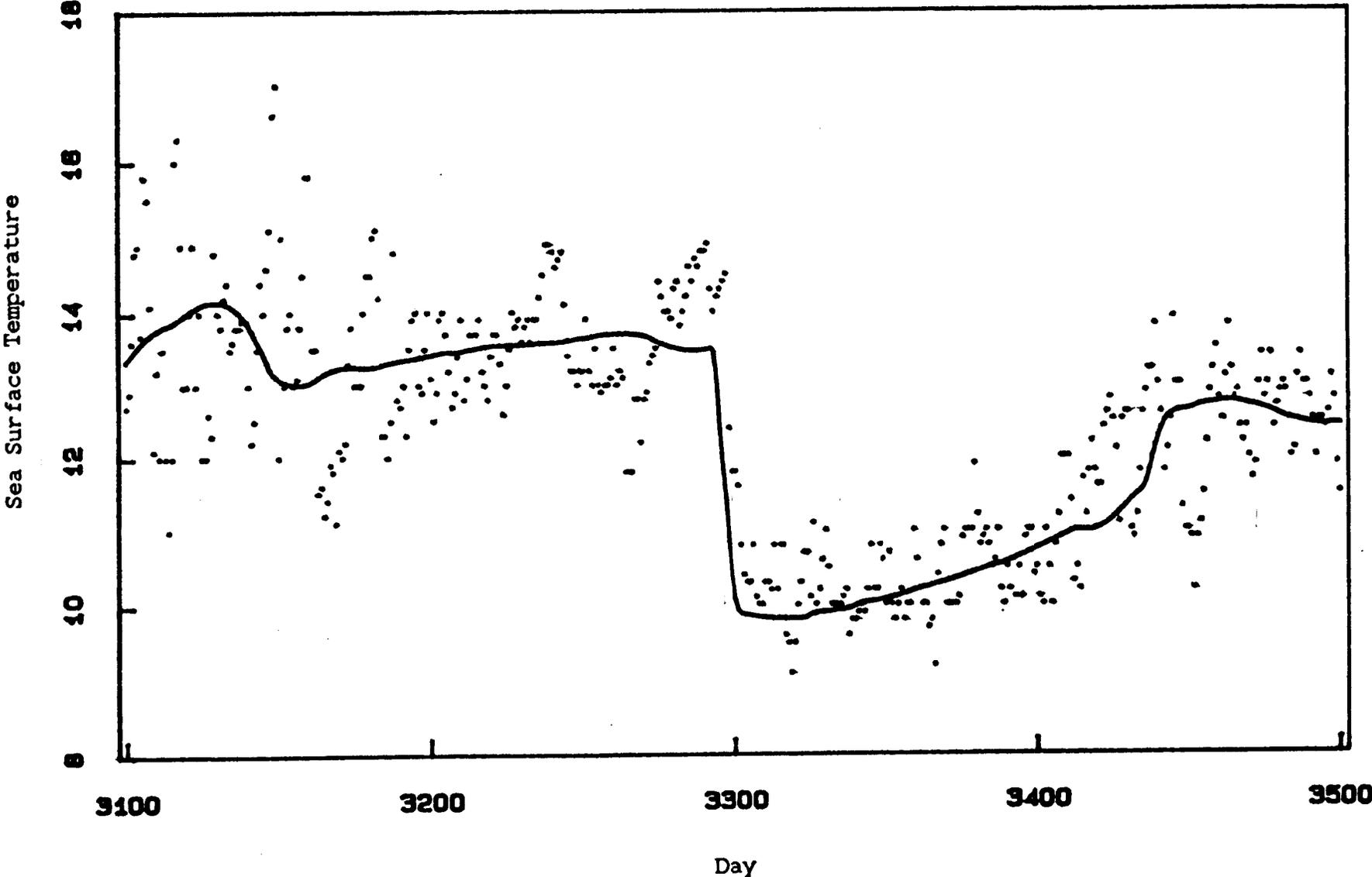Figure 14: Close-up of Spring Transition in Sea Surface Temperature Data
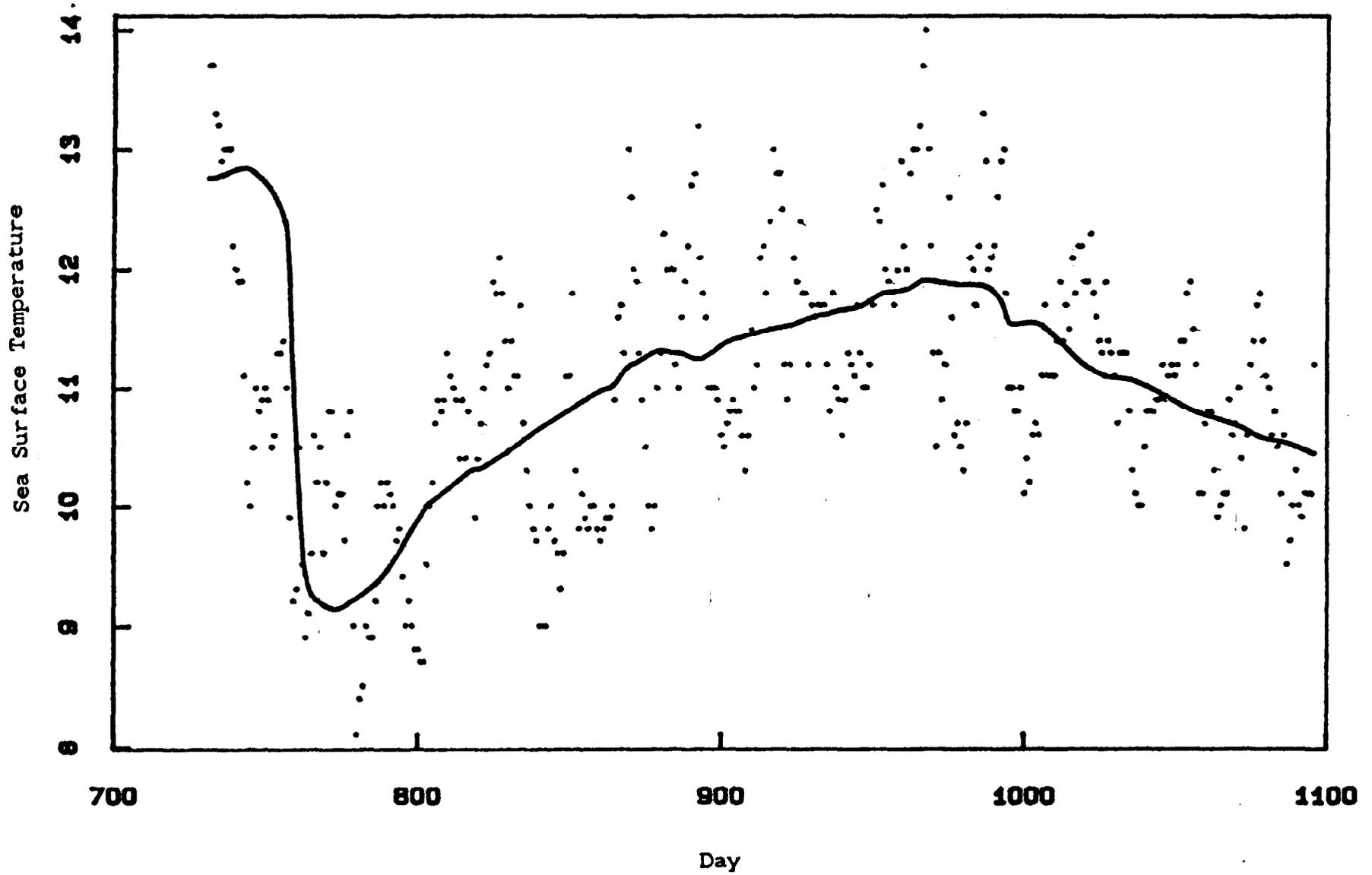
Figure 15: Close-up of Cold Year in Sea Surface Temperature Data

Figure 16: Cold Year with Small Window Split Linear Smooth