MULTI-SPECIES MODEL OF A DECIDUOUS FOREST[1]

Benjamin B. Stout

Cook College, Rutgers University, New Brunswick, New Jersey 08903

Jean Marc Deschenes

Research Station, Ste. Foy, Province of Quebec, Canada

AND

Lewis F. Ohmann

North Central Forest Experiment Station, USFS, St. Paul, Minnesota 55101

Abstract. A multi-species model of a deciduous forest is constructed that includes four species and three generic groups, based on one data set. Model equation coefficients are obtained using least squares, two-stage and three-stage least squares, and ridge regression statistical procedures. The resulting model equations are applied to a second data set and comparisons made of the residual mean squares from both data sets. Six of seven comparisons indicate essential equality. The model is examined in the light of the existing forest and conclusions regarding forest stand trends are drawn. On the basis of the interrelations in the model and supporting evidence in the literature it is concluded that there is not a unidirectional trend in composition, but rather the possibility of many trends. Which trend is actually followed will depend on the state of the environmental variables and the extent of disturbance.

Key words: Disturbance; forest model; least squares; unidirectional.

INTRODUCTION

Ecological studies in forest stands frequently are designed to examine the presence or absence, distribution, or productivity of a particular species in relation to environmental factors. Because of the underlying assumption upon which certain statistical procedures like multiple linear regression are based, consideration of the effect of other species on the study species is usually excluded. The purpose of this paper is twofold: (1) to test and apply a model not normally used in ecology wherein several dependent variables, in this case species, are considered simultaneously, and (2) to examine the results of the model in the light of existing knowledge regarding the forest stands of northern New Jersey.

AN ECONOMETRIC APPROACH

Our model is based on an econometric approach to such problems as supply and demand. Assume that two attributes of a particular commodity are of interest, supply (Y_1) and demand (Y_2). Also assume that other factors (X_j) are correlated with the supply and demand. We can write, for example,

(1)
$$Y_1 = f_1(Y_2, X_1, X_3, X_5)$$
$$Y_2 = f_2(Y_1, X_2, X_3, X_4, X_5)$$

The dependent variable in one equation appears as an "independent" variable in the other equation. The application of ordinary least-squares procedures to each of the equation in (1) violates the assump-

tion upon which the procedures are based and results in biased estimates of the coefficients. One assumption is that the "independent" variables are uncorrelated. Yet, in the example, X^3 and X^5 are used in both the Y^1 and Y^2 equations along with the respective Y's. Coefficients derived from such a system will be influenced in magnitude by an unspecified amount; they are biased.

Because of the seriousness of the bias problem econometricians have elaborated on least-squares procedures in an effort to get more realistic and less biased estimates of the coefficients in such a system of equations. One elaboration of least squares is to take predicted values of the dependent variables and use these instead of observed values to estimate coefficients; this is two-stage least squares (2SLS). A further elaboration is to take the array of differences between the observed and predicted values for all observations of the dependent variables (the residuals matrix) from the 2SLS analysis and divide that array into the variance-covariance matrix of "independent" variables used to predict the regression coefficients; this is three-stage least squares (3SLS) (Klein 1953, Friedman and Foote 1957). The 3SLS may be extended iteratively using each new residuals matrix generated each time a new set of coefficients is calculated (Furnival and Wilson 1971).

APPLICATION OF THE MODEL
TO A FOREST

The upland hardwood forest of northern New Jersey is a complex of several species growing on a

TABLE 1. Means and ranges of basal area per acre for the 24 stands sampled by Deschenes (1968) and the 24 stands sampled by Ohmann (1964). Prediction equations were developed from the Deschenes data and tested with the Ohmann data

Species	Deschenes		Ohmann	
	Mean	Range	Mean	Range
Acer saccharum	6.08	0.31 – 34.00	3.73	0.00 – 28.29
Fagus grandifolia	9.60	0.00 – 60.31	4.01	0.00 – 27.98
Quercus sp.	53.94	2.35 –102.96	62.02	10.49 –101.00
Fraxinus americana	11.14	0.00 – 58.44	3.41	0.00 – 27.91
Betula sp.	6.27	0.00 – 30.41	10.98	0.00 – 54.09
Carya sp.	4.61	0.00 – 15.69	4.81	0.00 – 14.57
Acer rubrum	6.76	0.00 – 51.88	5.09	0.66 – 15.95

complex of soil and topographic conditions. If we let some quantitative measure of the species be the dependent variables (Y_i) and the enviromental variables be the X_i we can write a system of equations comparable to (1) to be used as an equivalent model for forests:

$$Y_1 = f_1(Y_i; X_1, X_2, \ldots, X_j), \quad i \neq 1$$
$$Y_2 = f_2(Y_i; X_1, X_2, \ldots, X_j), \quad i \neq 2$$
(2)
$$Y_n = f_n(Y_i; X_1, X_2, \ldots, X_j), \quad i \neq n.$$

This model and the constitutent equations (2) express in a cryptic way the complex system of a forest that emerges if one discusses the intricate interrelationships that the experienced field ecologist perceives: each species is related to the other species and to the environment in complex but unspecified ways. Lotka (1924) proposed this concept in his discussion of rates of evolution.

METHODS

We used two sets of data, one of them, collected by Deschenes (1968), to calculate all of the equation coefficients. With those alone we could test only for statistical significance, the usual procedure. However, we chose to go one step further. The equations were applied to data collected by Ohmann (1964) to determine whether the coefficients were unique to the first data set or were applicable to the forest in general. Both data sets were collected in the area south of the Wisconsin terminal moraine in northern New Jersey.

The data set used to calculate equation coefficients had in it information on four species, *Acer saccharum*, *Fagus grandifolia*, *Fraxinus americana*, and *Acer rubrum*, and four generic groups, *Quercus* spp. (predominantly *Q. rubra*, *Q. alba*, and *Q. velutina*), *Betula* spp. (predominantly *B. lenta*), and *Carya* spp. Ohmann's data included these and a few others; the species or generic groups used in this study were common to both data sets.

Deschenes (1968) measured 18 soil and topo-graphic variables. Nine of those were used in this study: (1) Percent clay in the B horizon, (2) elevation in feet above mean sea level, (3) percent carbon in the A horizon, (4) carbon/nitrogen ratio in the A horizon, (5) the reciprocal of elevation, (6) the square root of carbon in the A horizon, (7) the reciprocal of exposure scaled (before taking the reciprocal) from 1 and the southwest to 7 on the north exposures with a value of 3.5 assigned to level plots, (8) nitrogen in the A horizon—percent by weight, (9) magnesium in the B horizon—mg/100 g. These nine were selected because of their apparent relationship with one or more of the species or generic groups on the basis of data plots and simple correlations.

Basal area in square feet per acre (BA/A) was used as the quantitative measure for each species or generic group $(1.0 \text{ ft}^2/A = 0.2296 \text{m}^2/\text{ha})$. Deschenes (1968) that a natural logarithm transformation of BA/A gave the best fit in his regression analyses. Because of this and to eliminate negative values we used $\text{Ln}(\text{BA}/A + 1)$ as the response variable. A summary of the data is given in Table 1.

Each species and generic group equation was developed through multiple stepwise least-squares regression analysis; all equations are linear additive. The response variable $(\text{Ln}(\text{BA}/A + 1)$ for each of the 7 species was regressed on the remaining 6 species, also $\text{LN}(\text{BA}/A + 1)$ and the 9 soil and topographic variables. Equations were accepted for use as model components if the multiple correlation coefficient, F ratio, and the SE of the coefficients were all significant at the 5% probability level.

The model equations from the regression analysis were further subjected to two-stage and three-stage least-squares (with iteration) procedures. We also examined the model equations using Ridge Regression modifications of standard least-squares regression procedures (Hoerl and Kennard 1970a, b). Because we knew that the basic assumption upon which least-squares regression is based had been violated, these other procedures were used to obtain other estimates of the coefficients. The several estimates of the coefficients for each of the 7 species or generic

TABLE 2. The seven model equations developed and accepted from Deschenes' data (1968). Also shown are comparable RMS from Ohmann's data (1964). The analysis that produced the Ohmann RMS is also shown. The n in both data sets is 24

Species	Coefficients and variables[a]	F	R	RMS Deschenes	RMS Ohmann	Analysis
Acer saccharum[b] A.s.	$1.206 - 0.435$ A.r. $+ 0.744$ Mg	10.3	.71	63.3	50.4	LS
Fagus grandifolia F.g.	$11.879 - 1.265$ Q. $- 0.970$ F.a. $- 0.310$ A.r. -0.005 E $- 281.6$ E^{-1}	19.3	.92	73.9	31.7	RT + .1[c]
Quercus sp. Q.	$5.763 - 0.363$ F.g. $- 0.562$ F.a. $- 0.273$ B.	24.1	.88	508.3	501.8	RT + .8
Fraxinus americana F.a.	$6.914 - 0.507$ Q $+ 0.002$ E $- 0.220$ C/N	45.6	.89	45.4	24.6	3 SLS 8[d]
Betula sp. B.	$1.105 - 0.656$ Q. $+ 0.153$ C/N					
Carya sp. C.	$0.287 + 0.082$ Cy $- 5.247$ Exp^{-1} $+ 0.882$ N $- 0.644$ Mg	4.8	.71	12.1	30.2	RT + .1
Acer rubrum A.r.	$12.331 - 0.409$ A.s. $+ 2.156$ C $- 9.688$ C$^{1/2}$	7.4	.72	48.4	29.5	3 SLS 1

[a] Mg = magnesium; C/N = carbon-nitrogen ratio; N = nitrogen; Cy = clay; E = Elevation; C = carbon; Exp = Exposure

[b] Form of variable is Ln $(BA/A + 1)$.

[c] The decimal value following RT indicates the k added to the diagonal in the $X^T X$ matrix.

[d] The value following 3 SLS is the iteration number.

group model equations were tested through prediction of the BA/A that Ohmann had measured in 24 stands in the same area that Deschenes had sampled. Ohmann had not measured carbon, nitrogen, or magnesium. We used the means of these variables from Deschenes' data in making the predictions for Ohmann's stands. All stands contained more than 50 ft² BA/A. Nineteen of 24 stands that Deschenes measured have more BA/A in *Quercus* spp. than any other species.

Comparisons between observed and predicted values of BA/A for stands from which the model equations and coefficients were determined, and the stands to which the predictive equations were applied was made on the basis of residual mean squares

$$[\text{RMS} = \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2]$$ of the untransformed data. The residual mean squares from a particular prediction was compared to the residual mean squares for the original model equation by use of an F-test to determine whether or not the residuals from the two data sets were essentially equal.

RESULTS

The model equations developed from Deschenes' data, using the three statistical criteria listed in Methods, fall into three classes. Five of the equations include both species and environmental variables, one equation includes only species as variables, and one equation has only environmental variables. The right hand side of the equations (noted as model components) and their pertinent statistics are presented in Table 2.

The solutions of the statistical procedures for Des-

chenes' data, when used on Ohmann's data, gave residual mean squares that were essentially equal for 4 of the 7 species—*Acer saccharum*, *Betula* sp., *Carya* sp., and *Acer rubrum*. For all species, except *Betula*, the residual mean squares in at least one case was not significantly different from the original estimate from Deschenes' data. In Table 2 the residual mean squares for Deschenes' data are shown with the comparable value from Ohmann's data; the statistical procedure that produced the Ohmann value is also indicated. The *Betula* sp. estimates were consistently high; the reason for this is not known.

The elaborated least-squares procedure estimates of the coefficients that produced the minimum residual mean square show that in all cases the change is usually to reduce the magnitude of both the constants (b_0 term) and the coefficients (b_i terms). Little change in the coefficients (b_i) regardless of procedure used is interpreted as an indication of a relatively unbiased estimate of the coefficient. The *Acer saccharum* and *Betula* sp. equations exhibit this stability, even though the latter does not have comparable residual mean squares. Figure 1 contains the coefficients that produce minimum residual mean squares when applied to Ohmann's data. Comparison with model equation coefficients can be made with the values in Table 1.

The relationships depicted in Fig. 1 show that *Carya* is not significantly correlated with another species in this forest. Nitrogen and clay have positive correlations with *Carya*, and magnesium is negatively correlated. The southwest exposure is correlated with less, and north and east exposures with more *Carya*. The small values of the equation co-
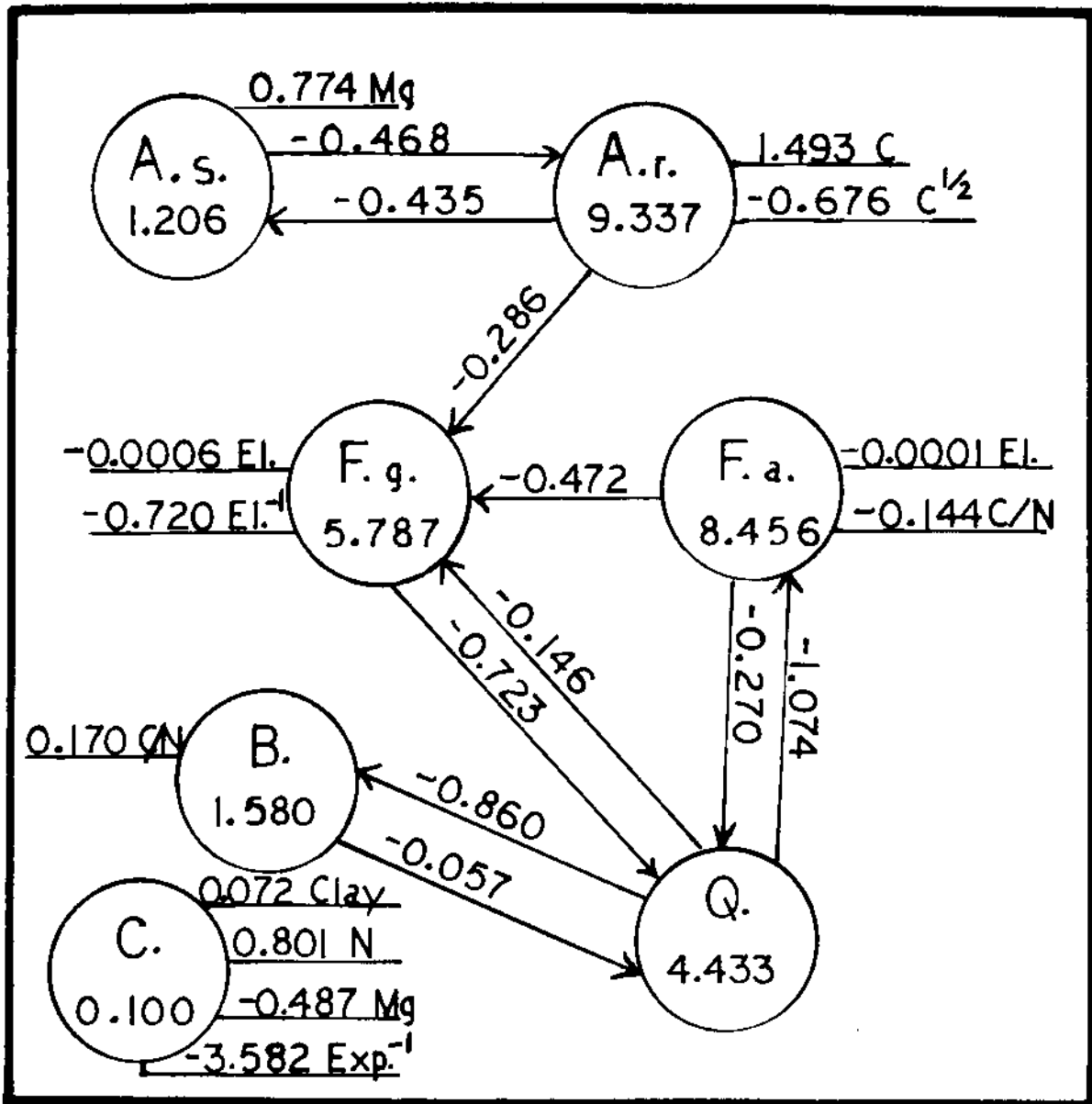
FIG. 1. A representation of the interrelatedness in the seven species model. Coefficients shown are those that gave minimum residual mean squares in the second data set. A.s. (*Acer saccharum*), A.r. (*Acer rubrum*), F.g. (*Fagus grandifolia*), F.a. (*Fraxinus americana*), B. (*Betula* sp.), Q. (*Quercus* sp.) C. (*Carya* sp.), Mg (Magnesium), C (Carbon), El. (elevation), C/N (carbon-nitrogen ratio), N (nitrogen), Exp. (Exposure).

efficients coupled with the values of the variables in the equation suggests that the quantities of *Carya* would tend to be small in any event. The range of *Carya* sp. in Deschenes' data was from 0.0 to 15.7 ft² BA/A, Table 1.

The *Quercus* equation includes only other species, all with negative coefficients. The *Quercus* species found in the study area sprout vigorously after cutting or fire (Fowells 1965). As the large constant value (b_o) in the equation (Table 2 and Fig. 1) suggests, if *Fagus*, *Fraxinus*, and *Betula* are absent or present only in small quantities, then large quantities

of *Quercus* can be expected. Figure 1 shows that the magnitude of the coefficients from *Quercus* is greater to the other two species than are the coefficients from *Betula* and *Fraxinus* to *Quercus*.

The remaining five species are correlated with environmental variables and other species. *Fraxinus* and *Betula* are both correlated with *Quercus*, the former more so than the latter (coefficients, Fig. 1). A high C/N status of the soil is correlated with more *Betula*. A low C/N status is correlated with more *Fraxinus*. The signs of the coefficients for C/N are different for these species (Fig. 1).

Fagus is correlated directly with *Acer rubrum*, *Fraxinus*, and *Quercus*. Indirectly *Fagus* is correlated with *Acer saccharum*, through *A. rubrum*, and with *Betula*, through *Quercus* (Fig. 1).

Both species of *Acer* are correlated with each other on an almost equal basis (Fig. 1).

DISCUSSION

Ecological

The classic view is that plant succession terminates with the forest in a climax condition. Therein species and environment are in equilibrium and a steady-state condition prevails. Buell et al. (1966) concluded, after exploring several types of analysis, that the long-term trend of the forests in the area (in the absence of disturbance) is to a larger component of *Acer saccharum*.

Presently, 19 of 24 of the stands that Deschenes measured, and he was studying *Acer saccharum*, have more *Quercus* sp. than any other species. Disturbance—whether fire, logging, or wind—favors *Quercus*. Christianson (1971) showed *Quercus* to be more fire resistant than is *Acer*. Charcoal at the surface of the mineral soil, cutting evidence (old stumps and sprout-origin trees), and blowdown mounds are common in these forests. Lack of these disturbances would favor the more tolerant species, as Buell, et al. (1966) concluded. The direction (signs) of the equation coefficients and the significant correlations (Fig. 1) corroborates the trend to *Acer*. But within this trend there may be alternate pathways. Our results indicate that where organic matter (carbon) tends to be high, *A. rubrum* might be favored. Where magnesium tends to be high, *A. saccharum* would be favored (Fig. 1). Deschenes (1968) reported that magnesium levels in the soil were related to the geomorphic history of the area, but even so the accumulation of organic matter over a period of nondisturbance could tip the scales to *A. rubrum* from *A. saccharum*.

The *Quercus* equation having only other species in it can be interpreted as an indication that any disturbance in the forest, such as fire, that would favor *Quercus* over other species would tend to delay the trend to *Acer* predicted by Buell, et al. (1966). The coefficients linking *Quercus* to the other species (Fig. 1) are larger from *Quercus* than to it. This is consistent with the concept of *Quercus* as an exploiter of disturbance. Given equal numbers of stems developing after a disturbance, Quercus would tend to dominate over time because crown dimensions in relation to stem dimensions are larger in these *Quercus* than in *Fraxinus americana*, for example (Stout 1967). We suggest that these larger space requirements per unit of basal area are being reflected in the relative size of the coefficients.

The species growing with *Quercus* and *Acer* will also vary from point to point. *Fraxinus* and *Betula* are correlated with the C/N ratio, with high favoring the latter and low the former (Fig. 1). Elevation, exposure, and carbon, clay, magnesium, or N content may favor one species over another at a particular point. The variation in these environmental factors, coupled with the species intercorrelations, suggests many possible combinations. Our forests can be viewed as points on an *n* dimensional hypersurface. When a change occurs in the value of any one of the variables which is a dimension of the hypersurface, a change in the forest composition can be expected.

From a pedagogical and public relations standpoint it might be good if the forest moved inexorably to a uniform climax. However, in the real world forests are subject to disturbance and variation in environmental factors. The method of analysis presented here, borrowed bodily from econometrics, can help pinpoint alternate points and, it is hoped, pathways over a complex hypersurface. An understanding of the correlations of these points and pathways takes us closer, we submit, to an understanding of the dynamics of our forests.

Statistical

The student of biological systems is faced with conditions in which there is little, if any, independence of the many variables. We have shown that in spite of the violations of the assumptions that underline regression procedures it is possible to get meaningful estimates of correlations in the forest. By the same token, the fact that in six of seven equations some estimate of the magnitude of the coefficients other than the least-squares estimates have lower residual mean squares when applied to data from other stands reinforces the warning from statistical theorists that to violate assumptions regarding regression analysis is perilous.

It would have been desirable to have found a particular procedure that produced best estimates of the model coefficients in all cases. This was not the case and it will rarely be the case, we speculate, because of the interrelatedness found in the data from any ecosystem. The interrelatedness is what makes the study of ecosystems so fascinating and, at the same time, so intractable when we use analytical techniques based on assumptions of independence. We have shown here that the use of procedures that have been developed to alleviate some of the difficulties encountered are helpful in that smaller residuals are obtained.

Methodological

Compartment models of ecosystems depend on, among other things, delineation of the compartments

and estimates of the transfer coefficients between compartments. The species and generic groups are in effect the compartments here and the coefficients the transfer coefficients, even though the coefficients are merely estimates of correlations.

Using the methodology suggested here allows the investigator to examine the interrelations among the forest species simultaneously. The critical test will come when a time dimension is added, but that is beyond the scope of this paper.

Analyses of comparable data from other facets of the ecosystem will show whether the pattern that emerges here (Fig. 1) is a prototype for ecosystem analysis or a relic of the system of analysis.

LITERATURE CITED

Buell, M. F., A. N. Langford, D. W. Davidson, and L. F. Ohmann. 1966. The upland forest continuum in northern New Jersey. Ecology 47:416–432.

Christianson, J. D. 1971. The effect of surface fires on three species of hardwood tree seedlings. Wm. L. Hutcheson Mem. For. Bull. 2:1–5.

Deschenes, J. M. 1968. *Acer saccharum* near and at the periphery of its range in New Jersey. Ph.D. Thesis. Rutgers Univ., New Brunswick, N.J.

Friedman, J., and R. T. Foote. 1957. Computation methods for handling systems of simultaneous equations. Agric. Handb. No. 94. USDA. Washington, D.C.

Fowells, H. A. 1965. Silvics of forest trees of the United States. Agric. Handb. No. 271. USDA, Washington, D.C.

Furnival, G. M., and R. W. Wilson. 1971. Systems of equations for predicting forest growth and yield, p. 43–57. *In* G. P. Patil [ed.] Statistical ecology. Vol. III. Many species, populations, ecosystems, and system analysis. The Penn. State Univ. Press, University Park and London.

Hoerl, A. E., and R. W. Kennard. 1970*a*. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12:55–67.

Hoerl, A. E., and R. W. Kennard. 1970*b*. Ridge regression: Applications to nonorthogonal problems. Technometrics 12:69–82.

Klein, L. R. 1953. A textbook of econometrics. Row, Peterson, Evanston, Ill. 355 p., illus.

Lotka, A. J. 1924. Elements of physical biology, Reissued 1956 as: Elements of mathematical biology. Dover Publications, New York.

Ohmann, L. F. 1964. Vegetation-topography relationships in New Jersey. Ph.D. Thesis. Rutgers Univ., New Brunswick, N.J.

Stout, B. B. 1967. Allometric constant variation between and within forest tree species. Ph.D. Thesis. Rutgers Univ., New Brunswick, N.J.