An Iterative Smoothing Method of Time Series: A Generalization of the Tukey's Method

Keisuke Suzuki

1. Introduction

John W. Tukey[1] proposed a very interesting smoothing method of time series. This method may be called "iterative moving median method" from the characteristics found in the method. When the method is applied to smooth a time series observed, the time series is smoothed by using iterative calculation. After sufficient times of the iterative calculation defined by this method have been done, the smoothed time series becomes a time series which is not changed even if the iterative calculation given by this method is additionally applied to the time series.

In this paper, I tried to expand and generalize the Tukey's method. And I could find a generalized smoothing method of time series by iterative calculation.

2. The Tukey's Method

When a time series which is observed at points of time and is written by

$$x_1,\ x_2,\ \cdots,\ x_t,\ \cdots,\ x_T$$

is given, the smoothing method of time series proposed by John W. Tukey consists of the following steps of calculation:

Step 1. Two new values x_1' and x_T' are added to the time series obtained by actual observation: $x_1,\ x_1,\ x_2,\ \cdots,\ x_T$. The new values are calculated by the following formula:

$$x_1'=3x_2-2x_3 \tag{2.1}$$

$$x_T'=3x_{T-1}-2x_{T-2} \tag{2.2}$$

Step 2. Median of 3 neighboring observed values, x_t, x_{t+1}, and x_{t+2} ($t=1,\ 2,\ \cdots,\ T-2$) is obtained.

Step 3. The median obtained at *Step* 1 is regarded as the smoothed value at time t+1, and it is expressed by $x_{t+1}(1)$ ($t=1,\ 2,\ \cdots,\ T-2$). At this stage of the process of calculation, the following new time series is obtained:

$$x_2(1),\ x_3(1),\ \cdots,\ x_{T-1}(1)$$

Step 4. $x_1(1)$ and $x_T(1)$ cannot be obtained by manipulation of *Step* 3. These values are given by the median of x_1', x_1, and x_2, and median of x_{T-1}, x_T, and x_T'. Therefore, the following formula can be written.

$$x_1(1) = \text{median}\ (x_1',\ x_1,\ x_2) \tag{2.3}$$

$$x_T(1) = \text{median}\ (x_{T-1},\ x_T,\ x_T') \tag{2.4}$$

Consequently, at this step, a new time series written by the following series is obtained :

$$x_1(1),\ x_2(1),\ \cdots,\ x_T(1)$$

The rule for obtaining the $x_1(1)$ and $x_T(1)$ which is constructed by the operations shown by equations (2.1) and (2.2) at *Step* 1 and equations (2.3) and (2.4) at *Step* 4 is called "end point rule".

Step 5. The $x_t(1)$ ($t=1, 2, \cdots, T$) is regarded as x_t.

Step 6. The operations which were done at the steps from *Step* 1 to *Step* 5 are again applied in turn to x_t obtained at *Step* 5, until the x_t is not changed even if the operations which were done at the steps from *Step* 1 to *Step* 5 are applied to the x_t obtained at *Step* 5.

Step 7. When the state in which x_t is not changed even if the manipulations which are done at the steps from *Step* 1 to *Step* 5 are applied is obtained, the x_t is regarded as a provisional smoothed time series.

Step 8. "Mesa" is found in provisional smoothed time series. Mesa is "pairs of adjacent points with a common value which is below or above the points on each side". in other words "a flat two point local maximum or minimum".[2)]

Step 9. The first value of two values which compose a mesa is regarded as the end point of a series given by a part of the provisional smoothed time series, which is composed by the first value of the mesa and the observed values x_t preceding the first value of the mesa. The second value of the mesa is regarded as the beginning point of a series given by the part of the provisional smoothed time series, which is composed by the second value of the mesa and the observed value x_t continued from the second value of the mesa. Therefore, if the first value of the mesa is x_τ, then the value x_τ is regarded as the end point of a time series, x_1, x_2, $\cdots$, $x_{\tau-1}$, x_τ, and if the second value of the mesa is $x_{\tau+1}$, then the value $x_{\tau+1}$ is regarded as the beginning of a time series, $x_{\tau+1}$, $x_{\tau+2}$, $\cdots$, x_T.

Step 10. The "end point rule" which is the rule for obtaining $x_1(1)$ and $x_T(1)$ by the operations given by *Step* 1 and *Step* 4 is applied to the x_τ and $x_{\tau+1}$, and the values newly obtained by the end point rule for x_τ and $x_{\tau+1}$ is written as $x_\tau(*)$ and $x_{\tau+1}(*)$.

Step 11. The component of the mesa composed by x_τ and $x_{\tau+1}$ in the provisional smoothed time series is replaced by $x_\tau(*)$ and $x_{\tau+1}(*)$. The time series obtained at this step is called here new provisional smoothed time series.

Step 12. The $x_\tau(*)$ and $x_{\tau+1}(*)$ which are found in components of the new provisional smoothed time series are regarded as x_τ and $x_{\tau+1}$. Therefore, at this step, the new provisional smoothed time series which is composed by the values :

$$x_1,\ x_2,\ \cdots,\ x_\tau(*),\ x_{\tau+1}(*),\ \cdots,\ x_T$$

is again expressed by the following time series.

$$x_1,\ x_2,\ \cdots,\ x_\tau,\ x_{\tau+1},\ \cdots,\ x_T$$

Step 13. New provisional smoothed time series which is composed by x_t ($t=1, 2, \cdots, T$) is again smoothed by the operations given by the steps from *Step* 1 to *Step* 12.

Step 14. When the state in which all the components of the new provisional smoothed time series x_t ($t=1, 2, \cdots, T$) are not changed even if the operations defined at the steps from *Step* 1 to *Step* 13 are applied is obtained, the new provisional smoothed time series is regarded as a final smoothed time series, which is expressed by the following time series.

$$x_1^*,\ x_2^*,\ \cdots,\ x_t^*,\ \cdots,\ x_T^*$$

The Tukey's method is the method as shown above. In this method, if the end point rule is not given, the time series which is composed by $x_{t+1}(1)$ ($t=1, 2, \cdots, T-1$) has not $x_1(1)$ and $x_T(1)$. But, fortunately, by the end point rule, the time series which is composed by $x_1(1), x_2(1), \cdots, x_T(1)$ is obtained.[3]

The process of the manipulations which are done in the Tukey's method can be depicted by Figure 1.

From original time series, augmented time series, $x'_1, x_1, x_2, \cdots, x_t, \cdots, x_T, x_T'$, is obtained by the operation of *Step* 1. When the operations of *Steps* 2 and 3 are given to the augmented time series, incomplete smoothed time series is obtained. And by the operation of *Step* 4, complete smoothed time series is obtained. At *Step* 5, the complete smoothed time series is regarded as the original time series observed. By the operations of *Steps* 6 and 7, provisional smoothed time series is calculated. At *Step* 8, mesas in the provisional smoothed time series are found. If mesas are found, they are splitted. And the new provisional smoothed time series is obtained by the operations of *Steps* 12 and 13. Lastly, after the operation of *Step* 14 is given to the new provisional smoothed time series, final smoothed time series is obtained. Of course, if mesas cannot be found, provisional smoothed time series is regarded as final smoothed time series.

By the Tukey's smoothing method, a smoothed time series which was called here final smoothed series is obtained by medians of 3 neighboring observed values. But, we would be able to obtain a smoothed time series which consists of medians of more than 3 neighboring observed values by generalizing the Tukey's method. Generalized smoothing rules of time series proposed in this paper are the rules which give smoothed time series which are composed by medians of $2n+1$ neighboring observed values or $2n$ neighboring observed values ($n=1, 2, \cdots$).

3. Decomposition of End Point Rule

In the Tukey's smoothing method, as mentioned above, the length of the final smoothed time series is not shortened by the end point rule. If this rule is generalized, we would be able to find a generalized smoothing method based on the Tukey's method.

The Tukey's end point rule is regarded as a combination of "extraporation rule" and "smoothing rule".

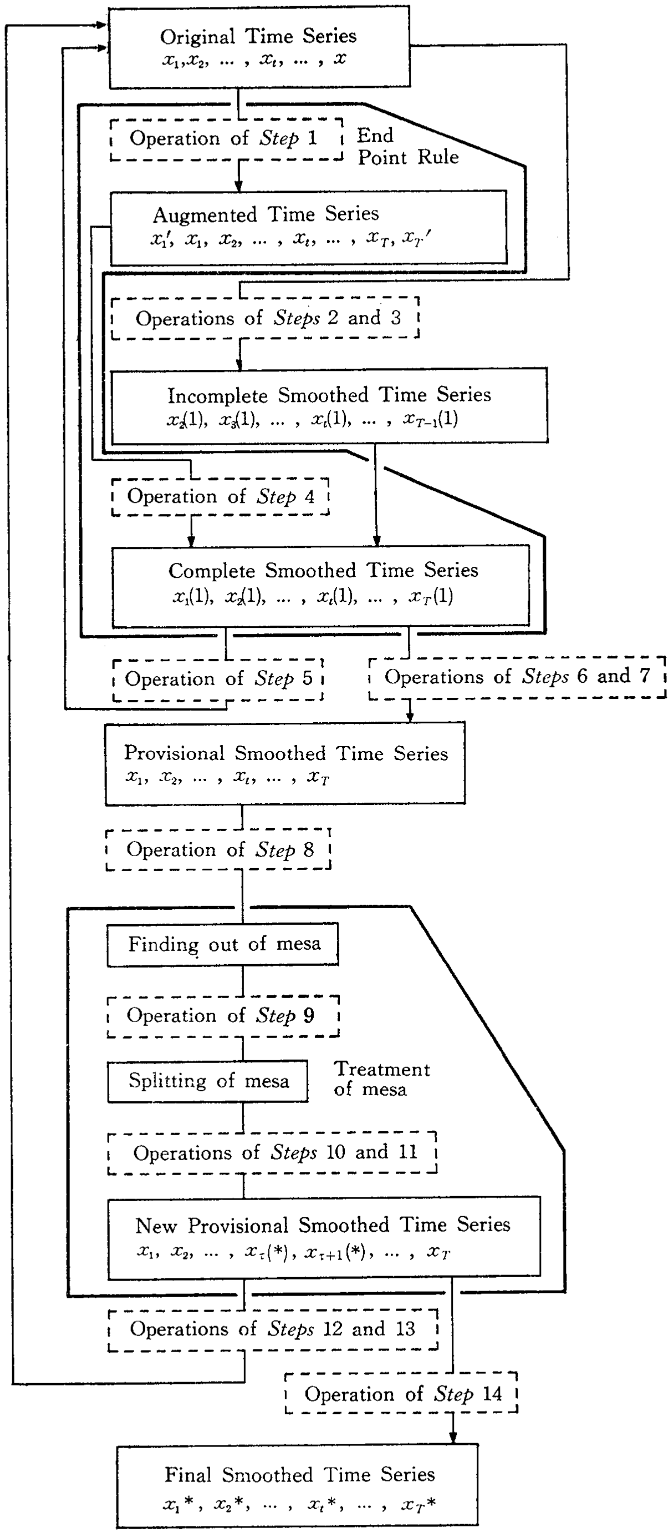

Figure 1 The process of the manipulations of the Tukey's method.

According to McNeil's interpretation, when the end point rule is applied to the time series analized, the value $3x_1-2x_2$ which is the value x_1 defined above is "actually obtained by fitting a straight line passing through the second point with slope double that of the line joining the second and third points and taking the value fitted by this relation" at the point of time 1.[4] For example, if the values of x_1, x_2, and x_3 are the values which are shown in Figure 2, the value x_1' is found at point P in this figure. But, this value x_1' $(=3x_2-2x_3)$ can be regarded as the value of x_0 obtained by the extrapolation of the time series observed, namely, the value can be actually obtained by fitting a straight line passing through the second point with slope of the line joining the second and third points and taking the value fitted by this relation at the point of "time 0". For example, if the value of x_1, x_2, and x_3 are the values which are shown in Figure 2, the value x_1' is found at point Q in Figure 3. Therefore, the value x_1' is regarded as x_0 obtained by the extrapola-

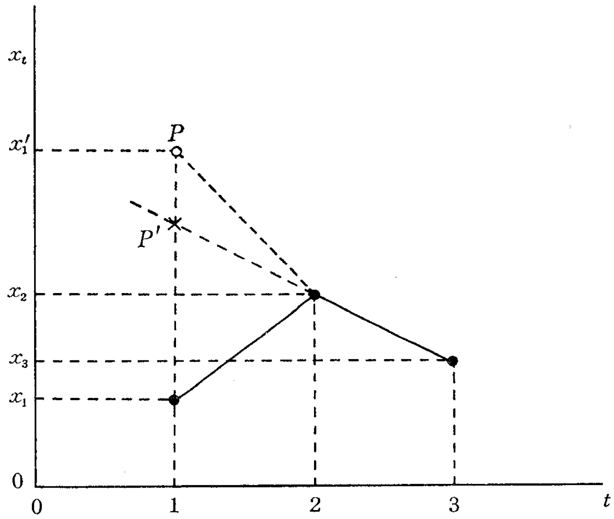

Figure 2 Determination of the value x_1'.

Notes : This figure is drawn, based on the figure shown in McNeil's book (McNeil, Donald R. : Interactive Data Analysis, New York, John Wiley, 1977, p. 122).

The point P' is the value for x_1' obtained by using the slope of the line combining the points expressing the values of x_3 and x_2.

tion of the time series observed, and the rule for obtaining the value x_0 can be called "extrapolation rule" of time series. The smoothed value for time 1 is given by the median of x_0, x_1, and x_2. This rule of obtaining the median is the same rule that the rule of smoothing of other part of the time series observed, which may be called "smoothing rule" in the Tukey's method.

The end point rule for obtaining the smoothed value of x_T is also regarded as a combination of "extrapolation rule" for obtaining the value of x_{T+1} and "smoothing rule" by which moving median for x_T is obtained.

Therefore, the Tukey's end point rule is regarded as a combination of "extrapolation

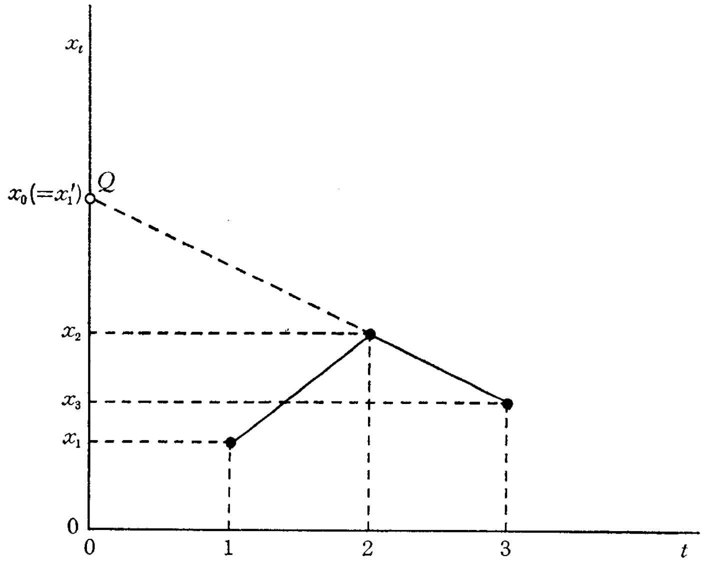

Figure 3 Determination of the value x_0.

Note : The value x_0 is exactly equal to the value x_1'.

rule" for obtaining the values x_0 and x_{T+1} and "smoothing rule" by which moving medians for x_1 and x_T are obtained.

4. Generalization of Extrapolation Rule (Genelalized Extrapolation Rule)

4.1 Generalization of the Tukey's Rule: Generalized Tukey's Extrapolation Rule

According to the Tukey's rule, x_0 is obtained by

$$x_0=3x_2-2x_3 \tag{4.1}$$

or

$$x_0=x_2-2\ (x_3-x_2) \tag{4.2}$$

The value (x_3-x_2) in equation (4.2) is the slope of the line combining the points expressihg the values x_3 and x_2. Therefore, we can generalize this rule for obtaining the value x_{-1}. The value x_{-1} is expressed by the following equation.

$$x_{-1}=x_2-3\left\{\frac{1}{2}(x_4-x_2)\right\} \tag{4.3}$$

The value $\frac{1}{2}(x_4-x_2)$ is the slope of the line combining the points expressing the values x_4 and x_2 are shown in Figure 4. Consequently,

$$x_{-1}=\frac{5}{2}x_2-\frac{3}{2}x_4 \tag{4.4}$$

Similarly, the value of x_{-2} is expressed by equation :

$$x_{-2}=x_2-4\left\{\frac{1}{3}(x_5-x_2)\right\} \tag{4.5}$$

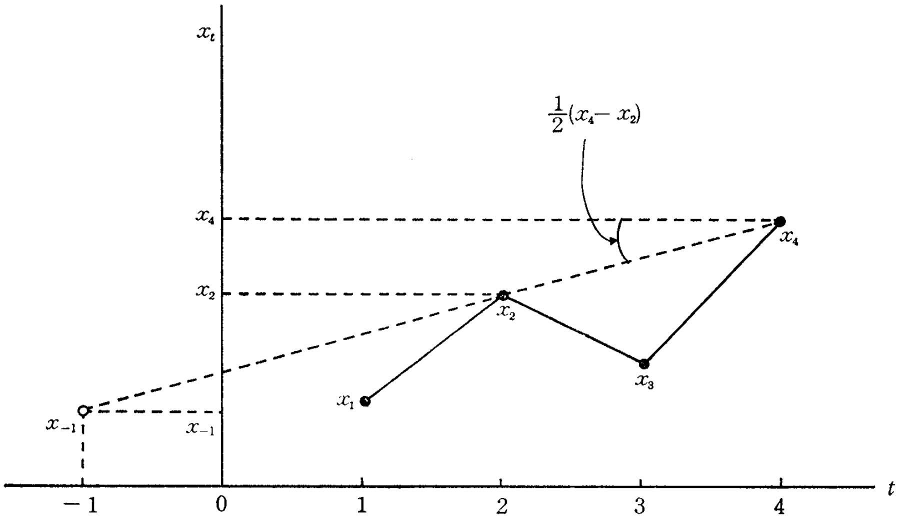

Figure 4 Determination of the value x_{-1}.

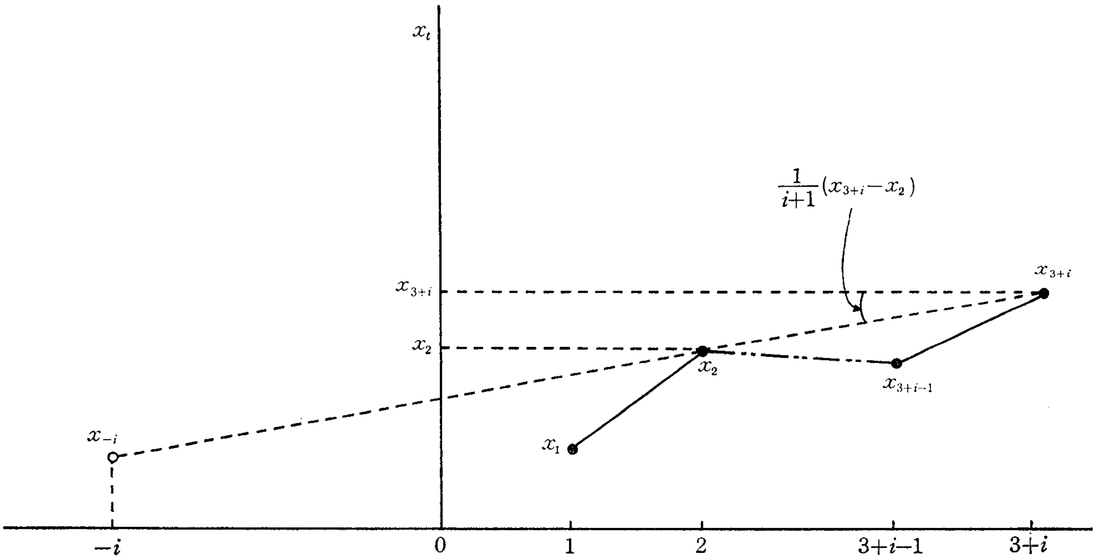

Figure 5 Determination of the value x_{-i}.

In this equation, the value $\frac{1}{3}(x_5-x_2)$ is also the slope of the line combining the points expressing x_5 and x_2. Therefore, x_{-2} is expressed by

$$x_{-2}=\frac{7}{3}x_2-\frac{4}{3}x_5 \tag{4.6}$$

In general, the value x_{-i} is expressed by equation :

$$x_{-i}=x_2-(i+2)\left\{\frac{1}{i+1}(x_{3+i}-x_2)\right\} \tag{4.7}$$

Figure 5 shows the mechanism of determination of the value x_{-i} expressed by equation (4.7). The value $\frac{1}{i+1}(x_{3+i}-x_2)$ in equation (4.7) is the slope of the line combining the points expressing the values x_{3+i} and x_2. From the equation (4.7), we can obtain the following function for the value x_{-i}.

$$x_{-i}=\frac{2i+3}{i+1}x_2-\frac{i+2}{i+1}x_{3+i} \tag{4.8}$$

Therefore, we can also obtain the following formula for the value x_{T+i} :

$$x_{T+i}=\frac{2(i+1)}{i}x_{T-1}-\frac{i+2}{i}x_{T-i-1} \tag{4.9}$$

which is obtained by the following relation between x_{T+i}, x_{T-1}, and x_{T-i-1}.

$$x_{T+i}=x_{T-1}-(i+2)\left\{\frac{1}{i}(x_{T-i-1}-x_{T-1})\right\} \tag{4.10}$$

The rule for obtaining x_{T-i} and x_{T+i} which is expressed by equations (4.7) and (4.10) is the "generalized Tukey's extrapolation rule".

4.2 Median Extrapolation Rule

Another extrapolation rule can be given, if we make a extrapolation rule for the values x_{-i} and x_{T+i}, by using an assumption that the value x_{-i} will appear in the vicinity of the values which are found at time $i+2$ and at points of time near time $i+2$, consequently, it will appear in the vicinity of the median of the actual values observed, $x_1, x_2, \cdots, x_{i+2}, \cdots, x_{2i+3}$; and the value x_{T+1} will appear in the neighborhood of the median of the actual values $x_{T-2i}, x_{T-2i-1}, \cdots, x_{T-1}, x_T$.

Therefore, the rule stated above can be expressed by the following equations :

$$x_{-i}=\text{median}\ (x_1, x_2, \cdots, x_{2i+3}) \tag{4.11}$$

and

$$x_{T+i}=\text{median}\ (x_{T-2i}, x_{T-2i-1}, \cdots, x_{T-1}, x_T) \tag{4.12}$$

where "median $(x_p, x_{p+1}, \cdots, x_m)$" means the median of the values $x_p, x_{p+1}, \cdots, x_m$. This rule expressed by equations (4.11) and (4.12) is called here "median extrapolation rule".

4.3 Median-Trend Extrapolation Rule

If a time series has a secular trend, it may be supposed that the x_{-i} and x_{T+i} will appear in the vicinity of the values on a line which expresses the secular trend of the series observed. The line which expresses the secular trend of the series observed may be expressed by a line which is obtained by combining a point expressing the value of x_1 and a median obtained from an adequate number of actual values which appear in the vicinity of time $-i$, or by combining a point expressing the value of x_T and a median obtained from an adequate number of actual values which appear in the vicinity of time $T+i$.

By using the presupposition stated above, we can obtain a rule for determining the values x_{-i} and x_{T+i}. The rule is written by the following equations :

$$x_{-i}=x_1-(i+1)\frac{1}{i+1}\{\text{median }(x_1,\ x_2,\ \cdots,\ x_{2i+3})-x_1\} \tag{4.12}$$

and

$$x_{T+i}=x_T+i\frac{1}{i}\{x_T-\text{median }(x_{T-2i},\ x_{T-2i-1},\ \cdots,\ x_{T-1},\ x_T)\} \tag{4.13}$$

Here, the values, median $(x_1, x_2, \cdots, x_{2i+3})$ and median $(x_{T-2i}, x_{T-2i-1}, \cdots, x_{T-1}, x_T)$ were regarded as the values on the line expressing secular trend at time $i+2$ and at time $T-i$, respectively ; and $\frac{1}{i+1}$ {median $(x_1, x_2, \cdots, x_{2i+3})$ $-x_2\}$ and $\frac{1}{i}$ $\{x_T-$median $(x_{T-2i}, x_{T-2i-1}, \cdots, x_{T-1}, x_T)\}$ were regarded as the slopes of the lines expressing secular trend found in the actual time series, as shown in Figure 6.

The value x_{-i} and x_{T+i} in equations (4.12) and (4.13) are regarded as the values on the lines of secular trend of the time series observed which are obtained by combining the points expressing the values of x_1 and median $(x_1, x_2, \cdots, x_{2i+3})$ and by combining the points expressing the values of median $(x_{T-2i}, x_{T-2i-1}, \cdots, x_{T-1}, x_T)$ and x_T.

The extrapolation rule defined by equations (4.12) and (4.13) is called here "median-trend extrapolation rule".

5. Generalization of Smoothing Rule: Generalized Smoothing Rule

According to the Tukey's smoothing rule, in general, except the end point, the smoothed value at time t, $x_t(1)$ is given by

$$x_t(1)=\text{median }(x_{t-1},\ x_t,\ x_{t+1}) \tag{5.1}$$

If we generalize this rule, we will be able to define a "generalized smoothing rule" expressed by the following expression :

$$x_t(1)=\text{median }(x_{t-q},\ x_{t-q-1},\ \cdots,\ x_t,\ \cdots,\ x_{t-q})\quad (q=1,\ 2,\ \cdots\cdots,\ n) \tag{5.2}$$

As shown equation (5.2), this rule can be applied to get smoothed value obtained by the median of $2n+1$ terms. The value defined by equation (5.2) is called here "$2n+1$ term moving median".

6. Generalized End Point Rule

In general, when we want to calculate smoothed series composed by $2n+1$ term moving median from a time series which is written by

$$x_1,\ x_2,\ x_3,\ \cdots,\ x_T$$

we must prepare new $2n$ values :

$$x_{-n+1},\ x_{-n+2},\ \cdots,\ x_{-1},\ x_0\ \text{ and }\ x_{T+1},\ x_{T+2},\ \cdots,\ x_{T+n-1},\ x_{T+n}$$

for calculating the following values :

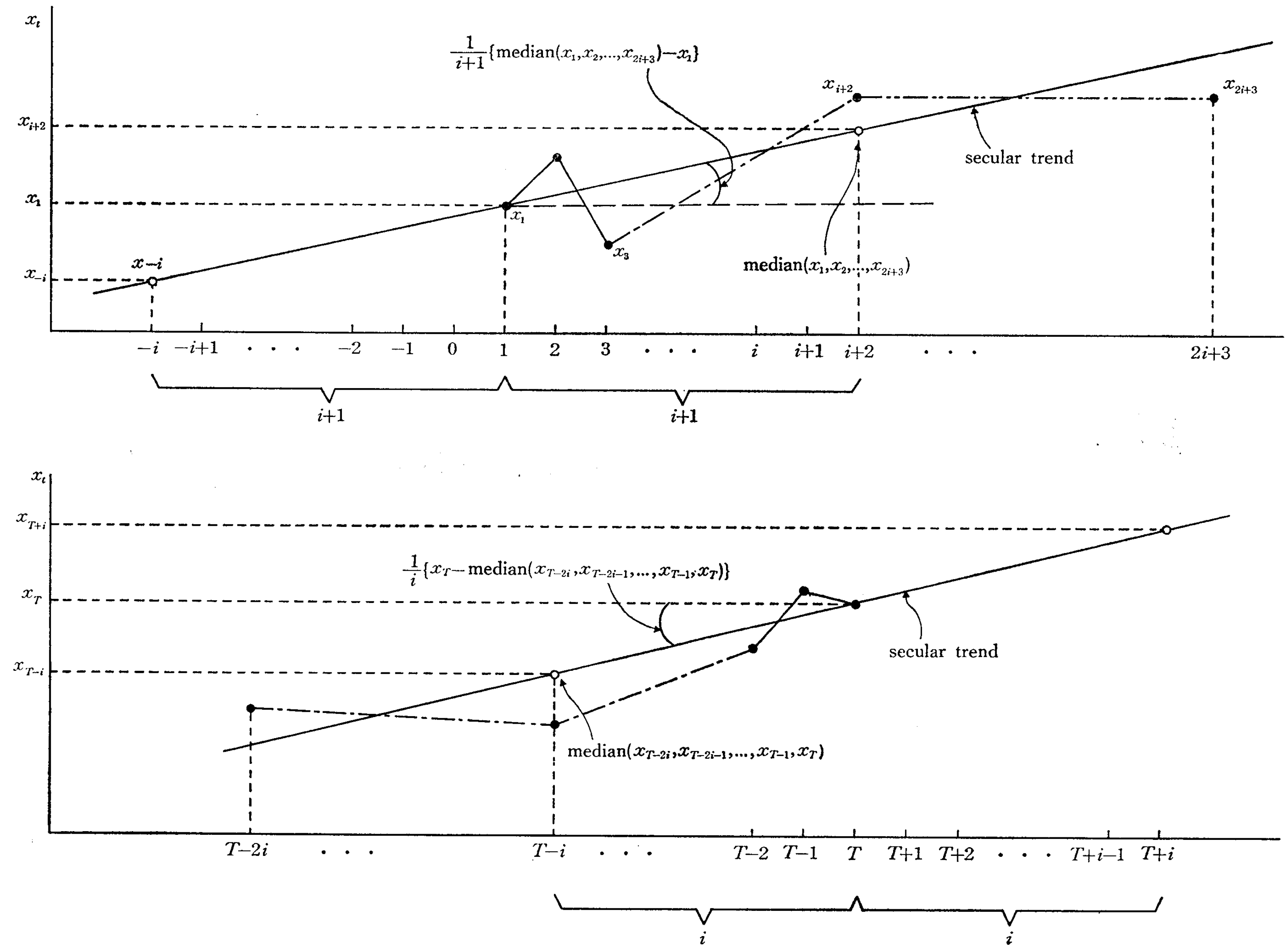

Figure 6 Determination of the values x_{-i} and x_{T+i}.

$$x_1(1),\ x_2(1),\ \cdots,\ x_n(1)$$

and

$$x_{T-n+1}(1),\ x_{T-n+2}(1),\ \cdots,\ x_T(1).$$

Generalized end point rule is the rule for obtaining the values (namely, the values of $2n+1$ term medians) which are

$$x_1(1),\ x_2(1),\ \cdots,\ x_n(1)$$

and

$$x_{T-n+1}(1),\ x_{T-n+2},\ \cdots,\ x_T(1)$$

by applying the generalized smoothing rule to the new $2n$ values :

$$x_{-n+1},\ x_{-n+2},\ \cdots,\ x_{-1},\ x_0$$

and

$$x_{T+1},\ x_{T+2},\ \cdots,\ x_{T+n-1},\ x_{T+n}$$

and already known values :

$$x_1,\ x_2,\ \cdots,\ x_n,\ x_{n+1}$$

and

$$x_{T-n},\ x_{T-n+1},\ \cdots,\ x_{T-1},\ x_T.$$

The new $2n$ values written above are calculated by one of the three kinds of generalized extrapolation rules. Therefore, we can construct three kinds of generalized end point rule : (1) the generalized Tukey's end point rule, (2) the median end point rule, and (3) the median-trend end point rule.

(1) Generalized Tukey's End Point Rule

In the generalized Tukey's end point rule, the generalized Tukey's extrapolation rule and the generalized smoothing rule are utilized. And the process of manipulation given by this rule is written by the following steps of operation.

Step 1. The following new $2n$ values are obtained by the generalized Tukey's extrapolation rule.

$$x_{-n},\ x_{-n+1},\ \cdots,\ x_{-1} \text{ and } x_{T+1},\ x_{T+2},\ \cdots,\ x_{T+n}$$

Step 2. $2n+1$ term medians which are shown by

$$x_1(1),\ x_2(1),\ \cdots,\ x_n(1) \text{ and } x_{T-n+1}(1),\ x_{T-n+2}(1),\ \cdots,\ x_T(1)$$

are obtained by the generalized smoothing rule.

(2) Median End Point Rule

In the median end point rule, the median extrapolation rule and the generalized smoothing rule are utilized. And the process of manipulation given by this rule is written by the following steps of operation.

Step 1. The following new $2n$ values are obtained by the median extrapolation rule.

$$x_{-n},\ x_{-n+1},\ \cdots,\ x_{-1} \text{ and } x_{T+1},\ x_{T+2},\ \cdots,\ x_{T+n}$$

Step 2. $2n+1$ term medians which are

$$x_1(1),\ x_2(1),\ \cdots,\ x_n(1) \text{ and } x_{T-n-1}(1),\ x_{T-n+2}(1),\ \cdots,\ x_T(1)$$

are obtained by the generalized smoothing rule.

(3) Median-Trend End Point Rule

In the median-trend end point rule, the median-trend extrapolation rule and the generalized smoothing rule are utilized. And the process of manipulation given by this rule is written by the following steps of operation.

Step 1. The following new $2n$ values are obtained by the median-trend extrapolation rule.

$$x_{-n},\ x_{-n+1},\ \cdots,\ x_{-1} \text{ and } x_{T+1},\ x_{T+2},\ \cdots,\ x_{T+n}$$

Step 2. $2n+1$ term medians which are

$$x_1(1),\ x_2(1),\ \cdots,\ x_n(1) \text{ and } x_{T-n+1}(1),\ x_{T-n+2}(1),\ \cdots,\ x_T(1)$$

are obtained by the generalized smoothing rule.

7. Generalization of Tukey's Smoothing Method: Iterative Smoothing Method by Median of $2n+1$ Terms

We can define generalized iterative moving median method by combining two rules: generalized end point rule and generalized smoothing rule.

According to equation (5.2), generalized smoothing rule is defined for obtaining the median of $2n+1$ terms. Therefore, we can easily define generalized iterative moving median method by median of $2n+1$ terms. Incidentally, we can define three kinds of methods as generalized iterative moving median method, since we have three kinds of generalized end point rules (the generalized Tukey's end point rule, the median end point rule, and the median-trend end point rule) and one generalized smoothing rule. These three kinds of methods are the "generalized Tukey's method", the "median method" and the "median-trend method".

(1) Generalized Tukey's Method

The generalized Tukey's method which is prepared for obtaining a smoothed time series is obtained by combining the generalized Tukey's end point rule and the generalized smoothing rule, and it is shown by the following steps of operation.

Step 1. The values $x_{-n+1},\ x_{-n+2},\ \cdots,\ x_0$, and $x_{T+1},\ \cdots,\ x_{T+n}$ are calculated by the generalized Tukey's end point rule or by equations (4.7) and (4.10) defined in the generalized Tukey's extrapolation rule.

Step 2. The smoothed value $x_t(1)$ ($t=1,\ 2,\ \cdots,\ T$) is obtained by equation (5.2) defined in the generalized smoothing rule.

Step 3. $x_t(1)$ is regarded as x_t.

Step 4. $x_t(1)$ is again calculated from x_t obtained at *Step* 3 by using the operations defined at steps from *Step* 1 to *Step* 3, until the x_t found at *Step* 3 is not changed even if the operations are given by the steps from *Step* 1 to *Step* 3.

The smoothed time series obtained by *Step* 4 is regarded as the final smoothed time series, $\hat{x}_t$ $(t=1, 2, \cdots, T)$.

"Mesa" will be also found in the final smoothed time series defined here. The end point rule for mesa which is constructed by manipulation : (i) splitting the mesa found in the final smoothed time series, and (ii) modification of the values of the components of the mesa by the generalized Tukey's extrapolation rule and generalized smoothing rule is applied to only the components of the mesa, if we want to modify the values of the mesa. When the mesa is modified, the values of the components of the whole series observed are again modified by the operations defined by the steps from *Step* 1 to *Step* 4 which are written above.[5]

(2) Median Method

The median method which is prepared for obtaining a smoothed time series is obtained by combining the median end point rule and the generalized smoothing rule, and it is shown by the following steps of operation.

Step 1. The values x_{-n+1}, x_{-n+2}, $\cdots$, x_o, and x_{T+1}, $\cdots$, x_{T+n} are calculated by the median end point rule or by equations (4.11) and (4.12) defined in the median extrapolation rule.

Step 2. The smoothed value $x_t(1)$ $(t=1, 2, \cdots, T)$ is obtained by equation (5.2) defined in the generalized smoothing rule.

Step 3. $x_t(1)$ is regarded as x_t.

Step 4. $x_t(1)$ is again calculated from x_t obtained at *Step* 3 by using the operations defined at the steps from *Step* 1 to *Step* 3, until the x_t found at *Step* 3 is not changed even if the operations are given by the steps from *Step* 1 to *Step* 3.

The smoothed time series obtained by *Step* 4 is regarded as the final smoothed time series, $\hat{x}_t$ $(t=1, 2, \cdots, T)$.

"Mesa" will be also found in the final smoothed time series defined here. The end point rule for mesa which is constructed by manipulations : (i) splitting the mesa found in the final smoothed time series, and (ii) modification of the values of the components of the mesa by the median extrapolation rule and generalized smoothing rule is applied to only the components of the mesa, if we want to modify the values of the mesa. When the mesa is modified, the values of the components of the whole series observed are again modified by the operations defined by the steps from *Step* 1 to *Step* 4 which are written above.

(3) Median-Trend Method

The median-trend method is also prepared for obtaining a smoothed time series is obtained by combination of the median-trend end point rule and generalized smoothing rule, and it is shown by the following steps of operation.

Step 1. The values x_{-n+1}, x_{-n+2}, $\cdots$, x_o, and x_{T+1}, $\cdots$, x_{T+n} are calculated by the median-trend end point rule or by equations (4.12) and (4.13) defined in the median-trend extrapolation rule.

Step 2. The smoothed value $x_t(1)$ ($t=1, 2, \cdots, T$) is obtained by equation (5.2) defined in the generalized smoothing rule.

Step 3. $x_t(1)$ is regarded as x_t.

Step 4. $x_t(1)$ is again calculated from x_t obtained at *Step* 3 by using the operations defined at steps from *Step* 1 to *Step* 3, until the x_t found at *Step* 3 is not changed even if the operations are given by the steps from *Step* 1 to *Step* 3.

The smoothed time series obtained by *Step* 4 is regarded as the final smoothed time series, $\hat{x}_t$ ($t=1, 2, \cdots, T$).

"Mesa" will be also found in the final smoothed time series defined here. The end point rule which is constructed by manipulations : (i) splitting the mesa found in the final smoothed time series, and (ii) modification of the values of the components of the mesa by the median-trend extrapolation rule and generalized smoothing rule is applied to only the components of the mesa, if we want to modify the values of the mesa. When the mesa is modified, the values of the components of the whole series observed are again modified by the operations defined by the steps from *Step* 1 to *Step* 4 which are written above.

8. Iterative Smoothing Method by Median of $2n$ Terms: Average Method

In 7, three kinds of iterative smoothing methods, namely, the generalized Tukey's method, the median method, and the median-trend method are defined for obtaining a smoothed time series by using median of $2n+1$ terms. Therefore, it is difficult to find a smoothed time series which consists of medians of $2n$ terms by applying the iterative smoothing methods mentioned above.

Consequently, we must prepare a method of smoothing by median of $2n$ terms specifically. The specific iterative smoothing method by median of $2n$ terms proposed here is given by the following steps of operation, and it is called here "average method".

Step 1. The final smoothed time series obtained by medians of $2n-1$ terms and $2n+1$ terms ($n=1, 2, \cdots$) are calculated by using one of the generalized Tukey's method, the median method, and the median-trend method. The values of the components of the final smoothed time series obtained by medians of $2n-1$ terms and $2n+1$ terms are expressed by $\hat{x}^{(2n-1)}$ and $\hat{x}^{(2n+1)}$.

Step 2. The final smoothed time series obtained by medians of $2n$ terms is calculated by the following equation :

$$\hat{x}^{(2n)}=\frac{1}{2}(\hat{x}^{(2n-1)}+\hat{x}^{(2n+1)}) \tag{6.1}$$

9. An Example of Application of the Median-Trend Method and the Average Method

Here, the median-trend method and the average method are applied to smooth an actual

Table 1 The process of obtaining the final smoothed time series composed by 3 term moving median

(million tons)

year	original value	calculation for smoothing	
		1st	2nd
	(3.4)	(3.4)	(3.4)
1950	3.4	3.4	3.4
1951	3.4	3.4	3.4
1952	3.7	3.5	3.5
1953	3.5	3.5	3.5
1954	3.4	3.4	3.4
1955	3.0	3.4	3.4
1956	4.0	4.0	4.0
1957	4.9	4.8	4.8
1958	4.8	4.9	4.9
1959	5.0	5.0	5.0
1960	5.8	5.8	5.8
1961	6.5	6.5	6.5
1962	6.9	6.6	6.6
1963	6.6	6.9	6.9
1964	7.7	7.7	7.7
1965	8.2	8.2	8.2
1966	8.6	8.6	8.6
1967	10.0	10.0	10.0
1968	11.9	11.9	11.9
	(12.8)	(12.8)	(12.8)

Note : The value in parentheses is the value obtained by the median-trend extrapolation rule.
The values obtained by the 2nd calculation for smoothing is the final smoothed time series.

time series. The actual time series observed is the time series of quantity of stock in commercial warehouse of Japan from 1950 to 1968[6]. In this analysis, the final smoothed time series composed by 3, 4, and 5 term moving median.

According to the statistics of the quantity of stock in commercial warehouse of Japan, the time series of the quantity is a time series which has cyclical fluctuations——local increasing and decreasing tendency——as shown in Tables 1, 2, and 3.

Tables 1 and 2 show the fact that the final smoothed time series composed by 3 term moving median was found after the first calculation for smoothing and the final smoothed time series composed by 5 term moving median was found after the second calculation for smoothing.

In the final smoothed time series composed by 5 terms moving median, cyclical fluctuations which were found in the original time series disappeared (Figure 7 (D)), while in the final smoothed time series composed by 3 and 4 term moving medians (espetially, at the beginning of these smoothed time series), weak cyclical fluctuations remained as shown in Figures 7(B) and (C). The reason why cyclical fluctuations disappeared in the final smoothed time series composed by 5 terms moving median is that the original time series had cyclical fluctuations whose period was 5 years (Figure 7(A)).

Table 2 The process of obtaining the final smoothed time series composed by 5 term moving median (million tons)

year	original value	calculation for smoothing		
		1st	2nd	3rd
	(3.4)	(3.4)	(3.4)	(3.4)
	(3.4)	(3.4)	(3.4)	(3.4)
1950	3.4	3.4	3.4	3.4
1951	3.4	3.4	3.4	3.4
1952	3.7	3.4	3.4	3.4
1953	3.5	3.4	3.4	3.4
1954	3.4	3.5	3.4	3.4
1955	3.0	3.5	4.0	4.0
1956	4.0	4.0	4.0	4.0
1957	4.9	4.8	4.8	4.8
1958	4.8	4.9	4.9	4.9
1959	5.0	5.0	5.0	5.0
1960	5.8	5.8	5.8	5.8
1961	6.5	5.8	5.8	5.8
1962	6.9	6.6	6.6	6.6
1963	6.6	6.9	6.9	6.9
1964	7.7	7.7	7.7	7.7
1965	8.2	8.2	8.2	8.2
1966	8.6	8.6	8.6	8.6
1967	10.0	10,0	10.0	10.0
1968	11.9	11.9	11.9	11.9
	(13.8)	(13.8)	(13.8)	(13.8)
	(15.2)	(15.2)	(15.2)	(15.2)

Note : The values obtained by the 3rd calculation for smoothing is the final smoothed time series.

Table 3 The process of obtaining the final smoothed time series composed by 4 term moving median

year	original value	final smoothed time series		Average
		Table 1	Table 2	
1950	3.4	3.4	3.4	3.40
1951	3.4	3.4	3.4	3.40
1952	3.7	3.5	3.4	3.45
1953	3.5	3.5	3.4	3.45
1954	3.4	3.4	3.4	3.40
1955	3.0	3.4	4.0	3.70
1956	4.0	4.0	4.0	4.00
1957	4.9	4.8	4.8	4.80
1958	4.8	4.9	4.9	4.90
1959	5.0	5.0	5.0	5.00
1960	5.8	5.8	5.8	5.80
1961	6.5	6.5	5.8	5.80
1962	6.9	6.6	6.6	6.60
1963	6.6	6.9	6.9	6.90
1964	7.7	7.7	7.7	7.70
1965	8.2	8.2	8.2	8.20
1966	8.6	8.6	8.6	8.60
1967	10.0	10.0	10.0	10.00
1968	11.9	11.9	11.9	11.90

Note : The values of the time series found in the column of "Average" is the average of the values of the final smoothed time series in Tables 1 and 2 and is the final smoothed time series composed by 4 term moving median.

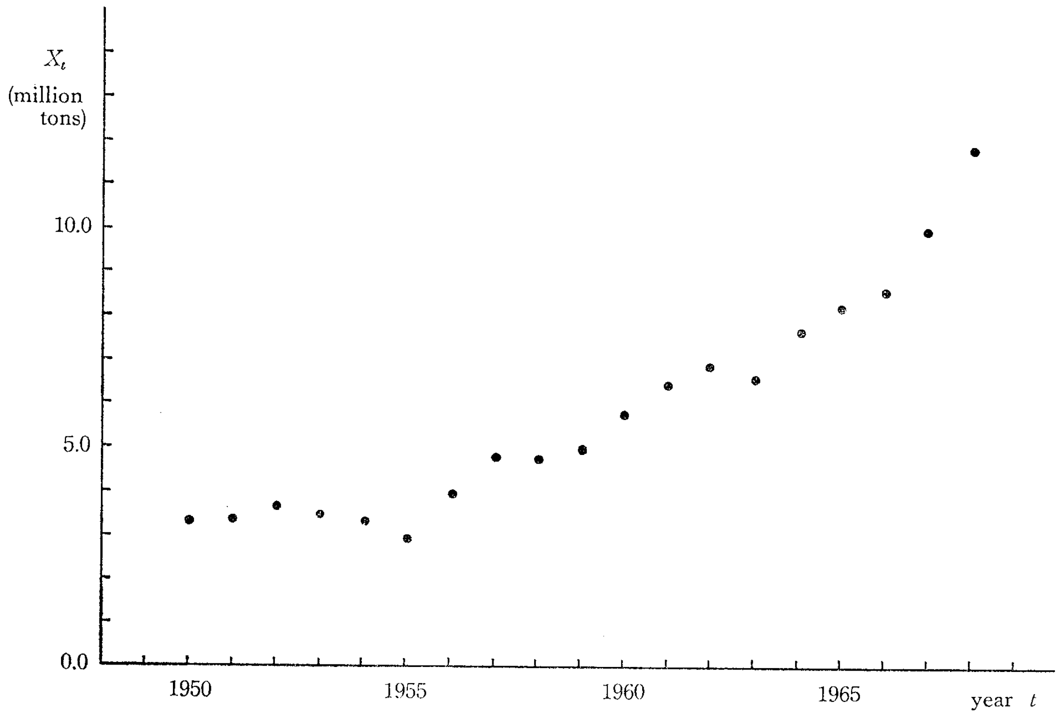

Figure 7(A) The original time series to be smoothed.

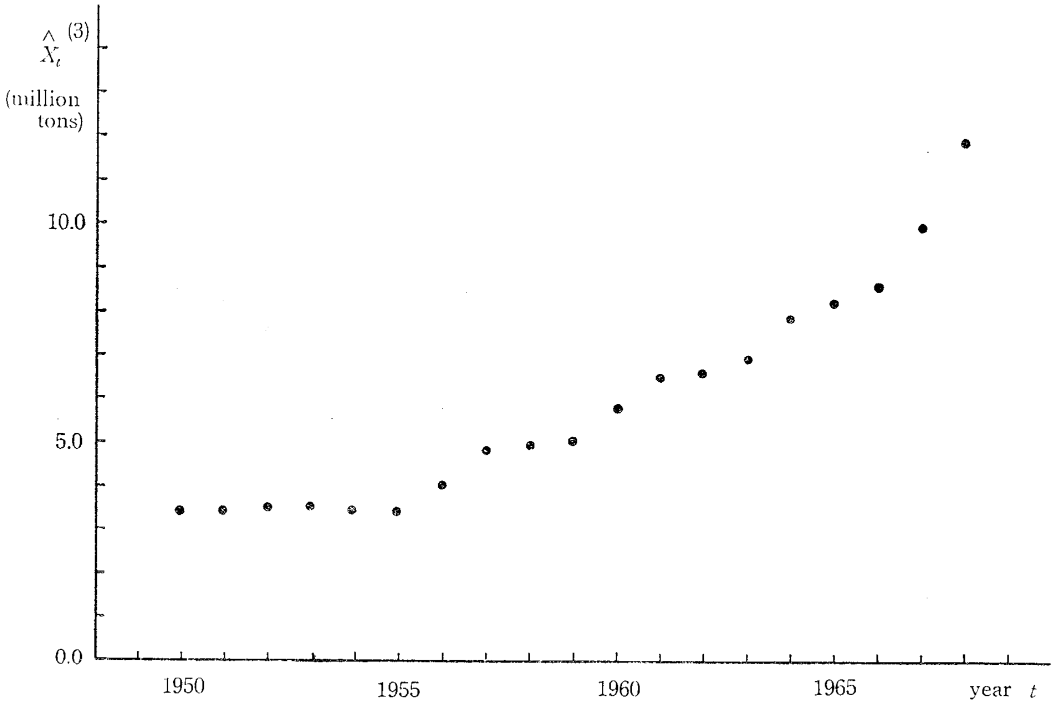

Figure 7(B) The final smoothed time series composed by 3 term moving median.

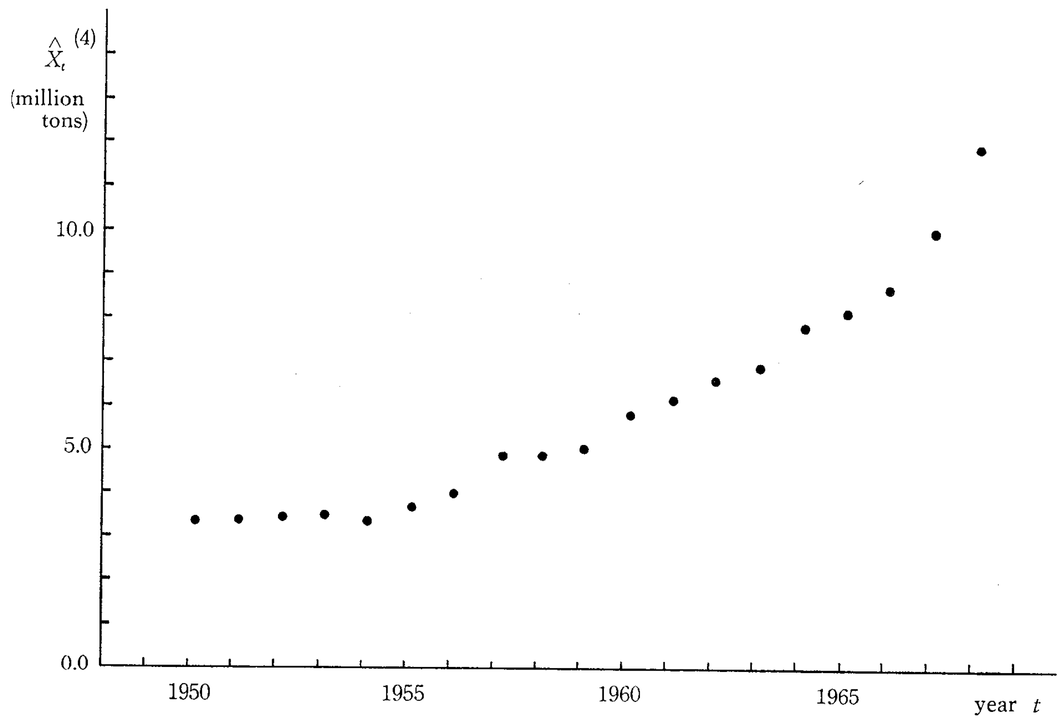

Figure 7(C) The final smoothed time series composed by 4 term moving median.

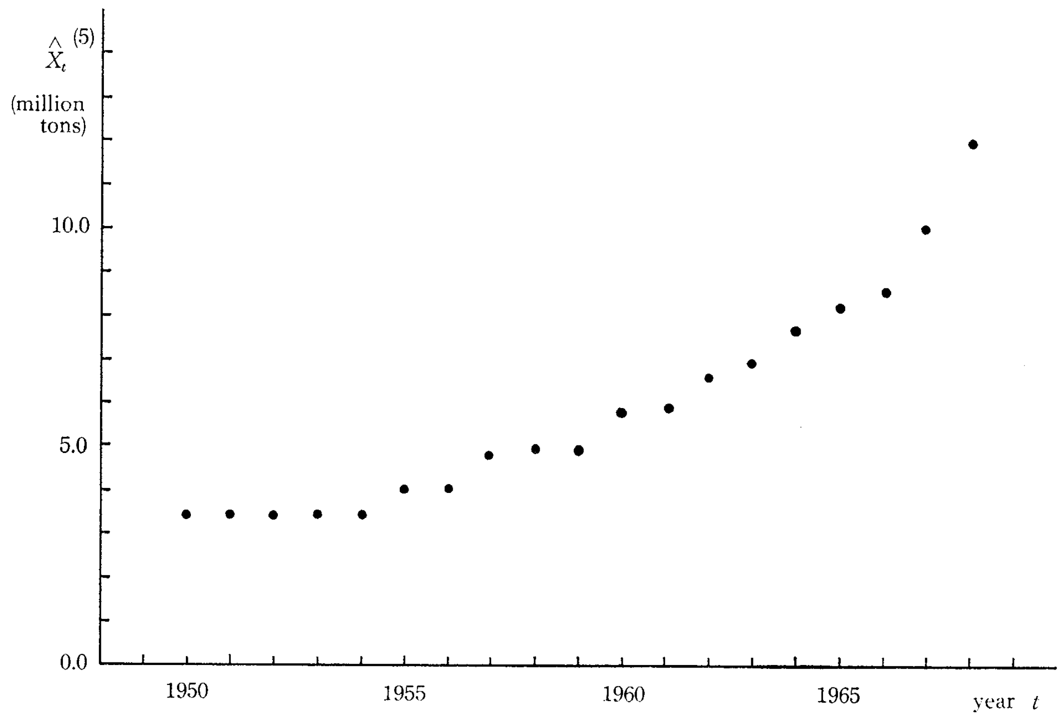

Figure 7(D) The final smoothed time series composed 5 term moving median.

10. Conclusion

The purpose of this paper was to expand and generalize the Tukey's method for smoothing time series and to find a new generalized smoothing method.

According to the Tukey's method, the final smoothed time series composed by 3 term medians which is obtained by iterative calculation is stable, and the length of the smoothed time series obtained is not shortened, namely, if the length of the original time series to be smoothed is T, then the length of smoothed time series is also T. These characteristics of the smoothed time series are very important.

The first characteristic of the smoothed time series——stability of the final smoothed time series——is given by iterative application of moving median to the original time series observed. And the second characteristic of the smoothed time series is given by application of the end point rule to the time series to be smoothed.

In this paper, when I generalize the Tukey's method, I tried to decompose the end point rule into (1) exptrapolation rule and (2) smoothing rule and generalized these rules.

As the result of generalization of the extrapolation rule, three kinds of generalized extrapolation rules, namely, (i) the generalized Tukey's extrapolation rule, (ii) the median extrapolation rule, and (iii) the median-trend extrapolation rule were obtained. Smoothing rule was also generalized, and as the result of the generalization, generalized smoothing rule was obtained. By combining each of the three kinds of generalized extrapolation rules and generalized smoothing rule, three kinds of generalized end point rules, (i) the generalized Tukey's end point rule, (ii) the median end point rule, and (iii) the median-trend end point rule were obtained. And, finaly, three kinds of generalized iterative smoothing method, (i) the generalized Tukey's method, (ii) the median method, and (iii) the median-trend method were constructed by utilizing each of the three kinds of generalized end point rules and generalized smoothing rule.

Incidentally, the three kinds of generalized iterative smoothing methods mentioned above can be used only for obtaining a smoothed time series which is composed by $2n+1$ term medians, since the generalized extrapolation rules or generalized end point rules and the generalized smoothing rule can be used for obtaining $2n+1$ term median. But, in this paper, a generalized iterative smoothing method which can be used for obtaining a smoothed series which is composed by $2n$ term medians was proposed.

From the application of the median-trend method, one of the generalized iterative smoothing method, to an actual time series, I could find an example in which the cyclical fluctuations of the original time series were cancelled out, when a smoothed time series which was composed by 5 term medians was obtained. From this result, it is concluded that the generalized iterative smoothing method proposed here is very effective to cancel out cyclical fluctuation found in the original time series observed whose period is not 3.

References

〔1〕 McNeil, Donald R. : *Interactive Data Analysis, A Practical Primer,* New York, John Wiley & Sons, 1977.

〔2〕 Tukey, John W. : *Exploratory Data Analysis,* Reading, Massachusetts, Addison-Wesley Publishing Company, 1977.

〔3〕 Unyu Sho, Daijin Kanbo, Tokei Chosa Bu (Japan, Ministry of Transportation, Minister's Secretariat, Department of Statistical Survey) ed. : *Unyu Keizai Tokei Yoran (Statistical Handbook of Transpotation Economics), Showa 42 nen Ban (1967),* Tokyo, Okura Sho, Insatsu Kyoku (Ministry of Finance, Printing Bureau), 1968.

〔4〕 Unyu Sho, Daijin Kanbo, Tokei Chosa Bu (Japan, Ministry of Transportation, Ministar's Secretariat, Department of Statistical Survey) ed. : *Unyu Keizai Tokei Yoran (Statistical Handbook of Transportation Economics), Showa 44 nen Ban (1969),* Tokyo, Okura Sho, Insatsu Kyoku (Ministry of Finance, Printing Bureau), 1970.

Notes

1) Tukey〔2〕pp. 204-264.
2) McNeil〔1〕p. 123.
3) Tukey〔2〕pp. 204-264, and, McNeil〔1〕p. 122.
4) McNeil〔1〕122.
5) When the "i" of the generalized Tukey's method is 1, this method becomes the Tukey's method. Therefore, it can be said that the Tukey's method is a specific case of the generalized Tukey's method.
   Incidentally, the deffinition of "mesa" is also, here, "a flat two point local maximum or minimum".
6) Unyu Sho〔3〕p. 90, and Unyu Sho〔4〕p. 80. The quantity of the stock in commercial werehouse is average of the quantity of the stock at the end of each month in a year.

要　約

鈴木啓祐：「時系列の反復平滑化法：テューキイの方法の一般化」,『流通経済大学論集』第18巻第3号, 1984年, 1-21ページ。

テューキイ（John W. Tukey）は，3項の中位数（3項移動中位数）による時系列の反復平滑化法を提唱した。この方法によると，興味あることには，完全に平滑化された時系列は，反復平滑化法が追加的に適用されても，もはや，それ以上平滑化されないという性質，すなわち，反復平滑化に対する「安定性」をもっている。

この論文では，このテューキイの反復平滑化法の拡張，あるいは，一般化が試みられた。一般化は，2つの点についてなされた。

その第1は，3項の中位数による平滑化から $2n+1$ 項の中位数による平滑化への一般化である。この一般化のためにテューキイの反復平滑化法において重要な役割を演じる端点ルール（end point rule）が一般化された。この一般化のためには，端点ルールを外挿ルール（extrapolation rule）と平滑化ルール（smoothing rule）とに分割し，それぞれのルールを一般化した。一般化された外挿ルールとしては，一般化テューキイ外挿ルール（generalized Tukey's extrapolation rule），中位数外挿ルール（median extrapolation rule），および，中位数・トレンド外挿ルール（median-trend extrapolation rule）の3種を提唱した。一方，一般化された平滑化ルールとしては，一般化平滑化ルール（generalized smoothing rule）を提唱し

た。これらを再び統合して，3種の一般化端点ルール，すなわち，一般化テューキイ端点ルール（generalized Tukey's end point rule），中位数端点ルール（median end point rule），中位数・トレンド端点ルール（median-trend end point rule）が構成された。

これらの一般化端点ルールと一般化平滑化ルールとを組合せて，3種の一般化された $2n+1$ 項の中位数による反復平滑化法，すなわち，一般化テューキイ法（generalized Tukey's method），中位数法（median method），中位数・トレンド法（median-trend method）が構成された。

第2の一般化は，$2n$ 項の中位数による反復平滑化法（iterative smoothing method by median of $2n$ terms）の構築である。これは，第1の一般化で得られた $2n+1$ 項の中位数による反復平滑化法（iterative smoothing method by median of $2n$ terms）を基礎として構成することとし，その方法を，その特徴から平均法（average method）と名づけることにした。

ここで提唱した新しい一般化反復平滑化法が現実のデータをどのように平滑化するかということを明確にするため，昭和25年から43年までの19項からなる年別全国営業倉庫貨物在庫量（各年の毎月月末在庫量の平均）に対して，中位数・トレンド法を適用して平滑化をおこなってみた。中位数の項数は，3項，4項，および5項である。予想されたように，平滑化の結果得られた時系列は安定化した。しかも，3項および4項の中位数による平滑化された時系列では，原系列に見られた波動的変動がまだ残存していたが，5項の中位数による平滑化された時系列では，その波動的変動がとり除かれるという事実を見いだすことができた。

このことから，テューキイの3項移動中位数による反復平滑化法の一般化は，週期が3項以外の長さをもつ波動的変動のみられる時系列から，その波動的変動をとり除くための有力な方法となり得るといえよう。