
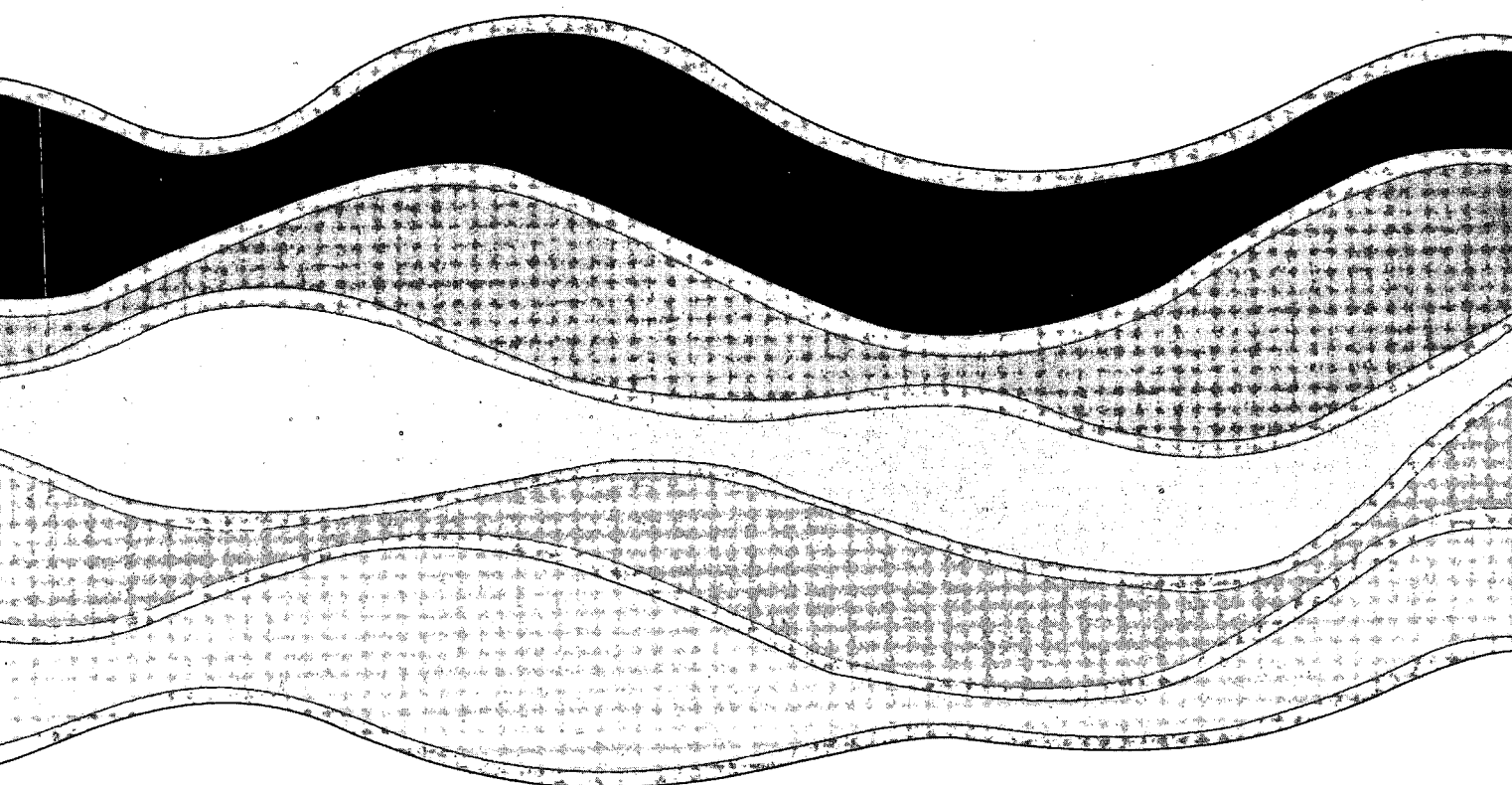



**NATIONAL
WATER
RESEARCH
INSTITUTE**



**INSTITUT
NATIONAL
de RECHERCHE
sur les
EAUX**



**THE ESTIMATION OF WATER QUALITY
CHANGES BETWEEN LOCATIONS
WITHIN A RIVER SYSTEM**

**S.R. Esterby, A.H. El-Shaarawi and
H.O. Block**

NWRI Contribution No. 90-45

TD
226
N87
No. 90-
45
c. 1

**THE ESTIMATION OF WATER QUALITY
CHANGES BETWEEN LOCATIONS
WITHIN A RIVER SYSTEM**

**S.R. Esterby¹, A.H. El-Shaarawi² and
H.O. Block³**

**¹Lakes Research Branch
²Rivers Research Branch
National Water Research Institute
Burlington, Ontario L7R 4A6**

**³Water Quality Branch
Western and Northern Region
Calgary, Alberta, Canada T2P 4A6**

NWRI Contribution No. 90-45

MANAGEMENT PERSPECTIVE

The problem of estimating differences in water quality between sampling locations, using data from a water quality sampling network, is addressed. Sodium and nitrate plus nitrite concentrations measured between 1984 and 1986 at five locations on the Bow and South Saskatchewan Rivers, varying from a headwater site to a site below agricultural and urban regions, are used to illustrate statistical methods. It is shown that differences between locations can be estimated from these data using models which account for seasonality due to the hydrological regime. For example, the two lower stations had significantly different sodium concentrations early and late in the year, but not in the summer. This case study should be of interest to anyone trying to assess spatial differences in water quality along a transect or a river system where data are collected over time at relatively long intervals (in this case study, one month). It is to be incorporated into the report of the departmental task force on the design and assessment of water quality data collection programs.

PERSPECTIVES DE GESTION

On examine le problème de l'estimation des différences dans la qualité de l'eau entre les stations d'échantillonnage, grâce à des données provenant d'un réseau d'échantillonnage pour la qualité de l'eau. Les concentrations de sodium et de nitrate ainsi que de nitrite, mesurées entre 1984 et 1986 en cinq points des rivières Bow et Saskatchewan-Sud, présentant des variations à partir d'une station en amont jusqu'à une station située en dessous de régions agricoles et urbaines, servent à illustrer les méthodes statistiques. On montre qu'il est possible d'évaluer les écarts entre les divers points grâce à ces données et à l'aide de modèles qui tiennent compte des caractéristiques saisonnières correspondant au régime hydrologique. Par exemple, les deux stations inférieures présentent des concentrations de sodium sensiblement différentes au début et à la fin de l'année, mais non en été. Cette étude de cas devrait pouvoir intéresser quiconque veut évaluer les différences spatiales dans la qualité de l'eau le long d'un transect ou d'un bassin de rivière, dont les données sont recueillies dans le temps à des intervalles relativement longs (dans la présente étude, un mois). Les résultats de l'étude seront incorporés au rapport du groupe de travail ministériel sur la conception et l'évaluation des programmes de collecte de données pour l'estimation de la qualité de l'eau.

ABSTRACT

Five locations on the Bow and South Saskatchewan Rivers, varying from a headwater site to a site below urban and agricultural regions, are compared for differences between concentration and flux of sodium and nitrate plus nitrite for the years 1984 to 1986. The hydrological regime displays marked seasonality and differences between upper and lower locations due to differing sources over the year. Two general methods of statistical analysis, both of which account for seasonality, are used to compare locations. One method, based on matching observations made at the same point on the hydrological regime, is used for the three upper locations. The sampling dates are matched and subsequently the differences are analyzed using matched pairs procedures such as the t test or the Wilcoxon signed-rank test. For locations where such matching of observations is not possible, concentration or flux at the same point on the hydrological regime is estimated by regression methods and the difference between estimates is assessed in the framework of regression analysis. Comparisons between locations, for each of three periods in the year, were made because the major flow events correspond to melting of snow on the prairies, melting of snow in the mountains and an autumnal increase at the lower stations, where irrigation waters are returned to the river. For the three years of data analyzed here, differences between the three upper and two lower locations are large, but depend upon the period of the year. Differences within the group of three upper or the group of two lower locations also depend upon the time of year.

RÉSUMÉ

Cinq points sur les rivières Bow et Saskatchewan-Sud, à partir d'une station en amont jusqu'à une station située en dessous de régions agricoles et urbaines, ont fait l'objet de comparaisons pour évaluer les écarts dans les concentrations et les flux de sodium et de nitrate ainsi que de nitrite pendant les années 1984 à 1986. Le régime hydrologique présente des caractéristiques saisonnières et des différences très marquées entre les points en amont et en aval, par suite de la variabilité des sources au cours de l'année. Deux méthodes générales d'analyse statistique, tenant chacune compte des caractéristiques saisonnières, servent à comparer les différents points. L'une des méthodes, consistant à comparer les observations faites en un même point du régime hydrologique, est utilisée pour les trois sites supérieurs. Les données d'échantillonnage sont appariées, et les différences sont ultérieurement analysées à l'aide de méthodes utilisant ces données appariées, comme le test-t ou le test de Wilcoxon pour observations appariées. Dans le cas des endroits où il n'est pas possible d'apparier les observations, la concentration ou le flux en un même point du régime hydrologique sont évalués grâce à des méthodes de régression, les écarts entre les évaluations étant déterminés dans le cadre de l'analyse par régression. On a procédé à des comparaisons entre les stations, pour chacune des trois périodes de l'année, car les principaux événements de flux correspondent au dégel dans les Prairies, au dégel dans les montagnes et aux hausses automnales dans les stations inférieures, où les eaux d'irrigation retournent à la rivière. Pour les trois années de données analysées ici, les différences entre les trois stations supérieures et les deux stations inférieures sont importantes, mais elles dépendent de la période de l'année. Les écarts à l'intérieur du groupe des trois stations supérieures ou du groupe des deux stations inférieures sont également fonction de la période de l'année.

INTRODUCTION

Along the course of a river system, changes occur in the hydrological regime and in the loading from point and non-point sources. To answer questions about the differences in water quality at different locations, these changes need to be taken into account in both the formulation of more specific questions or objectives and in the determination of the methods of data analysis. Answers to such questions are often sought using data from water quality sampling networks where samples have been collected at selected locations at some specified frequency. Data from the five sampling locations on the Bow and South Saskatchewan Rivers, which were described in the previous chapter, are used here.

Objective

Estimate differences in the concentration and flux of selected water quality variables measured at locations upstream and downstream from urban or agricultural areas, taking into account features of the hydrological regime.

Description of the Data

Nitrate plus nitrite ($\text{NO}_3 + \text{NO}_2$) and sodium (Na) concentrations, observed at the five previously described Bow and South Saskatchewan Rivers stations during 1984 to 1986, were chosen since the analytical methodology was consistent for both variables during these years. There is no overlap between the Na concentrations of the pair of locations BN0001

and AK0001, and the three locations BA0011, BE0013, and BH0017, but there is overlap between the two groups for $\text{NO}_3 + \text{NO}_2$ during the period of low $\text{NO}_3 + \text{NO}_2$ concentrations (Figures 1a and 2a). In the previous chapter, seasonal patterns were shown to exist for both variables at all locations, and this can be seen more clearly for three stations with lower concentrations in Figures 1b and 2b where only locations BA0011, BE0013, and BH0017 are plotted together.

Since these five stations span the Bow and South Saskatchewan Rivers from the headwater site in the mountains to the South Saskatchewan site below the irrigation region, the flow regime varies from location to location. Periods, during which the flow increases at some or all of the locations, appear to correspond to snow melting on the prairies, the later melting in the mountains and the return of irrigation waters. The importance of this is not merely the change in water volumes, but also the effects of difference sources of the water on the concentration of substances. The complete discharge record over a year would show these periods clearly, but even the discharge level on the day of water quality sampling, which is available in the water quality data files, is useful.

Each of the three years analyzed here was divided into three periods by using the plots of flow versus time and the following criteria. The first period of the year includes sampling times prior to the increased discharge at BA0011 and thus would appear to include increased flow due to the prairie snow melt but not due to the mountain snow melt. During this period, the flow remains approximately constant at BA0011, BE0013, and BH0017, but goes through a peak at BN0001 and

AK0001 (Figure 3). The second period shows a peak at all five stations and appears to result from the melting of mountain snow. In the third period, the flow is increasing again at BN0001 and AK0001 but is near spring levels at BA0011, BE0013, and BH0017. The dates of first and/or last sampling for each of the periods determined in this way and the number of sampling dates in each period are shown for each location and year in Tables 1 and 2.

The flux, obtained by multiplying concentration and flow at each sampling date, is plotted for Na and $\text{NO}_3 + \text{NO}_2$ in Figures 4 and 5, respectively. The major difference between $\text{NO}_3 + \text{NO}_2$ concentration and flux patterns over the year for BA0011, BE0013, and BH0017 is the absence of a summer flux minimum due to increased flow during this period. Increases in flux corresponding to spring or summer maxima are also seen at locations BN0001 and AK0001. Similarly for Na, spring or summer flow maxima produce higher summer flux values where concentrations were not high. Relationships between concentration and flow are suggested for sodium in the summer and fall and only for locations BN0001, AK0001, and BE0013, AK0001, respectively (Figures 6 and 7).

STATISTICAL METHODS

A major factor determining the statistical methods available for analysis of these data is the agreement in frequency and day of sampling at the different stations. To avoid confounding concentration differences, due to locations being at different stages of the hydrological regime, with differences due to location, comparisons need to be made

between measurements when each location is at the same stage of the hydrological regime. Sampling at locations BA0011, BE0013, and BH0017 was done monthly with the times of sampling in a given month not more than two days apart except for one three-day and two seven-day lags. The flow events also occurred at about the same time. Thus methods based on matching sampling dates can be used for comparing these three stations. There was poorer agreement between sampling dates at locations BN0001 and AK0001, and more frequent but irregular sampling at location AK0001. Hence comparisons between the latter two locations and among all five locations, require some method of estimating concentrations on a common date.

Regression Methods

A general method of analysis which permits conclusions to be drawn separately for the three periods corresponding to different flow events is a regression analysis based on the following model. For sampling done within one of the periods,

$$Y_{iks} = \mu + \alpha_{ik} + \beta_{ik} t_{iks} + \epsilon_{iks} \quad (1)$$

where Y_{iks} = concentration on day s in year i at location k
 μ = general mean
 α_{ik} = shift in mean due to location k and year i
 β_{ik} = change per unit time at location k in year i
 ϵ_{iks} = error term

and the model is fitted to the data from location k only. The second and third terms on the right-hand side account for non-constancy of concentration within the period, and linearity in time is an approximation which can be considered only because the year has been divided into periods.

If the average difference between locations over all three years is sought and the same slope is assumed, then the model is

$$Y_{iks} = \mu + \alpha_k + \beta_k t_{iks} + \epsilon_{iks}. \quad (2)$$

If the β_k are not significantly different from zero, differences between stations are estimated as differences between means via the α_k . If constancy of concentration within a period is an untenable assumption for at least one station, then either a regression line and a mean, or a pair of regression lines must be compared. This can be done by estimating the difference between concentrations at specific times, or in the case of non-intersecting lines, of testing whether one regression is uniformly greater than the other.

The latter test (Tsutakawa and Hewett, 1978) consists of calculating differences between locations k and k' at the two extremes of days within the period, say t_l and t_h . Let \hat{Y}_{kl} denote the estimated concentration at location k at time t_l , which is given by

$$\hat{Y}_{kl} = \hat{\mu} + \hat{\alpha}_k + \hat{\beta}_k t_l \quad (3)$$

where $\hat{\mu}$, $\hat{\alpha}_k$, and $\hat{\beta}_k$ are least squares estimates. The differences are then

$$W_l = \hat{Y}_{kl} - \hat{Y}_{k'l} \quad (4)$$

$$W_h = \hat{Y}_{kh} - \hat{Y}_{k'h}$$

and the variances are

$$\text{var}(W_l) = \text{var}(\hat{Y}_{kl}) + \text{var}(\hat{Y}_{k'l}) \quad (5)$$

$$\text{var}(W_h) = \text{var}(\hat{Y}_{kh}) + \text{var}(\hat{Y}_{k'h}).$$

$$\text{Then } T_r = |W_r| / \{\text{var}(W_r)\}^{\frac{1}{2}} \quad \text{for } r = l, h \quad (6)$$

$$\text{and } T = \min(T_l, T_h) \quad (7)$$

has a t distribution with degrees of freedom $n_k + n_{k'} - 4$, where n_k and $n_{k'}$ are the number of observations for locations k and k' used to fit model (2). A large value of T supports the hypothesis that the concentration at location k is greater than the concentration at location k' over the entire period.

A general test for nonconstancy of location mean level and change per unit time within a period can be accomplished for all five locations with an analysis similar to that of the previous chapter. The model in (1), which can be rewritten to emphasize testing for different means and slopes by station and year,

$$Y_{iks} = \mu + \alpha_{ik} + \beta_k t_{iks} + \beta_{ik} t_{iks} + \epsilon_{iks} \quad (8)$$

is fitted to the data from all five locations together. For

the present data, the variability of the three locations BA0011, BE0013, and BH0017 is much less than that of location BN0001 and AK0001, and thus a weighted analysis is required to satisfy the assumption of constant error variance. Let c_k , $k=1,2,\dots,5$, be constants such that

$$\text{var}(c_k \varepsilon_{iks}) = \sigma^2. \quad (9)$$

Then, if both sides of (8) are multiplied by c_k , regression analysis can be performed on the resulting terms (Draper and Smith, 1981).

Methods Based on Matching Dates

An alternative and simpler analysis is available for locations BA0011, BE0013, and BH0017. Let Y_{ijk} be the concentration at location k in month j of year i . Then comparisons between locations k and k' are based on $d_{ij} = Y_{ijk} - Y_{ijk'}$, where i and j take all possible values for the period of interest. The hypothesis that locations k and k' do not differ, $H_0: \mu_d = 0$, can be tested using the one-sample (or paired-comparison) t -test or the Wilcoxon signed-rank test (Snedecor and Cochran, 1980). In the former,

$$\sqrt{n} \bar{d} / s_d \quad (10)$$

follows a t -distribution with $(n-1)$ degrees of freedom, where \bar{d} and s_d are the mean and standard deviation calculated from the n differences d_{ij} .

The test statistic for the Wilcoxon signed rank test is V , the smaller in absolute value of V_+ =sum of the positive ranks and V_- =sum of the negative ranks, where the signed ranks are obtained by first ranking the absolute values of the differences, and then assigning the sign of d_{ij} to the corresponding rank. Small values of V provide evidence against the hypothesis of no difference between locations. $P(V \leq C)$ can be obtained from tables for $n \leq 20$ (Lehmann and D'Abrera, 1975) or from the normal approximation for $n > 20$, which is

$$Z = \frac{V + 0.5 - E(V)}{\sqrt{\text{var}(V)}} \quad (11)$$

where $E(V) = n(n+1)/4$ and $\text{var}(V) = n(n+1)(2n+1)/24$. If there are ties in the absolute differences, midranks are used, and if there are ties or zero differences, $E(V)$ and $\text{var}(V)$ become

$$E(V) = \{n(n+1) - d_0(d_0+1)\}/4 \quad (12)$$

and

$$\begin{aligned} \text{var}(V) = & \frac{1}{24} \{n(n+1)(2n+1) - d_0(d_0+1)(2d_0+1)\} \\ & - \frac{1}{48} \sum_{i=1}^e d_i(d_i-1)(d_i+1) \end{aligned} \quad (13)$$

where d_0 = number of zero differences and d_i = number of pairs tied at a single value (Lehmann and D'Abrera, 1975). The sign test could also be used, but, since it is less powerful than the Wilcoxon signed-rank test (Lehman and D'Abrera, 1975), it is recommended only when the ranks are not available. An

example of this is given by El-Shaarawi et al. (1985), where the results for some samples were reported as either a trace concentration or a concentration below the limits of detection.

RESULTS

Regression Analysis of Sodium and Nitrate Plus Nitrite Concentrations

The results of fitting model (8) separately for each year and then for all three years together are shown in Table 3. The constants used for weighting (footnotes, Table 3) are the ratios of the standard deviation for BA0011 to the standard deviation of the other locations, where data over the years 1984 to 1986 were used to calculate the standard deviations. With the exception of the summer period for $\text{NO}_3 + \text{NO}_2$ in 1986, all models indicate significant differences between locations. However, in most cases the models show that the mean level does not remain constant within the period (the presence of t_k or t under the column labelled Model in Table 3), and thus to estimate location differences, the month must be specified and the \hat{Y}_{1ks} calculated. Allowing for different mean levels and rates of change over time in each year accounted for significantly more variation only in the early and late periods for $\text{NO}_3 + \text{NO}_2$ concentrations and the late period for the Na concentrations (Table 4). The models assuming different means and time slopes for stations but only a mean shift between years are plotted together for the three periods in Figure 8.

An estimate of the difference between the concentration over the years 1984 to 1986 at stations k and k' can be

obtained from the difference $\hat{Y}_{iks} - \hat{Y}_{ik's}$, where \hat{Y}_{iks} is the estimate of the concentration on day s within the year. In the weighted analysis, the model

$$\begin{aligned} Z_{iks} = Y_{iks} c_k = & \mu c_k + d_{ik} c_k + \beta_k t_{iks} c_k \\ & + \beta_{ik} t_{iks} c_k + \epsilon_{iks} c_k \end{aligned} \quad (14)$$

or a reduced form, is fitted. The residual mean square, s^2 , estimates $\sigma^2 = \text{var}(\epsilon_{iks} c_k)$. Thus

$$\hat{Y}_{iks} - \hat{Y}_{ik's} = \frac{\hat{Z}_{iks}}{c_k} - \frac{\hat{Z}_{ik's}}{c_{k'}} \quad (15)$$

and

$$\begin{aligned} \text{var}(\hat{Y}_{iks} - \hat{Y}_{ik's}) &= \text{var}\left(\frac{\hat{Z}_{iks}}{c_k} - \frac{\hat{Z}_{ik's}}{c_{k'}}\right) \\ &= \frac{\text{var}(\hat{Z}_{iks})}{c_k^2} + \frac{\text{var}(\hat{Z}_{ik's})}{c_{k'}^2} - \frac{2 \text{cov}(\hat{Z}_{iks}, \hat{Z}_{ik's})}{c_k c_{k'}} \end{aligned} \quad (16)$$

The variance and covariances of \hat{Z} , or quantities required to calculate them, should be available from the computer program used to do the analysis.

Only some differences are of interest since, from Figures 1 and 2, it is clear that AK0001 and BN0001 have greater Na concentrations all year and greater $\text{NO}_3 + \text{NO}_2$ concentrations early and late in the year than the three locations BA0011, BE0013, and BH0017. Comparisons within a group

of locations can then be based on the unweighted regressions fitted to data from the individual locations. The results of testing for one location being uniformly greater than the other within a period, for the three periods and both $\text{NO}_3 + \text{NO}_2$ and Na are shown in Table 5. The fitted models are shown in Figure 9. Sodium concentrations were uniformly higher at location AK0001 than at BN0001 during the early and late periods of the year between 1984 and 1986, while $\text{NO}_3 + \text{NO}_2$ concentrations were not different over the entire period in the early part of the year and the summer. There is evidence of AK0001 concentrations exceeding those of BN0001 late in the year, but the evidence is weaker than observed for Na.

These two locations can also be compared by fitting the yearly mean and seasonal models as done in the previous chapter (Figure 10) and again differences between estimated values $\hat{Y}_{iks} - \hat{Y}_{ik's}$ can be obtained.

Matched Sampling Dates, Na and $\text{NO}_3 + \text{NO}_2$ Concentrations

The differences between pairs of locations for the three upper stations BA0011, BE0013, and BH0017 are plotted against time of sampling for both Na and $\text{NO}_3 + \text{NO}_2$ in Figure 10. Concentrations between all pairs of locations were found to differ significantly over the years 1984 to 1986 with consistent conclusions from the t, Wilcoxon, and sign tests (Table 6). The confidence limits for the mean difference in concentration between a pair of locations (Table 6) show that for $\text{NO}_3 + \text{NO}_2$, BH0017H < BA0011 < BE0013, and for Na,

BA0011 < BE0013 < BH0017. The plots suggest that the difference is not constant over the year. Division of the years into the periods described earlier resulted in significant differences for some periods but not others (Table 7), a result observed in the comparison of the two lower locations.

Na and NO₃ + NO₂ Flux

The seasonality observed in the concentrations is no longer evident in the flux (Figures 4 and 5), and attempts to fit a yearly mean and seasonal cycle to the flux data were not successful. Linear models within a period (equation (2)) were fitted to the flux data of locations BN0001 and AK0001, but the slope in time was significantly different from zero at the level of 0.07 or less only for NO₃ + NO₂ in the early period and for NO₃ + NO₂ and Na at location AK0001 in the summer. Thus comparisons between locations are based on the difference between location means (Figure 11 and Table 8), except for these two cases where two regression lines and a mean and a regression line are compared. The conclusions, similar to those based on concentrations, are that Na flux at AK0001 exceeds that at BN0001 during the early and late periods, but not in the summer, and NO₃ + NO₂ flux do not differ at the two locations.

The locations BA0011, BE0013, and BH0017 were analyzed for differences in the flux at pairs of locations by the same methods as used for concentrations (Figure 12 and Table 8). The ranking of locations changes from BH0017 < BA0011 < BE0013 on the bases of NO₃ + NO₂ concentrations to BA0011 < BE0013 and BH0017 for NO₃ + NO₂ flux, but doesn't

change for Na. Only the differences between Na fluxes at BA0011 and BE0013 during the summer period (Figure 14) were analyzed, since this provides an example of a linear time trend in the difference. The Na flux at BA0011 exceeds that at BE0013 over the summer period but the difference decreases over the period. The regression of the differences d_{ij} on month j is

$$d_{ij} = -120 + 7j$$

where d_{ij} is the difference between BA0011 and BE0013 in month j of year i and the largest difference was excluded from the analysis. The reason for taking differences between locations at matched dates is to remove the effect of season. If the seasonal effect is not the same at the two locations, differences will still retain some seasonality and the analysis will need to account for this. A marked linear trend in $\text{NO}_3 + \text{NO}_2$ differences in the early period, when location BH0017 is involved, is also evident (Figure 15).

DISCUSSION

The nonconstancy of the concentration of water quality variables within the year due to the hydrological regime means that to compare the concentrations at two locations, the locations must be at the same point on the hydrological regime. When matched sampling dates provide this correspondence, the data at the matched dates can be used directly and this has been done by taking differences between concentrations or between fluxes at matched dates for locations BA0011, BE0013 and BH0017. However, matched dates may

not always be available and then estimates of the concentration or flux at matched dates can be used instead of observed concentrations or fluxes, as was done for BN0001 and AK0001. Simple linear regression or a mean level were used to estimate the concentration or flux within a period of the year and fits of the seasonal cycle and yearly mean were also shown to be suitable for concentrations. Various smoothing procedures are another alternative.

Of the two water quality parameters analyzed here, Na reflects the urban and agricultural inputs to the rivers. Both Na concentrations and fluxes at BN0001 and AK0001 were greater than those at the three upper stations (with one exception, Figure 4). There was a consistent difference between BN0001 and AK0001 for Na concentration and flux, with AK0001 exceeding BN0001 early and late in the year, but no difference over the summer period, the period which was defined by the flow peak ascribed to the melting of snow in the mountains. On the basis of Na flux over an entire year, this data (1984 to 1986) indicated an increase in Na flux for locations further from the headwater location.

ACKNOWLEDGEMENTS

The authors would like to thank L.C. Keeler who performed many of the extensive computations required to complete this report.

REFERENCES

- Draper, N.R. and Smith, H. 1981. Applied Regression Analysis. Wiley, New York, 709 pp.
- El-Shaarawi, A.H., Esterby, S.R., Warry, N.D., and Kuntz, K.W. 1985. Evidence of contaminant loading to Lake Ontario from the Niagara River. Can. J. Fish. Aquat. Sci., 42, 1278-1289.
- Lehmann, E.L. and D'Abbrera, H.J.M. 1975. Nonparametrics. Holden-Day, Oakland, California, 457 pp.
- Snedecor, G.W. and Cochran, W.G. 1980. Statistical Methods. The Iowa State Univeristy Press, Ames, Iowa, 507 pp.
- Tsutakawa, R.K. and Hewett, J.E. 1978. Comparison of two regression lines over a finite interval. Biometrics, 34, 391-398.

FIGURE CAPTIONS

- Figure 1. Na concentrations, 1984 to 1986, for (a) all five locations, and (b) the three upper locations.
- Figure 2. $\text{NO}_3 + \text{NO}_2$ concentrations, 1984 to 1986, for (a) all five locations, and (b) the three upper locations.
- Figure 3. Discharge on the dates of water quality sampling, 1984 to 1986, for (a) all five locations, and (b) the three upper locations.
- Figure 4. Na flux, 1984 to 1986, for (a) all five locations, and (b) the three upper locations.
- Figure 5. $\text{NO}_3 + \text{NO}_2$ flux, 1984 to 1986, for (a) all five locations, and (b) the three upper locations.
- Figure 6. Na concentrations versus flow by period, 1984 to 1986.
- Figure 7. $\text{NO}_3 + \text{NO}_2$ concentrations versus flow by period, 1984 to 1986.
- Figure 8. Models fitted to Na and $\text{NO}_3 + \text{NO}_2$ concentrations by period over 1984 to 1986 shown together, with lines from different periods joined and lines for different locations distinguished. Data points for all locations are plotted with one symbol.

- Figure 9. Models used to test for concentrations at one location being greater than concentrations at the other over a period, for the three periods and 1984 to 1986 data. Observations and fitted lines are plotted against time as a fraction of the year.
- Figure 10. Yearly means and seasonal components fitted to Na and $\text{NO}_3 + \text{NO}_2$ concentrations, 1984 to 1986.
- Figure 11. Difference in Na and $\text{NO}_3 + \text{NO}_2$ concentrations at matched dates for the locations BA0011, BE0013, and BH0017.
- Figure 12. Models used to test for fluxes at one location being greater than those at the other location over a period, 1984 to 1986 data. Observations and fitted lines are plotted against time as a fraction of the year.
- Figure 13. Differences in Na and $\text{NO}_3 + \text{NO}_2$ fluxes at matched dates for locations BA0011, BE0013, and BH0017.
- Figure 14. Differences in Na flux at matched dates plotted against time as a fraction of the year, by period within the year, for 1984 to 1986 data at locations BA0011, BE0013, and BH0017.
- Figure 15. Differences in $\text{NO}_3 + \text{NO}_2$ flux at matched dates plotted against time as a fraction of the year, by period within the year, for 1984 to 1986 data at locations BA0011, BE0013, and BH0017.

TABLE 1: Division of the year into periods, based on discharge events, shown as first and/or last sampling dates for 1984 to 1986.

Station	First or Last Date of Sampling in Period Given as Month, Day					
	Prairie Melt		Mountain Melt Peak		Irrigation Return Flow	
	Last Date in Period	First Date in Period	Last Date in Period	First Date in Period	Last Date in Period	First Date in Period
Station	1984 1985 1986	1984 1985 1986	1984 1985 1986	1984 1985 1986	1984 1985 1986	1984 1985 1986
BA0011	0508 0402 0513	0613 0522 0603	0911 0813 0909	1016 0910 1015		
BE0013	0508 0403 0513	0614 0523 0604	0912 0813 0910	1017 0911 1015		
BH0017	0509 0403 0415	0611 0523 0521	0912 0814 0910	1017 0911 1016		
BN0001	0514 0433 0505	0618 0521 0616	0917 0821 0922	1015 0923 1020		
AK0001	0515 0416 0429	0529 0528 0527	0918 0801 0922	1023 0924 1028		

TABLE 2: Number of sampling dates per period for each station and year.

Location	Number of Sampling Dates														Location Total
	Early Period					Summer					Late Period				
	1984	1985	1986	Total		1984	1985	1986	Total		1984	1985	1986	Total	
BA0011	5	4	5	14	4	4	4	4	12	3	4	3	10	36	
BE0013	5	4	5	14	4	4	4	4	12	3	4	3	10	36	
BH0017	5	4	4	13	4	4	4	5	13	3	4	3	10	36	
BN0001	5	4	5	14	4	4	4	4	12	3	4	2	9	35	
AK0001	8	6	4	18	7	5	5	5	17	3	4	3	10	45	
TOTAL	28	22	23	73	23	21	22	22	66	15	20	14	49	188	

TABLE 3: Regression analysis based on model (8) for NO₃ + NO₂ and Na concentrations, 1984 to 1986.

Parameter	Year(s)	Early Period					Summer					Late Period				
		100			Residual		100			Residual		100			Residual	
		Model	R ²	p	df	SS	Model	R ²	p	df	SS	Model	R ²	p	df	SS
Na	1984	ℓ_k, tk	99.4	*	18	0.033	ℓ_k, tk	93.2	*	13	0.297	ℓ_k, tk	99.6	0.0002	4	0.016
	1985	ℓ_k, tk	98.1	*	10	0.051	ℓ_k	87.1	*	18	0.564	ℓ_k, tk	95.4	*	10	0.140
		ℓ_k, t	97.8	*	14	0.059	ℓ_k, tk	96.1	*	11	0.143	ℓ_k	93.6	*	15	0.194
	1986	ℓ_k, tk	94.7	*	13	0.177	ℓ_k, tk	85.4	0.0008	12	0.552	ℓ_k, tk	98.2	0.004	4	0.049
		ℓ_k	90.6	*	18	0.31	ℓ_k	73.2	0.0001	17	1.009	ℓ_k	97.4	*	9	0.069
1984-1986		ℓ_k, tk	95.1	*	61	0.55	ℓ_k, tk	85.7	*	56	1.69	ℓ_k, tk	92.3	*	38	0.756
												ℓ_k	91.7	*	43	0.821
NO ₃ + NO ₂	1984	ℓ_k, tk	84.0	*	17	0.00241	ℓ_k, tk	62.2	0.08	13	0.000274	ℓ_k, tk	95.5	0.007	5	0.000381
							ℓ_k	57.1	0.003	18	0.000311					
	1985	ℓ_k, tk	97.7	*	12	0.00052	ℓ_k, tk	82.1	0.005	11	0.000238	ℓ_k, tk	95.5	0.0002	10	0.00088
												ℓ_k, t	88.6	*	14	0.00134
	1986	ℓ_k, tk	90.4	*	13	0.00131	ℓ_k, tk	50.6	0.30	12	0.001482	ℓ_k, tk	98.2	0.004	4	0.000061
							t	18.6	0.045	20	0.002443	ℓ_k, t	90.4	0.0007	8	0.000320
	1984-1986	ℓ_k, tk	84.3	*	62	0.00842	ℓ_k, tk	28.1	0.02	56	0.003643	ℓ_k, tk	80.1	*	39	0.00474
							ℓ_k, t	24.8	0.004	60	0.003806					

p is the significance probability for the F test of the regression parameters; df the degrees of freedom; SS the sum of squares.

ℓ_k indicates different means were fitted for each location; t_k different slopes in time for each location; t one slope in time for all locations.

The sets of weights for BAO011, BE0013, BH0017, BN0001, AK0001 were (1.00, 0.55, 0.64, 0.03, 0.04) and (1.00, 0.42, 0.25, 0.05, 0.03) for NO₃ + NO₂ and Na, respectively.

TABLE 4: Tests for different means and time slopes in each year for NO₃ + NO₂ and Na, 1984 to 1986

Parameter	Early Period				Summer				Late Period			
	Residual		F		Residual		F		Residual		F	
	df	SS	df	(p)	df	SS	df	(p)	df	SS	df	(p)
Na	Separate	41 0.261	1.96	(>0.10)	36 0.992	1.27	18	0.205	2.42	(0.03)		
	Together	61 0.55			56 1.69		38	0.756				
	Difference	20 0.249			20 0.698		20	0.551				
NO ₃ + NO ₂	Separate	42 0.00424	2.07	(0.025)	36 0.001994	1.49	19	0.001322	2.46	(0.05)		
	Together	62 0.00842			56 0.003643		39	0.00474				
	Difference	20 0.00418			20 0.001649		20	0.003418				

p is the significance probability for the ratio of mean squares denoted by difference and separate in the table; df the degrees of freedom; and SS the sum of squares.

TABLE 5: Test of the hypotheses that the concentration at BN0001 equals that of AK0001 over the entire period based on 1984 to 1986 data.

Period	t_l	t_h	Na				NO ₃ + NO ₂			
			W_l	W_h	T(df)	p	W_l	W_h	T(df)	p
Early	13	136	-4.14	-6.18	2.99 (27)	0.006	0.11	0.34	0.56 (28)	0.29
Summer ^a	141	265	1.36	-12.04			-0.05	0.27		
Late ^b	266	350	-5.74		2.61 (17)	0.01	0.42	0.37	1.38 (23)	0.09

W = estimated concentration at BN0001 minus estimated concentration at AK0001.

t_l, t_h are the earliest and latest days in the period. p is the significance probability for the statistic, T , which is defined by equation (7), and df indicates degrees of freedom.

^a Lines cross hence one line is not uniformly above the other.

^b Difference in mean levels; T then becomes usual t statistic.

TABLE 6: Tests for a different in concentration between locations and estimation of the average difference, for years 1984 to 1986

Result	Sodium				Nitrate + Nitrite			
	BA0011 -BE0013	BE0013 -BH0017	BA0011 -BH0017	BA0011 -BE0013	BE0013 -BH0017	BA0011 -BE0013	BE0013 -BH0017	BA0011 -BH0017
Number of pairs	34	33	33	35	34	33		
Significance probability t-test	<0.0001	0.001	<0.0001	0.0009	<0.0001	0.012		
Wilcoxon signed ranks	<0.0001	0.0004	<0.0001	0.002	<0.0001	0.014		
Sign test	<0.0001	<0.0001	<0.0001	0.002	<0.0001	0.012		
Confidence limits, $t^{(1)}$	-0.89,	-0.50,	-1.27,	-0.025,	0.014,	0.001,		
	-0.64	-0.09	-0.86	-0.005	0.042	-0.020		

(1) These are based on the Bonferroni t statistics and thus with probability 0.95, the true difference will lie within the given interval simultaneously for all three differences calculated for one water quality parameter.

TABLE 7: Wilcoxon signed-rank test for the difference between locations by period, 1984 to 1986

Period	Significance Probability (Number of Pairs)					
	Sodium			Nitrate + Nitrite		
	BA0011 -BE0013	BE0013 -BH0017	BA0011 -BH0017	BA0011 -BE0013	BE0013 -BH0017	BA0011 -BH0017
Early	0.002 (13)	0.064 (13)	0.003 (12)	0.18 (13)	0.009 (13)	0.08 (12)
Summer	0.002 (12)	0.002 (12)	0.65 (12)	0.16 (12)	0.10 (12)	0.25 (12)
Late	0.009 (9)	0.01 (8)	0.009 (9)	0.007 (10)	0.001 (9)	0.23 (9)

TABLE 8: Test of the hypotheses that the flux at BN0001 exceeds that of AK0001 over the entire period based on 1984 to 1986 data.

Period	t_2	t_h	Na				NO ₃ + NO ₂			
			W_ℓ	W_h	T(df)	p	W_ℓ	W_h	T(df)	p
Early	13	136	-825 ^b		4.56 (29)	<0.0005	-18.7	6.4 ^a		
Summer	141	265	-1678	-165	0.36 (26)	0.64	-19.3	29.3 ^a		
Late	266	350	-872 ^b		3.11 (17)	0.004	6.8		0.53 (17)	0.69

See footnotes Tables 5.

TABLE 9: Tests for the difference in flux between locations and estimation of the average difference for years 1984 to 1986

Result	Sodium				Nitrate + Nitrite			
	BA0011 -BE0013	BE0013 -BH0017	BA0011 -BH0017	BA0011 -BE0013	BA0011 -BE0013	BE0013 -BH0017	BA0011 -BH0017	BA0011 -BH0017
Number of pairs	34	33	33	35	34	33	33	33
Significance probability t-test	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.08	0.0001	0.0001	0.0001
Wilcoxon signed ranks	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.05	< 0.0001	< 0.0001	< 0.0001
Confidence limits, $t^{(1)}$	-61.89, -34.15	-91.42, -54.26	-149.44, -91.29	-3.75, -1.44	-2.47, 0.43	-5.97, -1.58	-5.97, -1.58	-5.97, -1.58

See footnote Table 6.

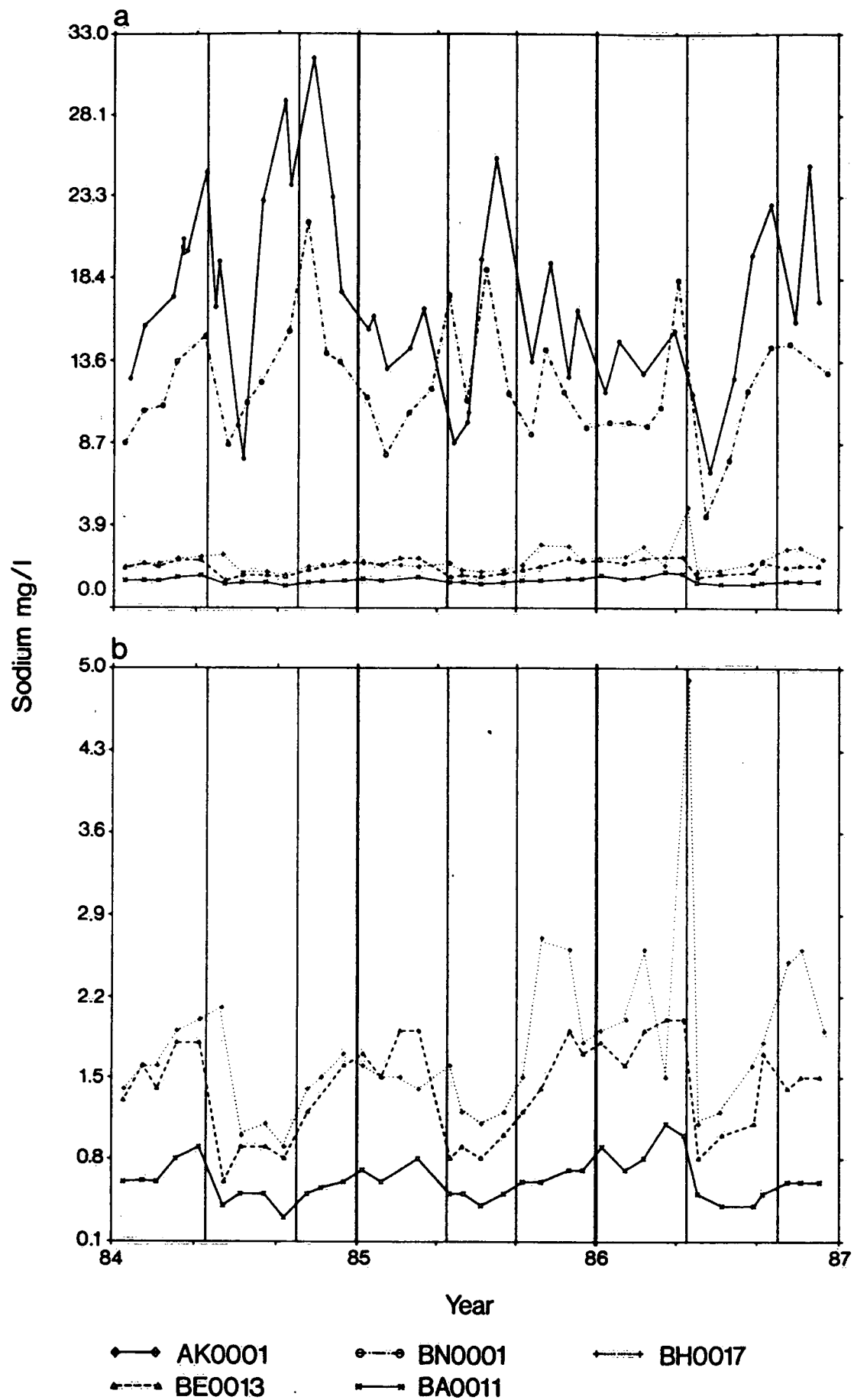


FIGURE 1

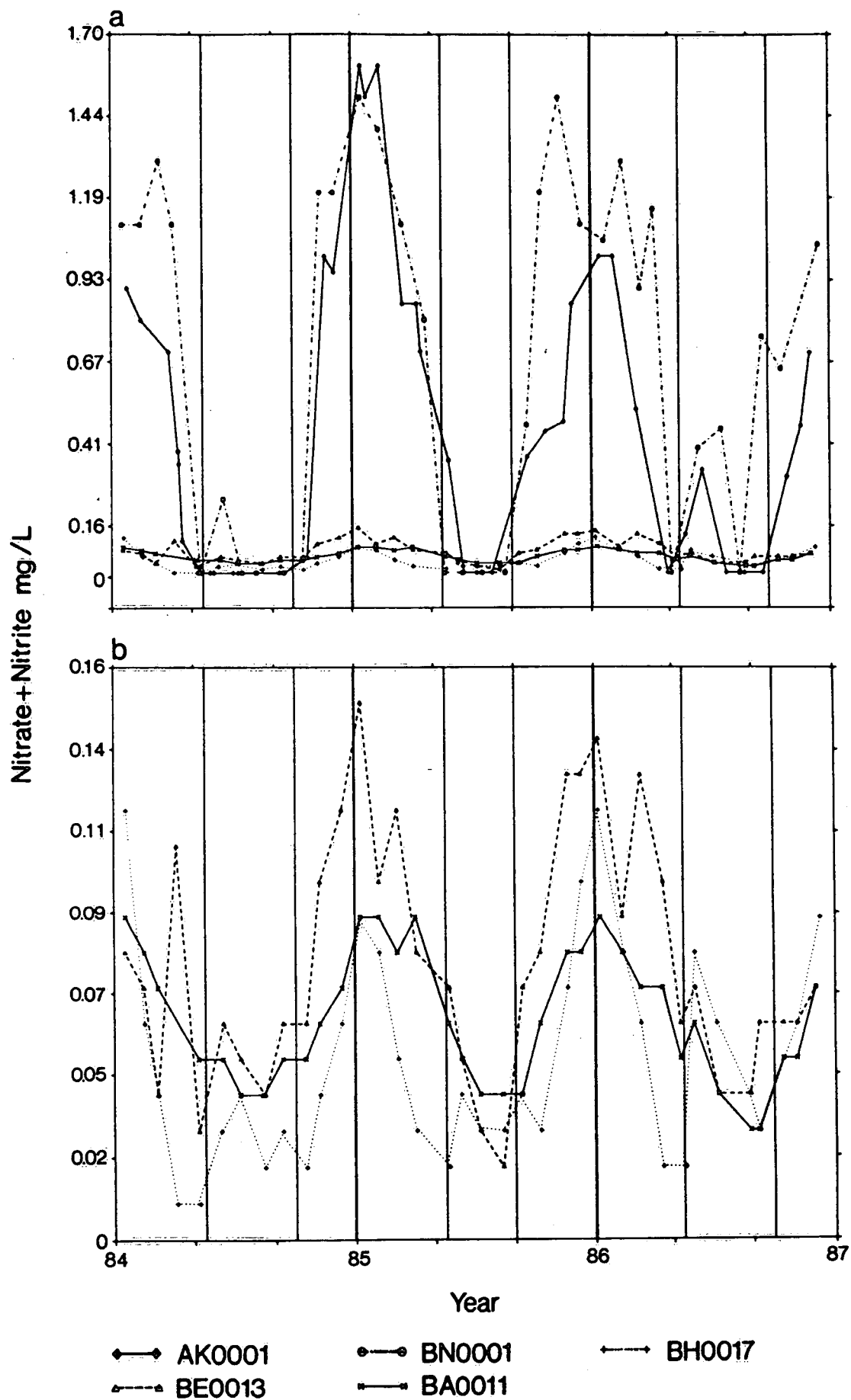


FIGURE 2

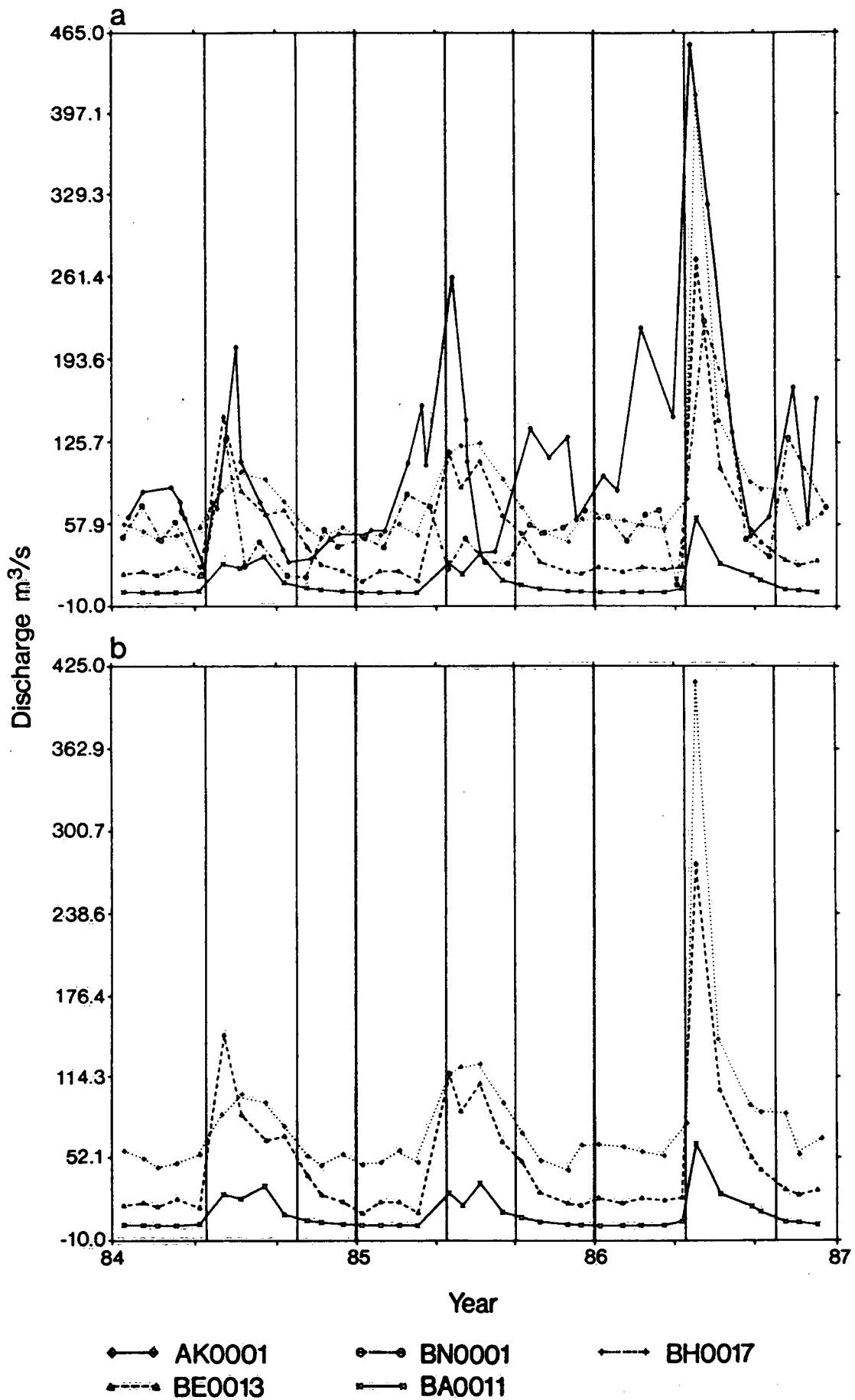


FIGURE 3

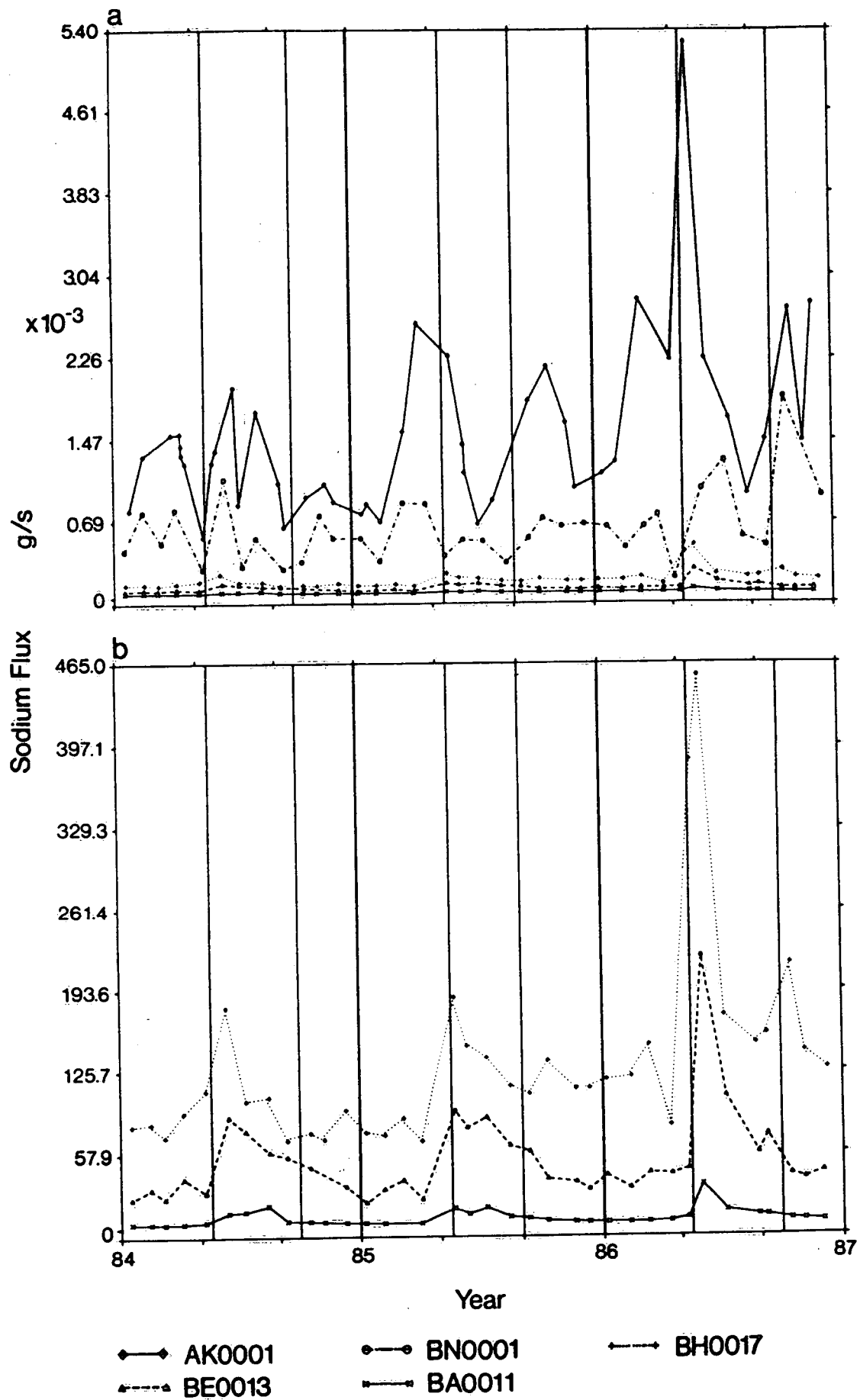


FIGURE 4

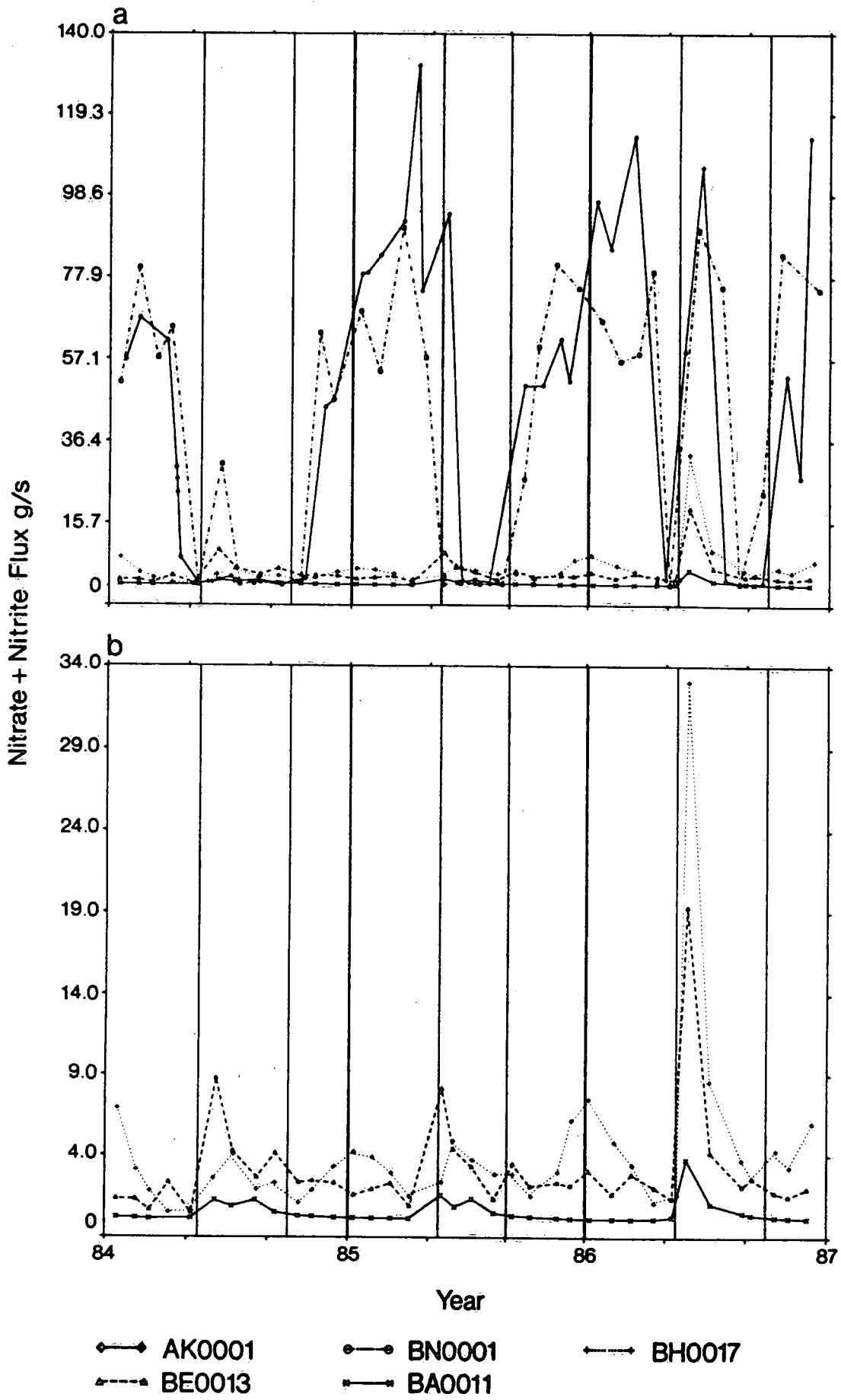


FIGURE 5

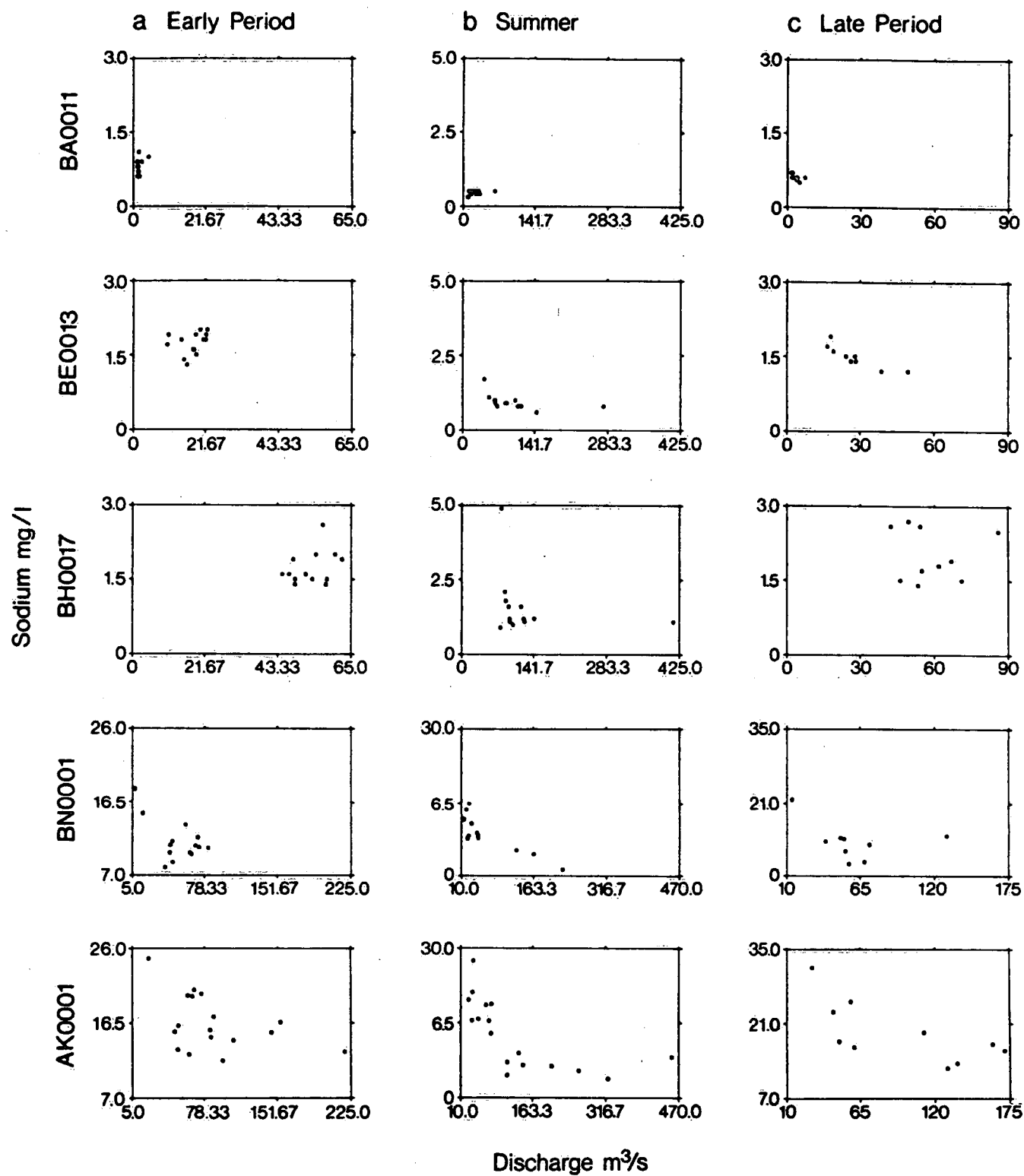


FIGURE 6

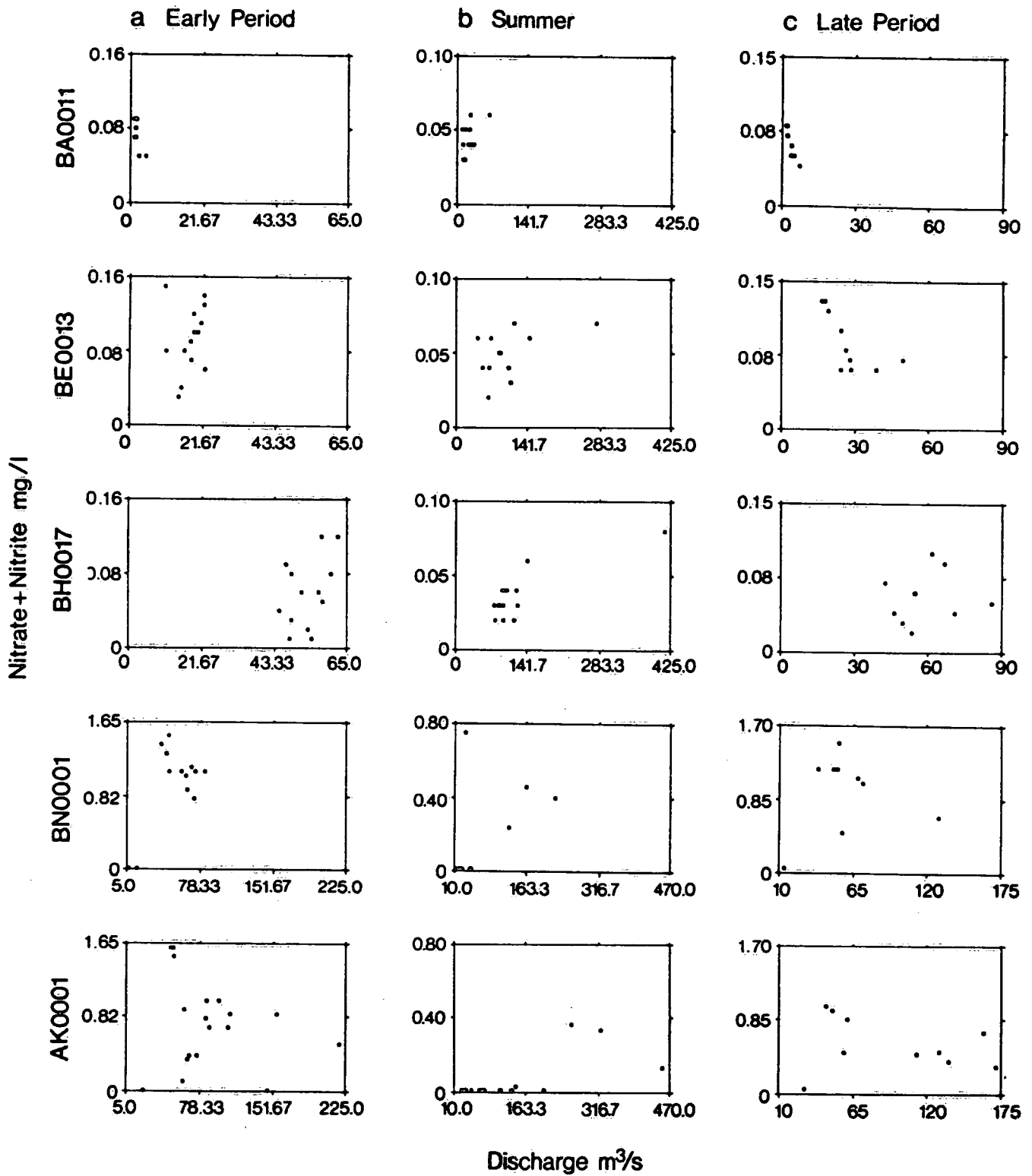


FIGURE 7

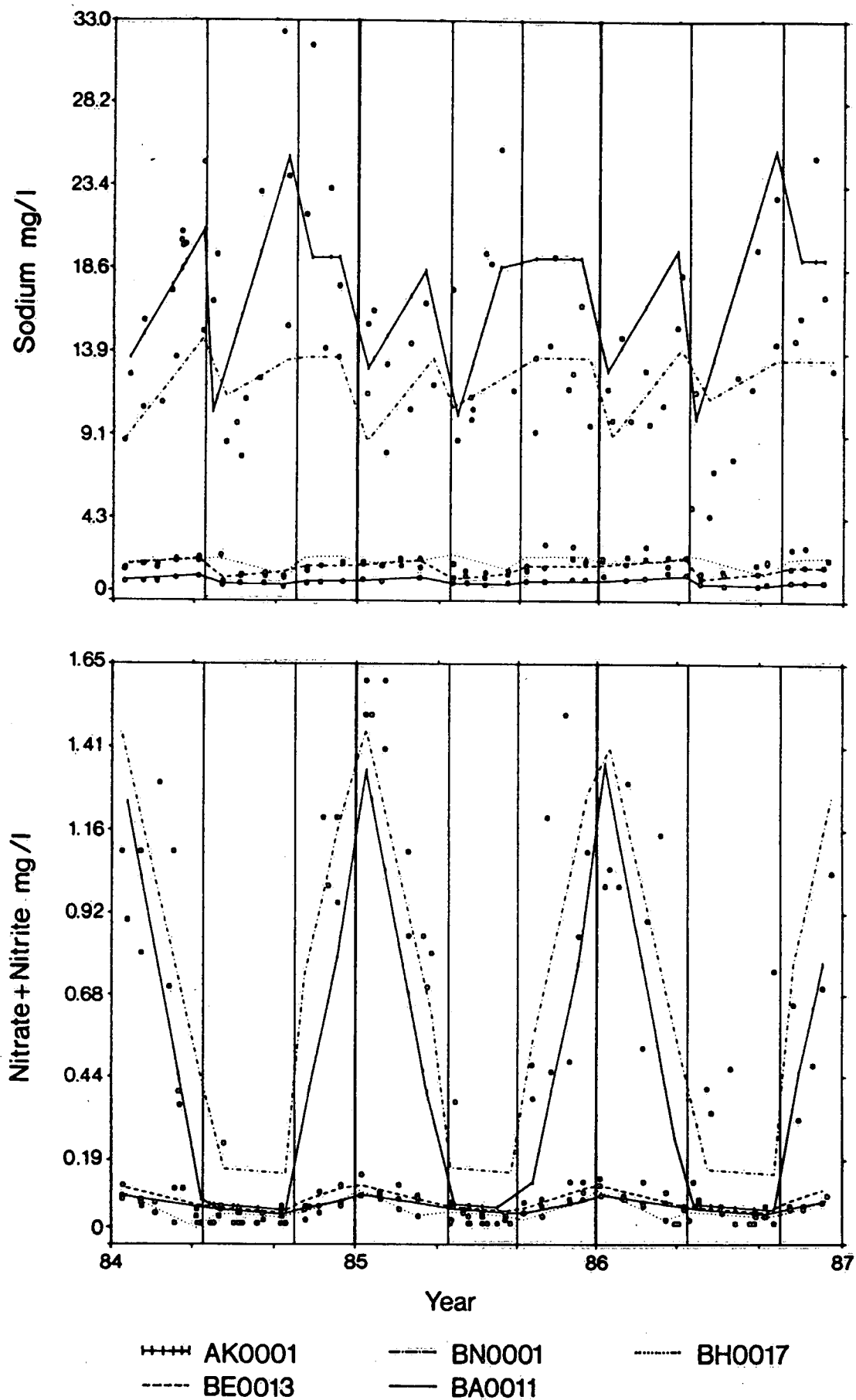


FIGURE 8

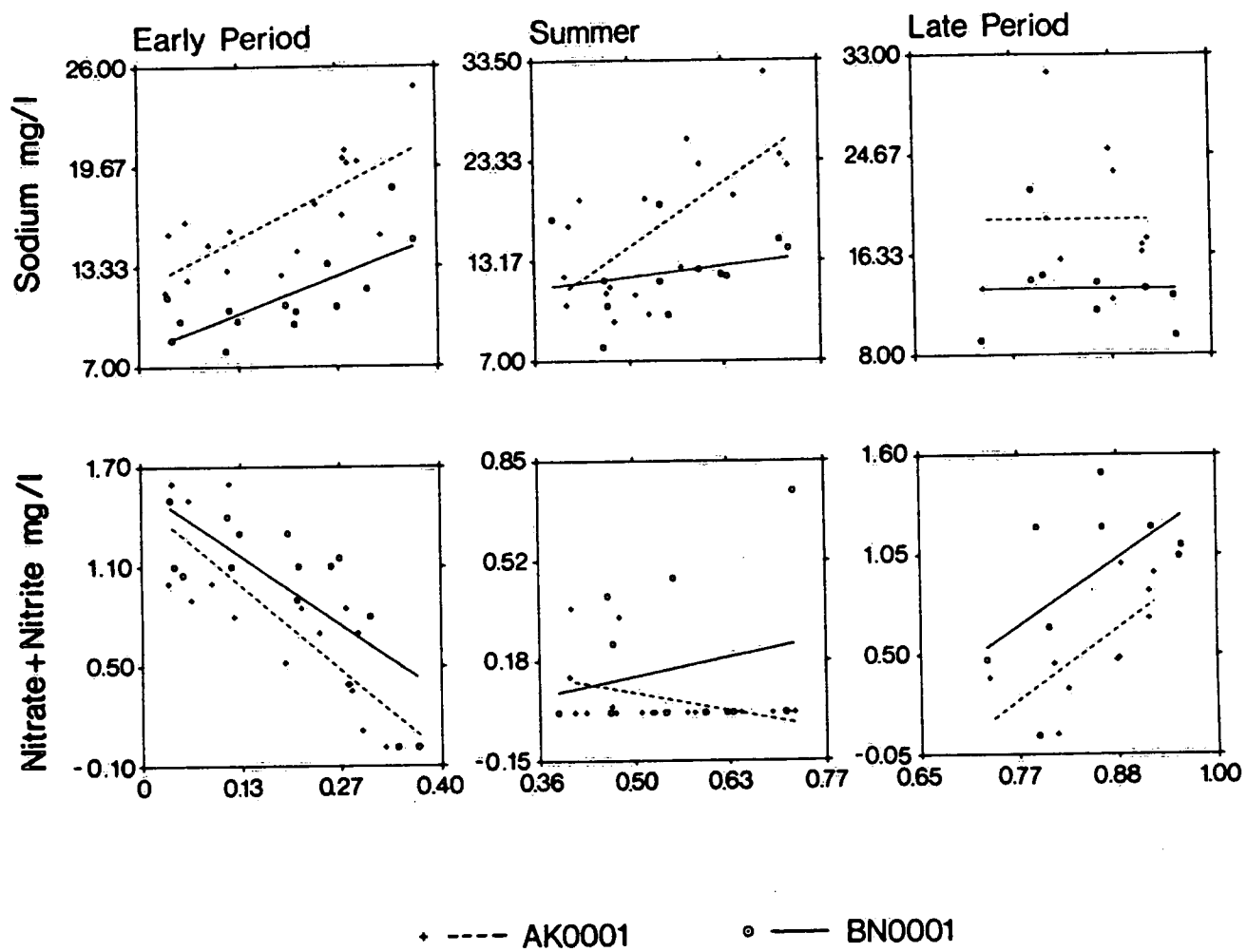


FIGURE 9

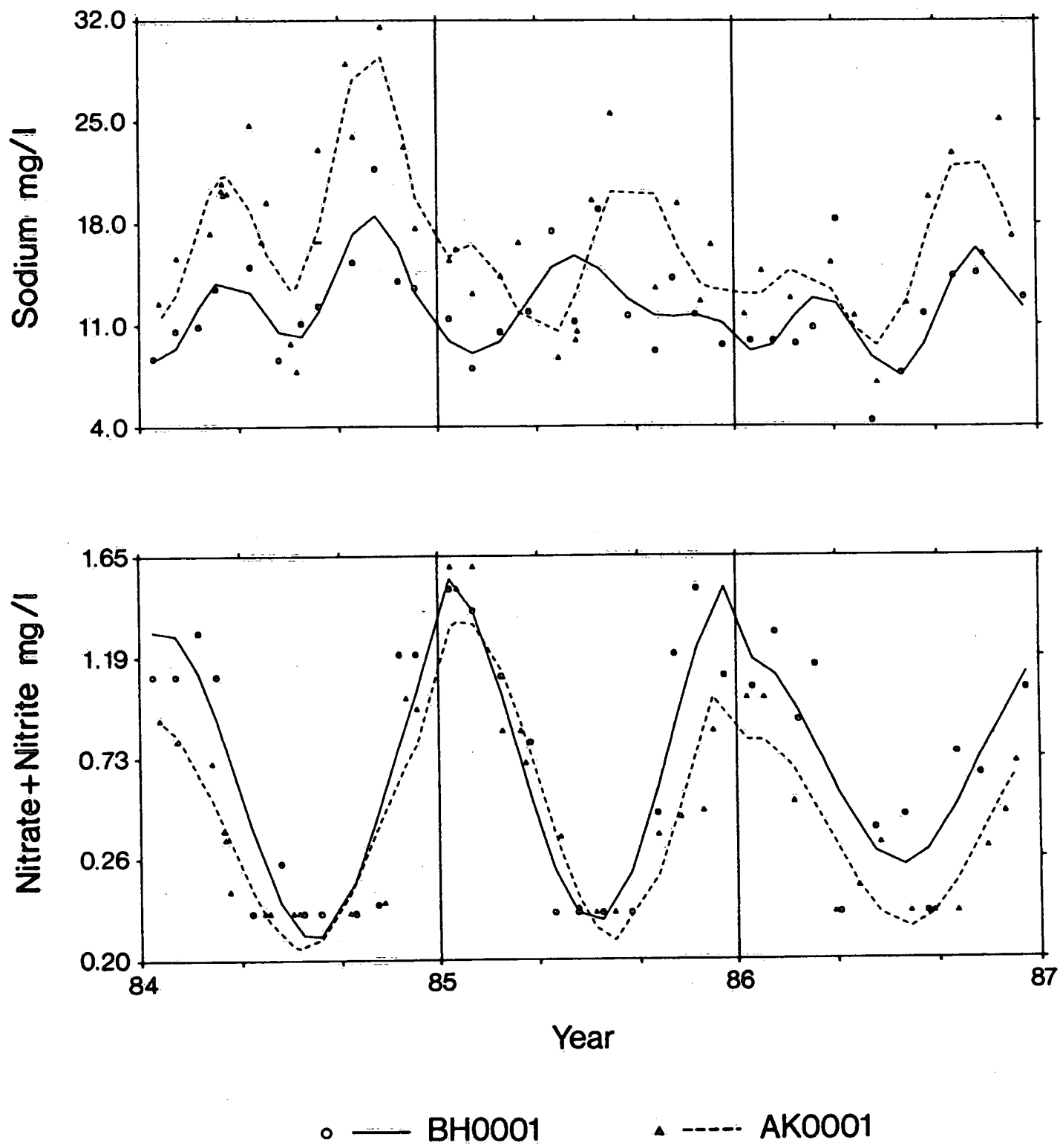


FIGURE 10

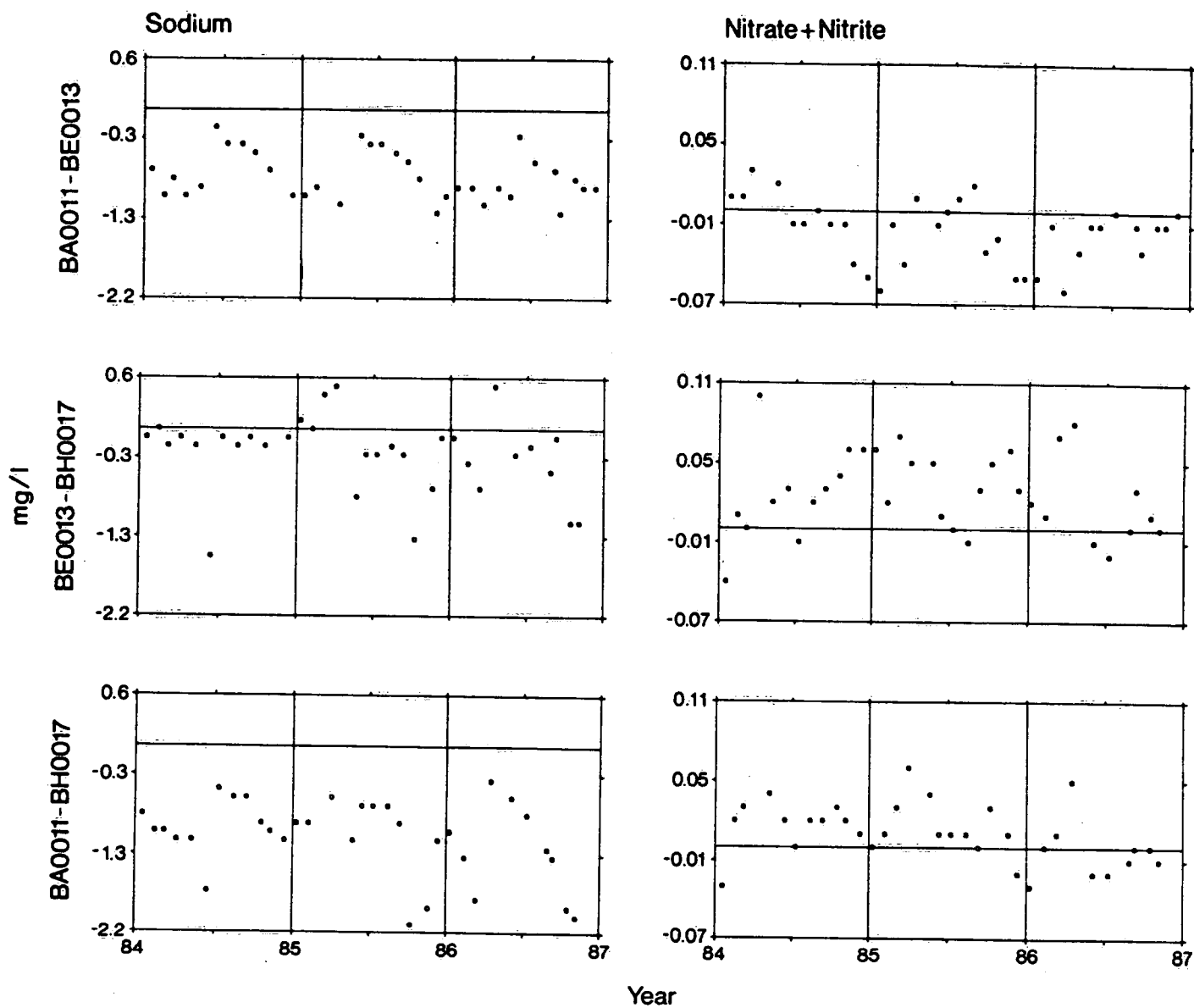


FIGURE 11

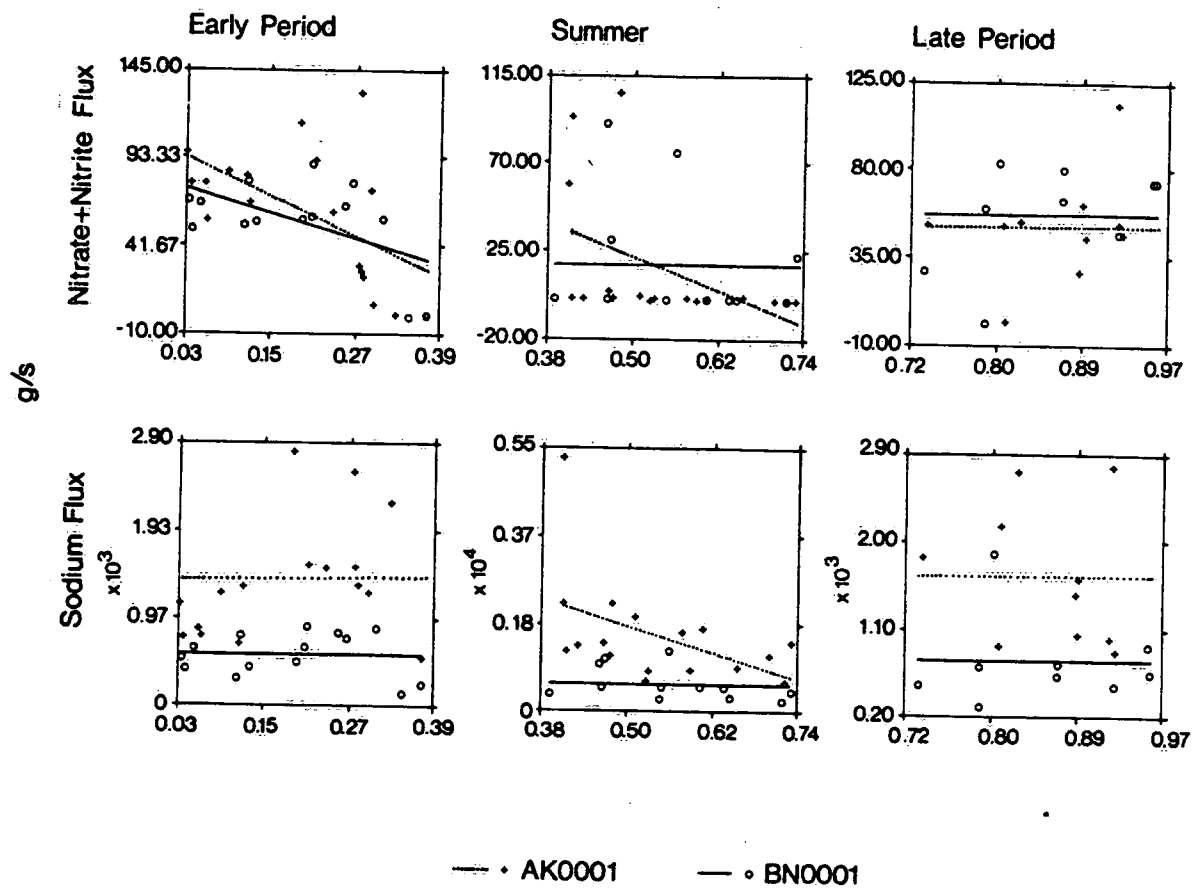


FIGURE 12

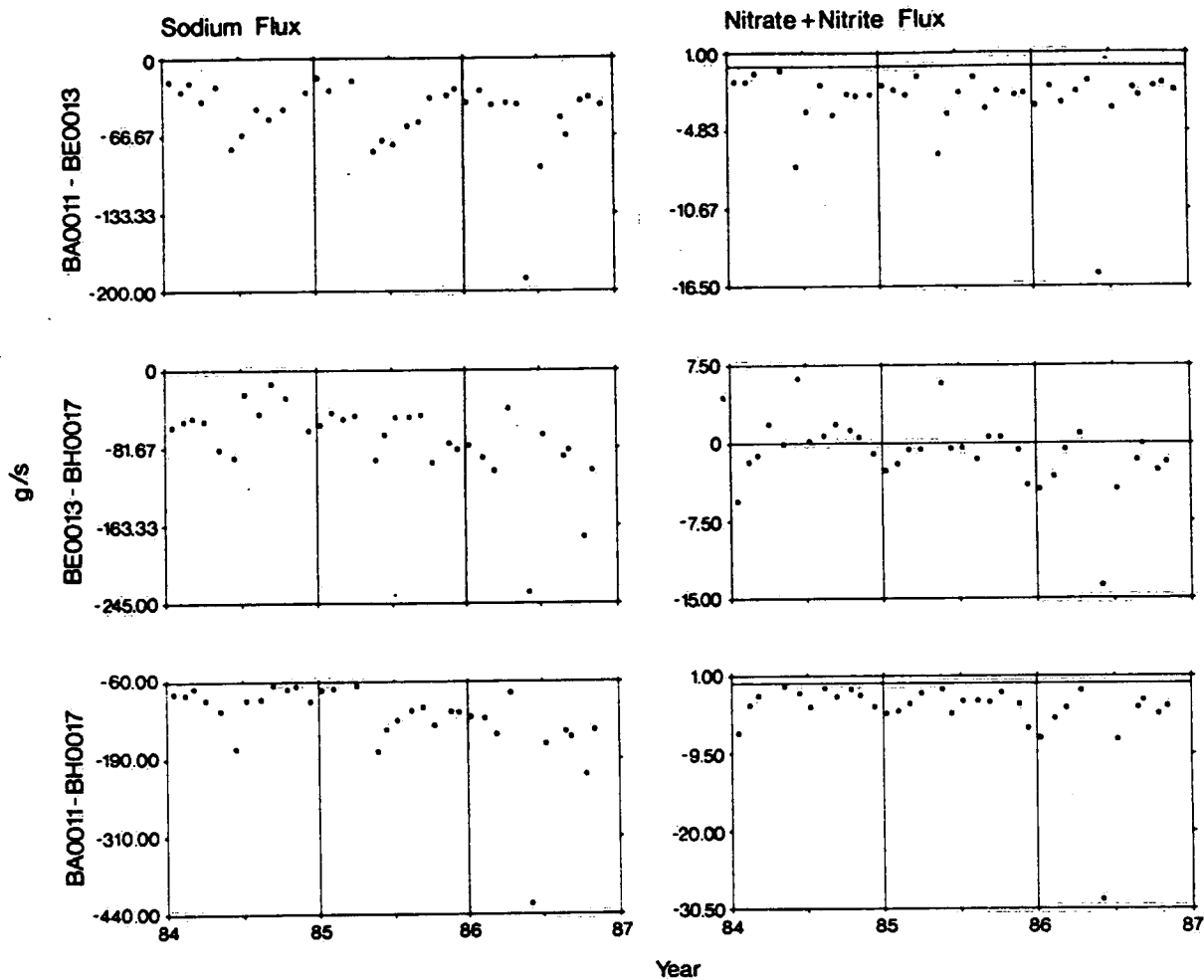


FIGURE 13

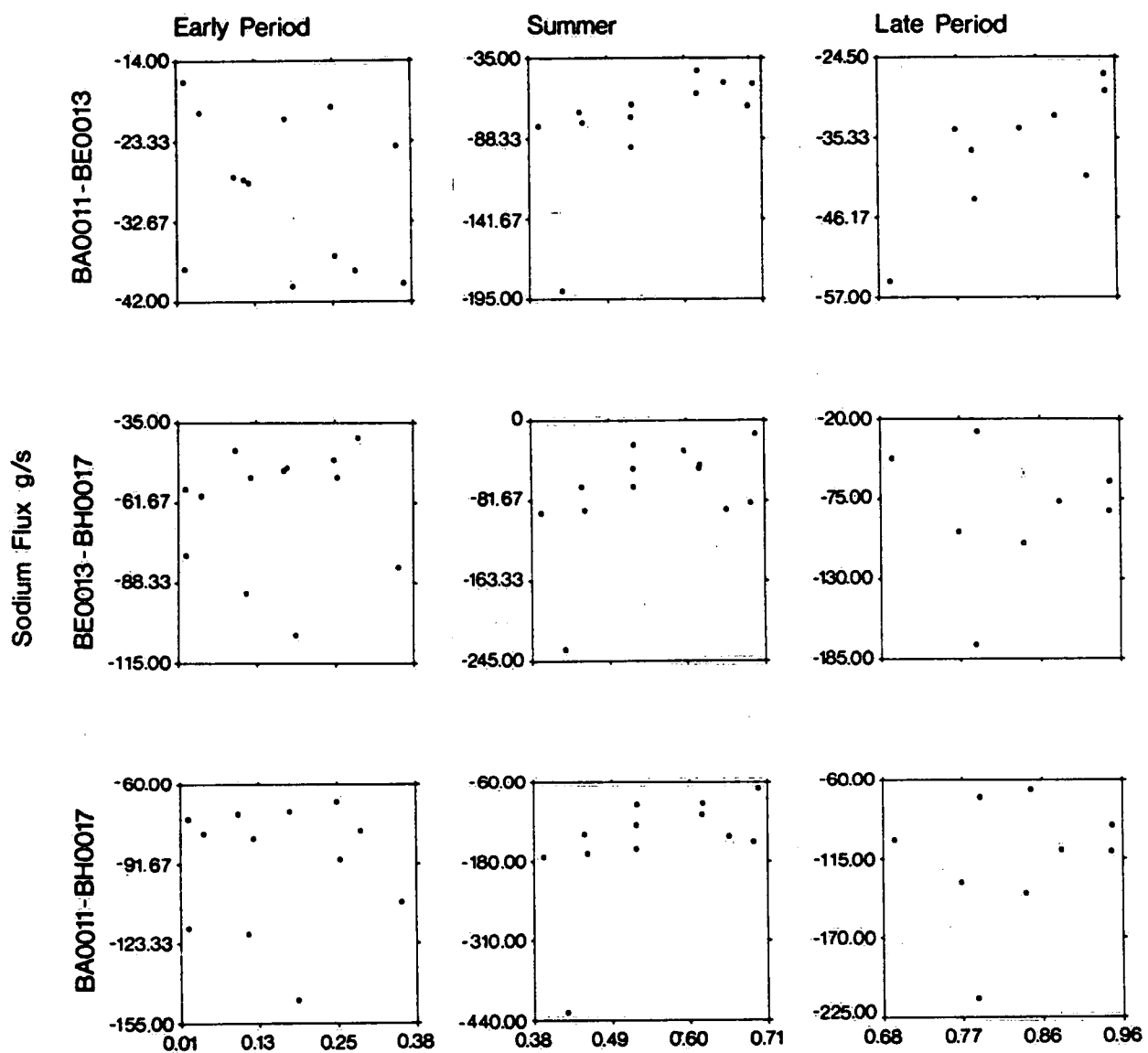


FIGURE 14

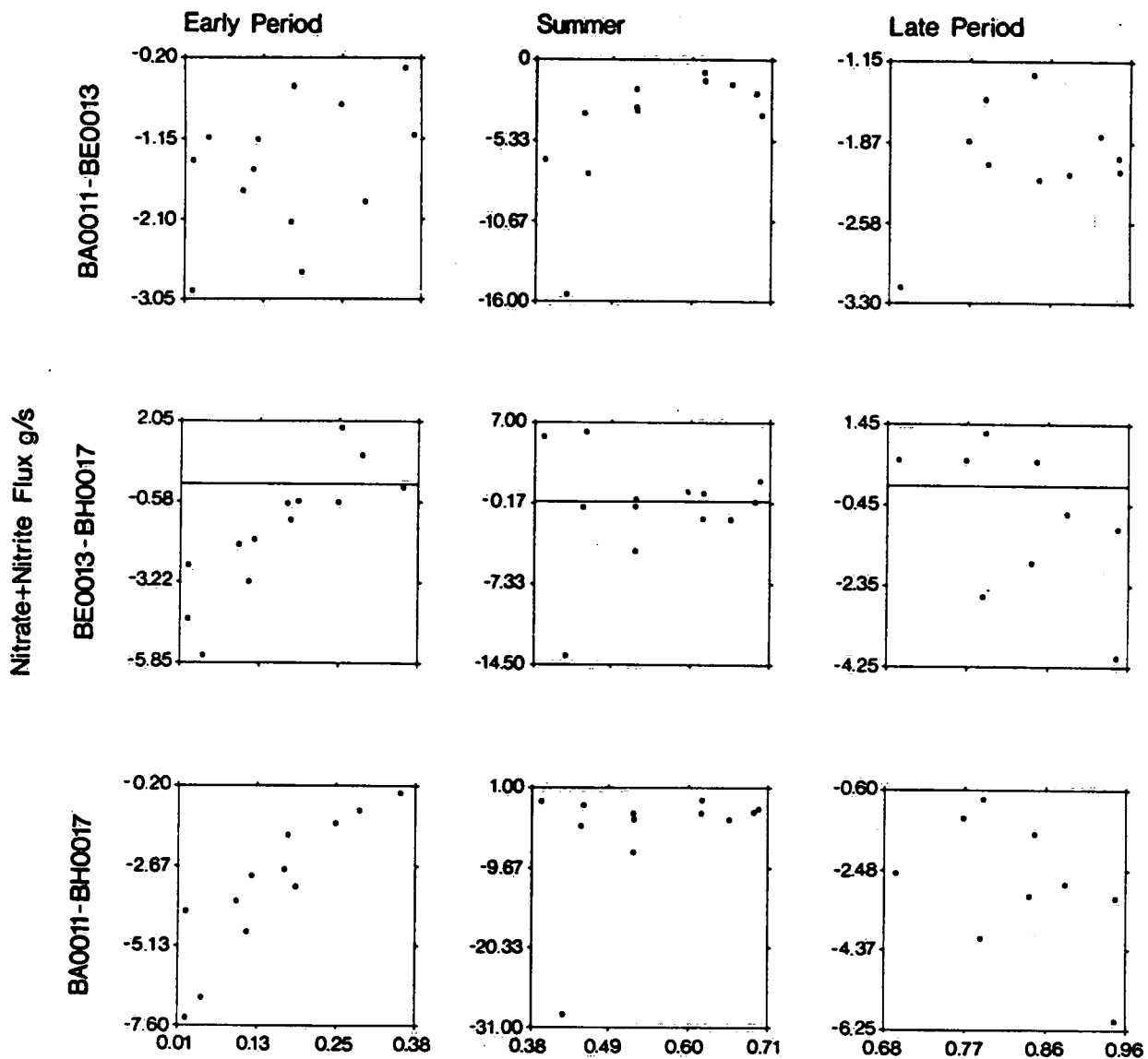
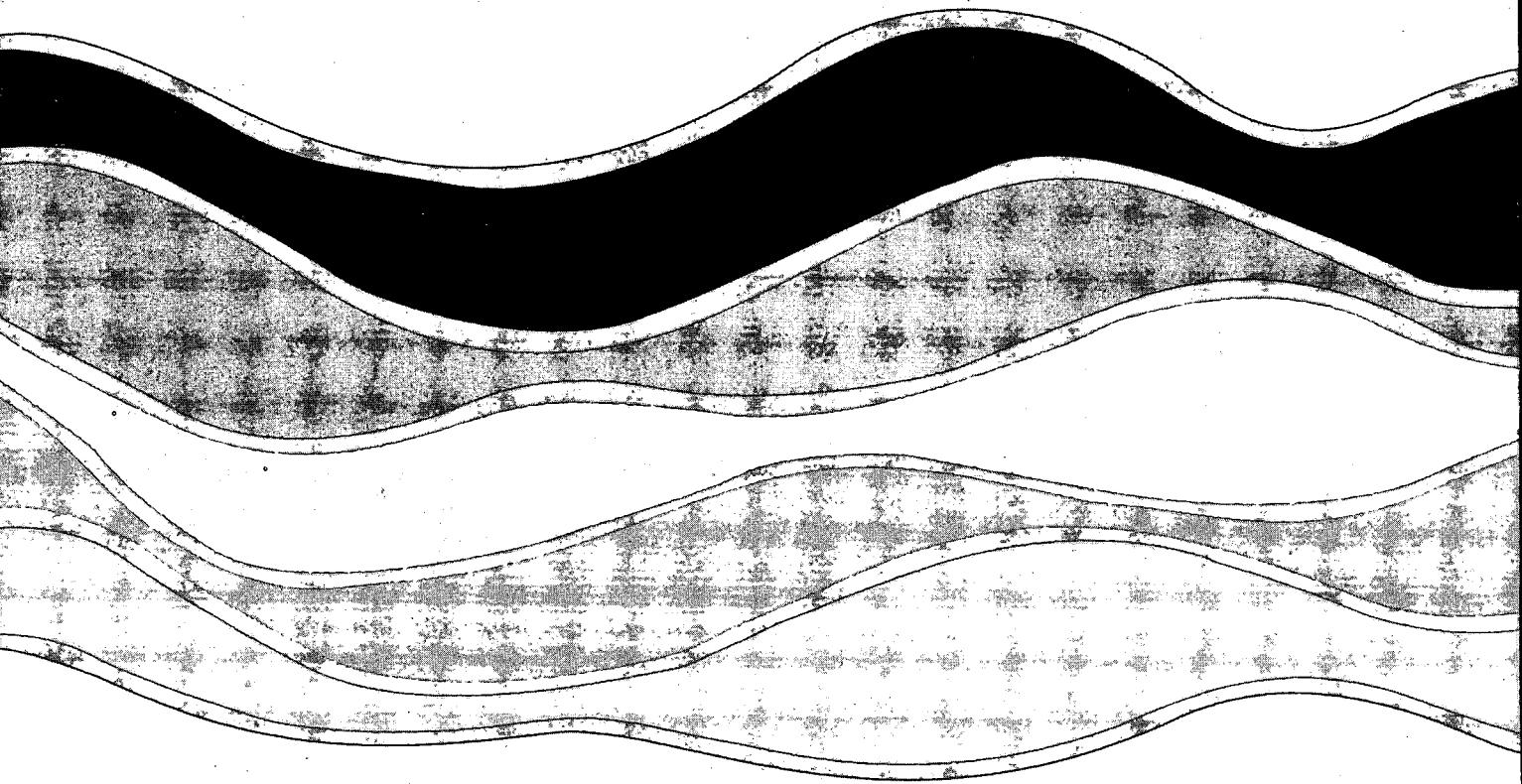


FIGURE 15

Environment Canada Library, Burlington



3 9055 1017 0639 7



NATIONAL WATER RESEARCH INSTITUTE
P.O. BOX 5050, BURLINGTON, ONTARIO L7R 4A6



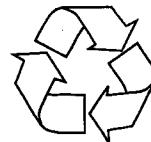
Environment
Canada

Environnement
Canada

Canada

INSTITUT NATIONAL DE RECHERCHE SUR LES EAUX
C.P. 5050, BURLINGTON (ONTARIO) L7R 4A6

Think Recycling!



Pensez à Recycling!