Applied Geography (1992), *12*, 261–279

Climate, soil and crop yield relationships in Cascade County, Montana

John P. Wilson

Department of Earth Sciences, Montana State University, Bozeman, MT 59717, USA

Kristin E. S. Gerhart

Montana State Office, Soil Conservation Service, Bozeman, MT 59715, USA

Gerald A. Nielsen

Department of Plant and Soil Science, Montana State University, Bozeman, MT 59717, USA

and **Christine M. Ryan**

Geographic Information and Analysis Center, Montana State University, Bozeman, MT 59717, USA

Abstract

The Productivity Index (PI) model estimates the productivity effects of erosion by the simulated removal of surface soil and consideration of available water-holding capacity, bulk density and pH. Although it has performed well in the US Corn Belt and elsewhere, further testing is required to demonstrate its applicability in other semi-arid environments. Evaluation of model performance in four fields in Hill and Jefferson counties, Montana, revealed a weak relationship between PI and small grain yield, possibly due to local conditions. Therefore, soils and crop data were extracted from the USDA–SCS SOILS-5 and Montana Agricultural Potentials System (MAPS) databases, as well as the county soil survey, to evaluate PI model performance and indicate appropriate changes in its design in Cascade County, Montana. Results indicate that model performance can be improved with the addition of factors to account for water balance, slope, growing degree days and calcium carbonate content. Regression of barley, spring wheat and winter wheat yield data against PI values from the original model accounted for 34, 31 and 31 per cent of the variability in yields of these three crops, respectively. R^2 increased an average of 77 per cent and accounted for 54, 59 and 58 per cent of the variations in yields when the four new factors were added to the model. These results have encouraged further efforts to develop a modified version of the PI model that uses computerized databases for county-scale assessments of semi-arid environments.

The loss of productive land as a result of soil erosion has been a worldwide concern for years and many commentators believe that it threatens our ability to feed the world's increasing population adequately (see, for example, Brown and Wolf, 1984; Pimental *et al.* 1987). Although losses of topsoil to erosion are widely

0143–6228/92/03/0261–19

considered to reduce crop productivity, this has not been well quantified for a wide range of soils (Williams 1981; Crosson 1983; Rijsberman and Wolman 1985). This situation is now changing as researchers devote more effort to quantifying the relationship between long-term soil erosion and crop productivity. Much of this work has occurred in the US, where two new modelling approaches for evaluating changes in the productivity of soils in relation to long-term erosion have emerged: the Erosion-Productivity Impact Calculator (EPIC) developed by Williams *et al.* (1984) and the Productivity Index (PI) model developed by Pierce *et al.* (1983) at the University of Minnesota, based on earlier work by Neill (1979) and Kiniry *et al.* (1983) at the University of Missouri.

EPIC is a process model which operates on a daily time step and simulates erosion, plant growth and economic components to determine erosion costs and optimal management strategies. Numerous weather, soil, tillage and crop parameters are required as inputs. The difficulty and cost of generating these inputs have so far restricted its application to national assessments in the US.

The Productivity Index (PI) model, like EPIC, considers changes in soil profile characteristics with depth. It can be used to examine the vulnerability or rate of productivity decline of a soil subjected to simulated erosion over time. A vulnerable soil is one which quickly becomes less favourable to crop growth as surface materials are removed. Relationships in the PI model have been derived and tested primarily in the US Corn Belt using corn as the target crop. The model uses the US Department of Agriculture–Soil Conservation Service (USDA–SCS) SOILS-5 database to estimate current soil properties and crop yields and assumes that the major effect of erosion is to change the soil micro-environment for root growth and, consequently, future yields. The model incorporates factors for available water-holding capacity (AWC), soil reaction (pH) and bulk density (BD), with adjustments for family particle-size class and permeability, and uses an idealized corn root distribution to weight model parameters for different horizon depths. Erosion is simulated by the incremental removal of surface soil. This soil removal produces a new potential root environment with a new combination of soil properties.

Pierce *et al.* (1984a) evaluated model performance in the Corn Belt by regressing PI against corn yields from county soil surveys and Minnesota crop equivalency ratings (CERs). CER represents the relative economic return per acre when managed for cultivated crops, permanent pasture or forestry, whichever use is computed as giving the highest net return (Pierce *et al.* 1984a). R^2 values ranged from 0·63 to 0·71 and were increased by 26 per cent (0·7 to 0·90) when histosols, frequently flooded and depressional soils, and soils with slopes exceeding 6 per cent were excluded. In addition, modified versions of the PI model (with new model statements to account for the response of crops to local soil characteristics) have been successfully applied to several locations outside the Corn Belt, including Nigeria, Hawaii, Mexico and India (Rijsberman and Wolman 1985).

Results of previous research on dryland small grain yields in Montana indicate that several new model statements may be required to achieve adequate model performance in semi-arid small-grain production regions. Burke (1984), for example, regressed selected climate and soil variables against small-grain yield at several sites in Montana under a high degree of management in terms of fertility, weed and pest control, and identified available water-holding capacity and depth to $CaCO_3$ as the most important variables in predicting yield. Larson (1986) examined the influence of soil series on small-grain yield in Montana, finding depth to $CaCO_3$ and organic matter content to be positively correlated with grain yield and pH to be

inversely correlated with yield. These properties were also highly predictive of soil test phosphorous and available water-holding capacity. Depth to $CaCO_3$ and organic matter content are not incorporated in the original (Minnesota) version of the PI model.

The calcareous nature of many agricultural soils in Montana is derived from sedimentary parent materials (soft black shales and Tertiary valley-fill) of diverse origin and composition or from wind-deposited fine sand and silt (Montagne *et al.* 1982). Although researchers in Montana (Schweitzer 1980; Munn *et al.* 1982; Burke 1984; Larson 1986) and in other parts of the world (Spratt and McIver 1972; Karathanasis *et al.* 1980) have demonstrated that the presence of carbonates in soils has an adverse effect on crop growth, there remains disagreement about the exact cause–effect relationships. The presence of carbonates can have a dominating influence on many physical and chemical soil properties, including the available water-holding capacity, pH and bulk density factors considered by the PI model. Excessive amounts of $CaCO_3$ may affect soil fertility by decreasing phosphorous and micronutrient availability. Nutrient deficiencies may influence crop rooting patterns and grain yield. Similarly, bulk density values may not predict rooting distributions within horizons of $CaCO_3$ accumulation. Roots penetrate the soil by displacing soil particles or by following existing pores or channels (Aubertin and Kardos 1965). Pores and planes of weakness may be rare in horizons where secondary $CaCO_3$ has accumulated and it may act to maintain soil strength and inhibit root penetration.

Sandor (1989) and Wilson *et al.* (1991) recently evaluated PI model performance with soils and crop data collected from four fields in Hill and Jefferson counties, Montana. Their results confirmed those of earlier Montana studies in that model performance was improved by the addition of factors to account for the location and content of $CaCO_3$ and organic mater in the soil profile. The regression of small-grain yield against the productivity indices generated with the original model accounted for 64, 67, 63 and 1 per cent of the variability in yield within fields 1 to 4, respectively. However, R^2 increased by an average of 44 per cent and accounted for 77, 69, 60 and 75 per cent of the variation in grain yield in the same fields when modified PI values were used. Moreover, R^2 increased slightly (0·69 to 0·75) for all fields collectively when cropping history was considered.

These field-scale results encouraged the efforts to develop a modified and improved version of the PI model that used computerized climate and soils databases for county-scale soil erosion/crop productivity assessments in the semi-arid environments reported in this paper. The specific objectives of this study were to:

1. evaluate the performance of the existing PI model in Cascade County, Montana, where soil properties and crop yields vary widely;
2. identify climate, soil and other site variables important to small grain yield in Montana;
3. suggest modifications to improve the performance of the PI model in Montana and other northern Great Plains environments.

Study area description

Cascade County is situated in north central Montana (47°22′N, 111°20′W), bordering the eastern slopes of the Rocky Mountains (Fig. 1). The county covers 6177 km^2 and consists of nearly flat or rolling plains in the north (900–1200 m

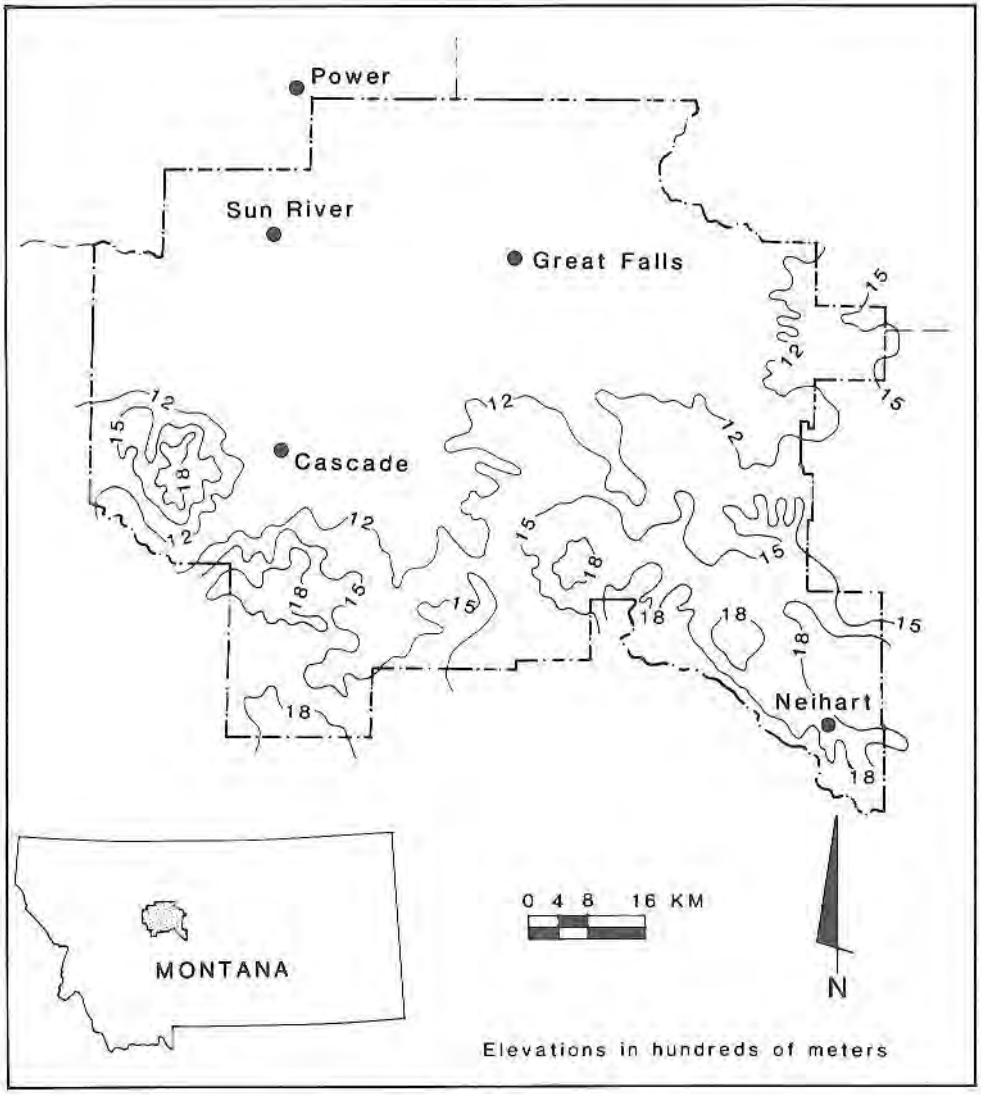

Figure 1. Cascade County location map

elevation) and benchlands and mountainous areas in the south and east (1200–2000 m). Approximately 35 per cent of the county is cropped and most of the remainder is range. The principal crops are wheat, barley, hay and pasture. A small proportion of cropland (7 per cent) is irrigated. The range areas support short and mid-height grasses as well as some shrubs, while some higher elevations are forested (Gerhart 1989).

The climate is characterized by low humidity, low winter and high summer temperatures, and mostly sunny days. Temperatures range from mean monthly lows of −12°C to highs of 28°C and the frost-free period varies from 85 to 135 days. Average wind speeds vary between 17 and 26 km hr^{-1} and flow predominantly from

Table 1. 1987 Cascade County crop data

Crop	Dryland crops		Irrigated crops		Total production (t)
	Area harvested (ha)	Average yield (t ha^{-1})	Area harvested (ha)	Average yield (t ha^{-1})	
Barley	32 820	1·7	2 390	4·0	65 355
Durum wheat	120	2·5	–	–	300
Oats	365	1·7	120	2·7	945
Spring wheat	10 120	1·7	2 025	3·9	25 100
Winter wheat	47 510	2·4	245	3·6	114 905

the southwest. The highest wind velocities occur in the autumn and spring, when soil surfaces are exposed to erosive winds, although strong, warm chinook winds along the Rocky Mountain front in winter serve to melt protective snow covers and leave the soil vulnerable to wind desiccation and removal by wind and water. Average annual precipitation ranges from 25–40 cm on the flat, rolling plains to 40–100 cm on the benchlands and mountains in the south and east. It falls mainly as snow between November and March, although snow may fall as early as September and as late as July. The maximum precipitation occurs in June and is followed by occasional thunderstorms throughout the summer. Precipitation varies greatly from year to year, with detrimental effects for agriculture in dry years. The soils are not strongly developed (Haploborolls are common) and fall mostly within the Ustic soil moisture regime and frigid temperature regime (Montagne *et al.* 1982). The cropped soils occur on the flat and gently sloping plains, fans, benches and terraces. Most of these agricultural soils are classified as mollisols and entisols.

Cascade County forms the southern corner of Montana's 'Golden Triangle', an area well known for its extensive cultivation of small grains. The northern half of the county contains the majority of the cropland, although precipitation is generally lowest there. The southern half enjoys a more abundant moisture supply, but its soils and topography are less suited to grain crops and highly mechanized agriculture. The production of small grains in Cascade County is dominated by dryland winter wheat, spring wheat and barley (Table 1). Yields are relatively low as Cascade County farmers cope with several productivity problems related to soil chemistry (high $CaCO_3$ concentrations and salinity), cobbly and gravelly soils, soil crusting and a substantial wind erosion hazard (Clark *et al.* 1982), in addition to drought and hail. The sale of small grains makes a substantial contribution to the county's economy in most years despite these problems; for example, farmers received \$19·3 million from the sale of crops in 1987 (when most of the county received normal precipitation).

Methods and data sources

Productivity Index model

The Minnesota PI model (Pierce *et al.* 1983, 1984a, 1984b, 1984c) can be written as:

$$\mathrm{PI} = \sum_{i=1}^{n} (A_i \times C_i \times D_i \times WF_i) \qquad (1)$$

where

A_i = sufficiency of AWC
C_i = sufficiency of bulk density
D_i = sufficiency of pH
WF_i = weighting factor representing an idealized rooting distribution
n = number of horizons in the depth of rooting.

The sufficiency curves are based on root response to each variable normalized over the range 0·0–1·0 (Fig. 2). Essentially a sufficiency value estimates how deficient a soil environment is for optimum root development in terms of a dimensionless numerical index. The weighting factor modifies the 'importance' of each horizon's PI values according to that horizon's depth within the rooting zone (Fig. 2d).

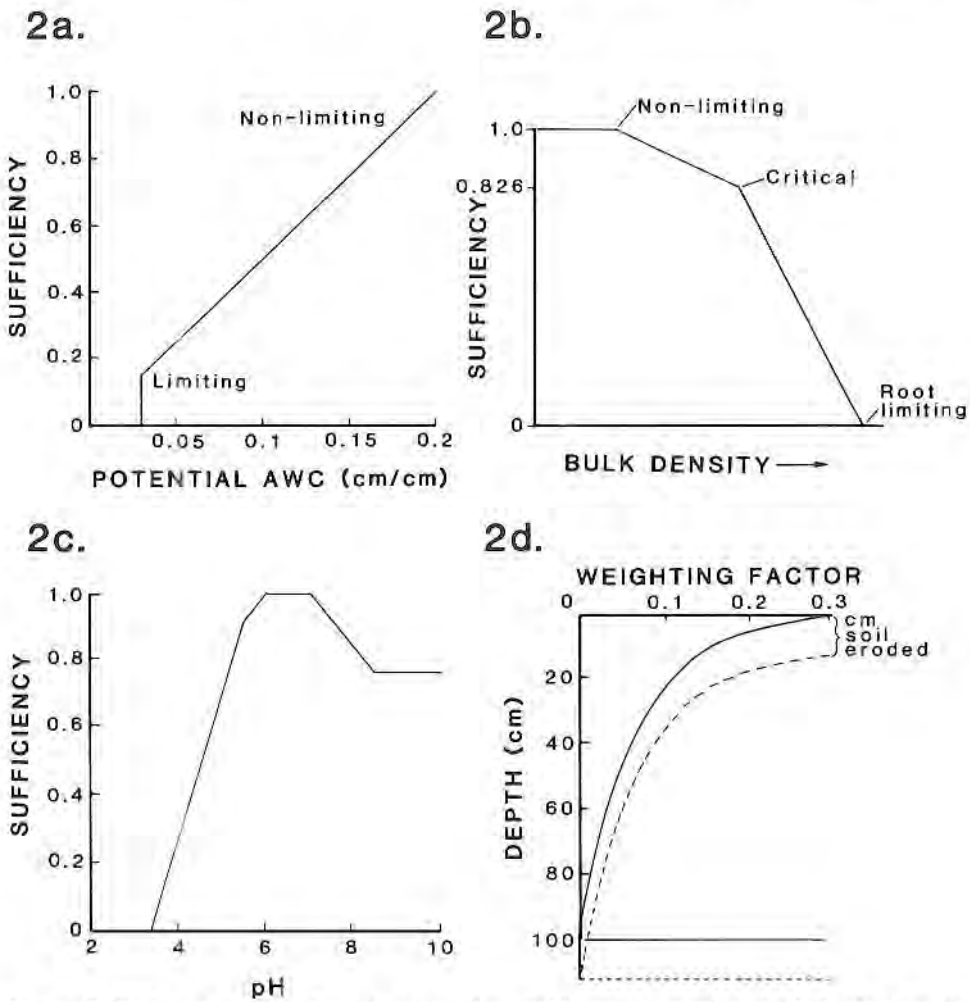

Figure 2. Sufficiency curves for (a) potential available water capacity, (b) bulk density, (c) pH and (d) the sliding weighting factor used by the PI model (after Pierce *et al* 1983). No units are reported on the X-axis in (b) because the non-limiting, critical and root-limiting bulk densities vary with soil family texture class

Underlying the PI approach is the premise that crop yields are closely related to the rooting environment provided by the soil. The model focuses, therefore, on inherent soil properties, and based on these variables it calculates the productive capacity of the soil represented by PI values. By examining the physical and chemical characteristics of the soil profile as soil is removed, PI provides a quantitative evaluation of the changes in a soil's rooting environment and overall productive value that occur in response to erosion processes.

Data requirements and model testing

Copies of the PI model programs and manual (Winkelman *et al.* 1984) were obtained from the Department of Soil Science at the University of Minnesota. The Montana SOILS-5 database contained most of the soils data required to run the model and was provided on magnetic tape by the USDA–SCS. The first two pre-processing programs (GENERS5 and GPIFORM) are run on a mainframe computer and accomplish four operations as follows:

1. the SOILS-5 data required to run the PI model are extracted from magnetic tape;
2. the sand and clay contents of each horizon are used to assign soil family texture classes;
3. the relationship between bulk density and available water-holding capacity, clay content and air-filled porosity are compared, and the bulk densities are adjusted for horizons with highly developed soil structure (Pierce *et al.* 1983; Winkelman *et al.* 1984);
4. the soils data used by the PI model are formatted and linked with the appropriate series names and phase numbers for downloading and subsequent analysis on an IBM-compatible PC.

A look-up table was added to the GENERS5 program to estimate bulk densities, since these values were missing for 95 per cent of the soil series in Cascade County. This look-up table was derived from the bulk density triangle developed by Grossman and Baumer at the USDA–SCS National Soil Laboratory in Lincoln, Nebraska (Fig. 3). The verbal soils descriptions in the Cascade County Soil Survey (Clark *et al.* 1982) were used to select and manually retain the most appropriate surface textures in shortened SOILS-5 records. The original records contained data for surface textural classes that did not occur in Cascade County.

The TODIRECT program was used to convert the downloaded data file from a sequential to a random access file format and the GRAPHPI2 program was used to compute PI values. GRAPHPI2 generates several PI values for each soil series. First, a current productivity rating is calculated. Next, PI values are computed for successive soil profiles as computer-simulated erosion removes 2 cm of topsoil at a time. These calculations continue until 100 cm are removed or PI reaches zero. GRAPHPI2 also plots the sequence of PI values against centimetres eroded in a standard graph format. Finally, the program displays current PI values as well as PI values after 50 and 100 cm of soil have been lost, and a vulnerability index expressing the rate at which productivity diminishes with decreasing profile depth.

Pierce *et al.* (1984a) regressed PI values against current crop yields to evaluate model performance in the Corn Belt, concluding that R^2 values of at least 0·70 demonstrated that PI values provided good estimates of crop productivity. A similar approach was used to evaluate model performance in Cascade County. The barley, spring wheat and winter wheat yield data used for these comparisons were obtained from the Cascade County Soil Survey (Clark *et al.* 1982). These yields

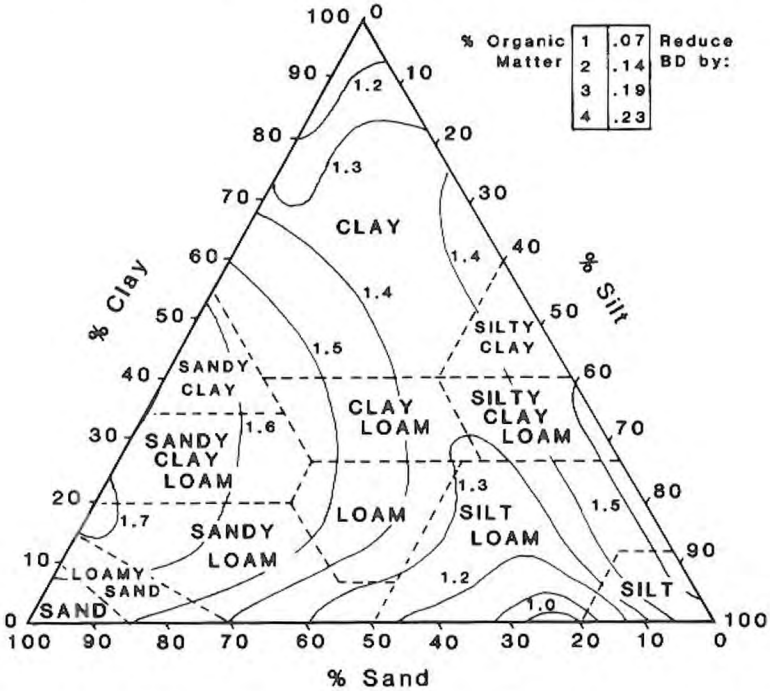

Figure 3. Bulk density estimated from soil textural class triangle developed by Grossman and Baumer, USDA–SCS National Soil Science Laboratory, Lincoln, Nebraska

were preferred over those reported in the SOILS-5 database because the latter represented generalized estimates collected from several states which, in some cases, reflect little of the local climatic, topographic, farming and soil characteristics which influence yields in Cascade County. Dryland crop yields were used since yield data for irrigated crops were less abundant (see Table 1) and, more importantly, because highly variable irrigation practices result in highly variable yields and soil degradation processes unrelated to erosion.

Assumptions similar to those made in previous applications of the PI model were also made for the first set of comparisons tried in this study. First, local climatic variations were presumed to have no effect on model performance (that is, crop productivity) in Cascade County. Secondly, the influence of landscape position (as described by Onstad *et al.* 1985) was ignored except for the exclusion of steeply sloping soils. Thirdly, it is assumed that a high level of farming technology was applied and, therefore, that farming technology could not explain variations in crop yields. Finally, plant nutrients and moisture conditions were presumed to be adequate for crop growth. Some of these assumptions were waived when new components were added to the original (Minnesota) PI model.

Data requirements and statistical analysis required for model additions

The low R^2 values obtained with the original PI model led to the addition of four new variables representing water balance, growing degree days, slope and $CaCO_3$ content, which may strongly influence crop yields.

Data for annual precipitation, potential evapotranspiration (PET) and growing degree days (GDD) are not recorded by soil series in county soil survey reports. Fortunately, these reports do list range and township locations for 'typical' soil profiles representing each soil series. These locations were used to assign annual precipitation, potential evapotranspiration and growing degree days totals from the Montana Agricultural Potentials System (MAPS) database (Nielsen *et al.* 1990) to specific soil series. The MAPS database divides the State of Montana into 18 000 cells (each three minutes latitude by three minutes longitude) and stores annual precipitation and potential evapotranspiration in inches of water. The water balance factor was obtained by subtracting this from precipitation. The growing degree days factor counted each degree fahrenheit above 50°F, since this mean daily temperature best represents the minimum daily temperature required for small-grain production. The slope and $CaCO_3$ factors were both computed from soil series data reported in the Cascade County Soil Survey (Clark *et al.* 1982). However, the slope ranges (for example, 0–2 per cent, 4–8 per cent, etc.) were converted to single values (mid-points) and the ratings of soil reactivity with hydrochloric acid were translated into depths to affected layers and effervescence classes.

These four factors and the original PI values were combined with the SAS stepwise regression procedure (Freund and Littell 1986) in an attempt to explain yield variation more completely than with PI values alone. Calcium carbonate content was treated as a qualitative variable with four classes as follows: no $CaCO_3$ in the top 20 cm of the soil profile; slight $CaCO_3$; strong $CaCO_3$; and violent effervescence occurring within 20 cm of the surface. This depth is recognized as the plough layer where root development and nutrient uptake by plants are most critical. Calcium carbonate content was represented by three indicator variables (X1, X2 and X3 for the last three classes) in the regression models, each taking on the values 0 and 1. Interaction effects were introduced into the models in the usual manner, by including cross-product terms linking each of the indicator variables with each of the continuous variables (Neter *et al.* 1986). Given this approach, four continuous variables (PI, slope, water balance and growing degree days), three indicator variables (representing four $CaCO_3$ classes) and 12 interaction terms were regressed against the dependent variable (yield) for each of the three crops used in this study. The 0·1 significance level was used to add and retain variables in all three regression models.

Results

PI model evaluation

Three sets of scatterplots, regression equations and coefficients were produced from the simple bivariate regression of PI against barley, spring wheat and winter wheat yield, respectively (Fig. 4). The variations in PI explained less than 35 per cent of the variations in yields for all three crops. These low R^2 values and large scatter of data points suggest that the original version of the PI model does not satisfactorily estimate crop productivity in Cascade County.

Pierce *et al.* (1984a) obtained similar correlation coefficients (R^2 values of 0·23–0·68) when they first regressed PI values against corn yields in several Iowa, Indiana, South Dakota and Minnesota counties. However, they improved their correlations and obtained R^2 values of 0·70–0·77 when they excluded histosols, frequently flooded, depressional and slope soils from subsequent model tests. This

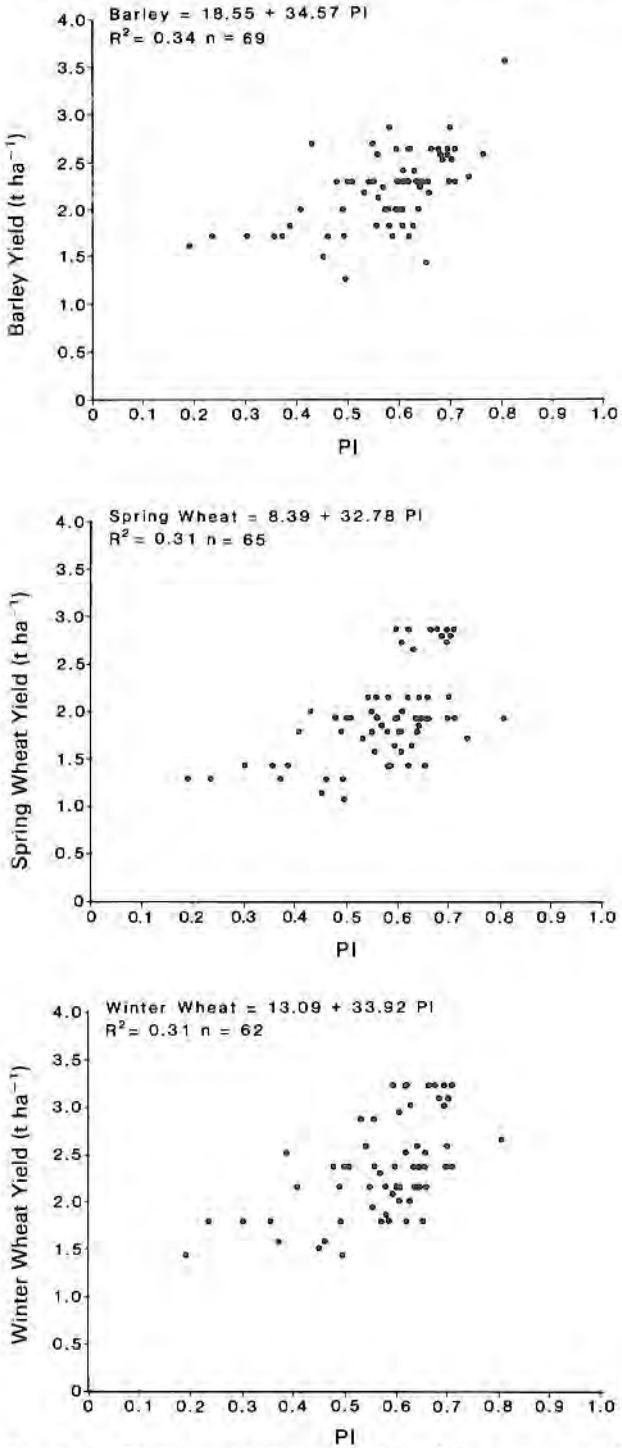

Figure 4. Bivariate regression scatterplots comparing PI values with yields of barley, spring wheat and winter wheat

strategy could not be repeated in Cascade County for two reasons. First, three of the four exclusion cases (histosols, frequently flooded and depressional soils) seldom occur in Cascade County and, secondly, the scatterplots did not produce outliers that could be ascribed to other special conditions. The exclusion of soils with slopes exceeding 6 per cent eliminated 75 per cent of the soil series from further evaluation and produced R^2 values of 34, 3 and 11 per cent for barley, spring and winter wheat, respectively. These results prompted the search for additional soil and climate factors to improve PI model performance in semi-arid environments.

Multiple regression outcomes

Water balance, growing degree days, slope and $CaCO_3$ factors, in addition to the available water-holding capacity, bulk density and pH factors in the original (Minnesota) PI model, were used in an attempt to explain crop yields. Table 2

Table 2. Cascade County crop yields by soil series and selected soil/climatic parameters[a]

	Barley		Spring wheat		Winter wheat	
New variables	**No.**	**Average yield (t ha^{-1})**	**No.**	**Average yield (t ha^{-1})**	**No.**	**Average yield (t ha^{-1})**
Water balance (cm)						
10·1–20	1	2·15	1	1·81	1	2·22
0·1–10	0	–	0	–	0	–
−9·9–0	1	3·33	1	1·81	1	2·49
−19·9–−10	6	2·12	5	1·94	6	2·30
−29·9–−20	12	2·23	10	2·01	9	2·43
−39·9–−30	33	2·07	32	1·84	29	2·18
−49·9–−40	16	1·87	16	1·66	16	2·01
Growing degree days (degrees celsius above 10°C × days)						
1315–1424	9	2·21	9	2·04	7	2·43
1205–1314	44	2·05	41	1·81	40	2·16
1095–1204	14	1·97	14	1·77	13	2·14
985–1094	2	2·74	1	1·81	2	2·25
Slope (%)						
10·6–14·0	5	1·74	5	1·55	5	1·90
7·1–10·5	8	2·11	8	1·75	8	2·15
3·6–7·0	19	1·98	17	1·67	18	2·03
0·0–3·5	37	2·16	35	1·97	31	2·35
Calcium carbonate content (in top 20 cm of soil profile)						
Violent effercescence	1	2·15	1	1·81	1	2·22
Strong effervescence	18	2·04	18	1·74	14	2·05
Slight effervescence	8	1·90	6	1·44	6	1·93
No reaction	42	2·12	40	1·93	44	2·28

[a] No. refers to number of soil series used to grow specific crops

shows the yields associated with different combinations of soils grouped according to these new factors. These and similar results from other studies linking variations in yields to these variables encouraged the search for model additions using these factors. The results of regressing PI, slope, water balance and growing degree days (represented by three indicator variables and 12 interaction terms) against the dependent variable (yield) for each of the three crops are summarized in Table 3. All three models used between four and six independent variables and explained at least 54 per cent of the variability in crop yields (using the 0·1 level of significance) and at least 52 per cent using the usual 5 per cent significance level.

PI explained 31·3–34·0 per cent of the variability in crop yield and was positively correlated with yield in all three cases. (The regression equations are similar and regression coefficients identical to those obtained with simple bivariate regression

Table 3. Multiple regression results[a]

Step	Variable	Parameter estimate	Partial R^2	Model R^2	Prob>F
	Dependent variable: BARLEY (t ha^{-1})				
	INTERCEPT	32·0296			
1	PI	30·2707	0·340	0·340	0·0001
2	SLOPE	−0·8757	0·084	0·424	0·0028
3	WBAL	0·5397	0·097	0·521	0·0006
4	WBAL1	0·2125	0·020	0·541	0·1007
	Dependent variable: SPRING WHEAT (t ha^{-1})				
	INTERCEPT	−17·1865			
1	PI	40·6476	0·313	0·313	0·0001
2	SLOPE	−0·7304	0·126	0·439	0·0004
3	WBAL2	1·1718	0·039	0·478	0·0370
4	M50GDD2	0·0062	0·045	0·523	0·0202
5	M50GDD	0·0110	0·039	0·562	0·0255
6	SLOPE1	−0·7613	0·025	0·587	0·0670
	Dependent variable: WINTER WHEAT (t ha^{-1})				
	INTERCEPT	−0·7358			
1	PI	38·4072	0·313	0·313	0·0001
2	SLOPE	−1·1160	0·121	0·434	0·0007
3	WBAL	0·4403	0·060	0·494	0·0110
4	M50GDD	0·0099	0·039	0·534	0·0327
5	WBAL2	0·3681	0·023	0·557	0·0915
6	SLOPE2	0·6140	0·027	0·583	0·0668

[a] The independent variables are abbreviated as follows: M50GDD = growing degree days using 10°C threshold; M50GDD2 = growing degree days × X2 (i.e., soils with strongly effervescent surface horizons); PI = Productivity Index; SLOPE = slope in percent; SLOPE1 = slope × X1 (i.e., soils with slightly effervescent surface horizons); SLOPE2 = slope × X2 (i.e., soils with strongly effervescent surface horizons); WBAL = water balance (i.e., annual average precipitation − potential evapotranspiration); WBAL1 = water balance × X1 (i.e., soils with slightly effervescent surface horizons); WBAL2 = water balance × X2 (i.e., soils with strongly effervescent surface horizons)

because PI was entered first in every case.) The lack of interaction terms including PI in Table 3 indicates that the different combinations of PI and $CaCO_3$ did not affect yields in significantly different ways.

Slope, by itself, was inversely correlated with yield in all three cases and explained 8·4–12·6 per cent of the variability in crop yield. In addition, two interaction terms (SLOPE1 and SLOPE2) explained 2·5 and 2·7 per cent of the variability in spring and winter wheat yield, respectively. The slopes with slightly effervescent soils (SLOPE1) followed the general trend in that they were inversely related to spring wheat yield, whereas slopes with strongly effervescent soils (SLOPE2) were positively correlated with winter wheat yield. Water balance was positively correlated with yield and appeared in all three models by itself or as part of interaction terms. Hence, the water balances for soils with no $CaCO_3$ in the top 20 cm (WBAL) and slightly effervescent soils (WBAL1) combined to explain 11·7 per cent of the variability in barley yield. Similarly, the water balance for soils with no $CaCO_3$ in the top 20 cm and/or strongly effervescent soils (WBAL2) combined to explain 3·9 and 8·3 per cent of the variability in spring and winter wheat yields, respectively.

The growing degree days variable (M50GDD) was the least important of the four continuous variables. It explained 3·9 per cent of the variability in yield in each of the wheat models, and one interaction term (M50GDD2) representing growing degree days for strongly effervescent soils explained 4·5 per cent of the variation in spring wheat yield. The six significant interaction terms used in the models (Table 3) demonstrate that the creation and analysis of multiple combinations of soils grouped by $CaCO_3$ content improved model performance in all three instances.

Discussion

Overall, these new and expanded regression models explained 54·1, 58·7 and 58·3 per cent of the variability in Cascade County barley, spring wheat and winter wheat yields, respectively. Although these results indicate that the new and expanded models performed much better than the original (Minnesota) model in Cascade County, they do not match the results obtained by Pierce *et al.* (1984a) using the original model in the US Corn Belt. The model extensions and data used in this study may have contributed to this result in at least four different ways. The first two problems are concerned with data limitations and the last two consist of model changes that might improve model performance further.

PI model input data

The latest version of the PI model developed by Pierce *et al.* (1983, 1984a, 1984b) used the national USDA–SCS SOILS-5 database to minimize data collection, increase computational efficiency and minimize questions about data reliability and accuracy. However, two problems encountered in this study suggest that there may be regional variations in the quality of these databases which affect model performance.

One problem concerned the lack of bulk density values in the Montana SOILS-5 database. The missing values were estimated from soil textures (see Fig. 3); however, Mausbach and Gamble (1984) have shown that accurate bulk density estimates require information about parent material (loess, glacial till, alluvium, residuum, and so on) as well as soil texture. The extent to which the solution used in this study contributed to poor model performance in Cascade County (compared

to the Corn Belt) depends on two factors: the relative proportions of soils with measured bulk densities in different parts of the country; and the levels of accuracy that accompany the methods used by USDA–SCS soil survey personnel to estimate bulk densities for the large numbers of soils lacking measured values.

A second potential problem follows from the use of several typifying pedon (profile) locations to prepare soil series or phase descriptions in the SOILS-5 database. This approach favours the inclusion of local conditions from states and counties with older soil surveys, including some affecting soil development and productivity. In contrast, many of the soils identified in Cascade County had already been described in other parts of the country where these soils were first surveyed, so that local (Cascade County) environments are probably not well represented in the SOILS-5 database. The broader consequences of these limitations are discussed below in conjunction with those affecting the yield data.

Yield data used for model testing

The yield data used for model verification in Cascade County may have contributed to poor model performance as well. Pierce *et al.* (1983) preferred the use of long-term average yields over yields collected from 'typical' years for model verification because they are less likely to reflect variations in weather and farming practices. However, the accuracy of the 'long-term average yields' reported in county soil surveys may vary regionally along with the magnitude of variations in climate and soils within counties and the local knowledge of soil survey personnel. This problem may have contributed to the lower R^2 values obtained in this study since Cascade County was chosen for its diversity of climatic and soils conditions (Gerhart 1989). Hence, some of the yields reported by Clark *et al.* (1982) in the Cascade County soil survey may reflect local conditions and farming practices that are not documented in the survey. There is obviously no way to incorporate these conditions and practices in model applications that rely exclusively on computerized databases and county soil survey information.

There may be additional (and more substantial) problems accompanying the use of these yield data for model verification, given the tendency for yields to be expressed in multiples of twos and fives (28, 30, 32, 35, 40 bushels, and so on) in soil survey reports. Pierce *et al.* (1984a) recognized some of these problems with soil survey yields in the Corn Belt and used the crop equivalency ratings developed by Rust and Hanson (1975) in addition to corn yields for model verification. However, this approach could not be used in this study because similar ratings have not been compiled for Montana soils.

Both the data input and yield problems make for 'noisy' data, with the result that some of the clarity in climate–soil–yield relationships in Cascade County may have been lost. This may also help to explain regional variations in model performance because these problems are probably less severe in other parts of the country (such as the Corn Belt) where soil surveys are more detailed and agricultural land uses are less extensive. Notwithstanding these problems, the results from PI and yield studies in different parts of the world suggest two kinds of additional modifications to PI that may further improve performance in semi-arid environments.

Existing PI model structure

One set of changes concerns the structure of the original (Minnesota) PI model. Two aspects of this model, the depth of rooting used with the weighting factor and

the equal importance assigned to the available water-holding capacity, bulk density and pH factors, may not be suited to applications in semi-arid, small-grain environments.

The consequences of changing the depth of rooting used with the weighting factor was considered by Sandor (1989) as part of his evaluation of PI model performance in four fields in Hill and Jefferson counties, Montana. Although barley and wheat have different rooting depths and water use patterns from corn (Rickman *et al.* 1977; Proffitt *et al.* 1985), the relative distribution of weights within the rooting zone means the choice of rooting depth has a minor impact on model output. Sandor (1989) substituted a rooting depth of 178 cm for barley and wheat, but found that PI was not very sensitive to changes in rooting depth so long as the new weighting functions emphasized those soil characteristics closest to the surface.

Model performance might also have been improved in this study by varying the equal weights assigned to the available water-holding capacity, bulk density and pH inputs in the original version. These weights were not altered due to the relatively small data set representing agricultural soils in Cascade County and a desire to first evaluate the effects of adding additional parameters. The effects of varying the weights will be considered when sufficiency curves are developed for the new model factors.

New sufficiency curves

The multiple regression results show a need to expand the PI model to include at least four additional factors which affect crop productivity in Cascade County (see Table 3). The next challenge involves the development and evaluation of sufficiency curves, so that these factors can be represented in the PI model in the same way as the original inputs. Several limitations with the present study, some of them unique to specific factors, must be overcome to build these curves.

The first two limitations (left unattended) would diminish the prospects of building appropriate sufficiency curves for all four environmental factors. The first of these limitations, concerned with the small number of agricultural soils examined in this study, has already been rectified by expanding the next phase of the study to examine agricultural soils in 18 Montana counties. The reliance on linear regression models to explain climate–soil–yield relationships in Cascade County constituted the second general limitation. Solving this problem will require the development and evaluation of non-linear as well as linear models to explain climate–soil–yield relationships. Several precedents suggest that non-linear models will perform better than linear models. The three sufficiency curves used in the original (Minnesota) version of the PI model incorporate non-linear relationships (see Fig. 2). The PI and water use factors proposed by Gantzer and McCarty (1987) as part of a modified PI model to predict corn yields on a claypan soil in Missouri incorporate power curve linear transformations, and the generalized sufficiency curves proposed by Pierce *et al.* (1984c) for water balance and growing degree days envisaged non-linear relationships (Fig. 5). Great care must be exercised in this search so that the final model does not achieve statistical success at the expense of biologically interpretable results.

The remainder of the problems are specific to one or more of the relevant environmental variables. The major challenge to building the water balance curve envisaged by Pierce *et al.* (1984c) is deciding how to represent the water balance on the X-axis. For example, it could be represented as the quotient of annual precipitation and potential evapotranspiration or as centimetres of water (as shown

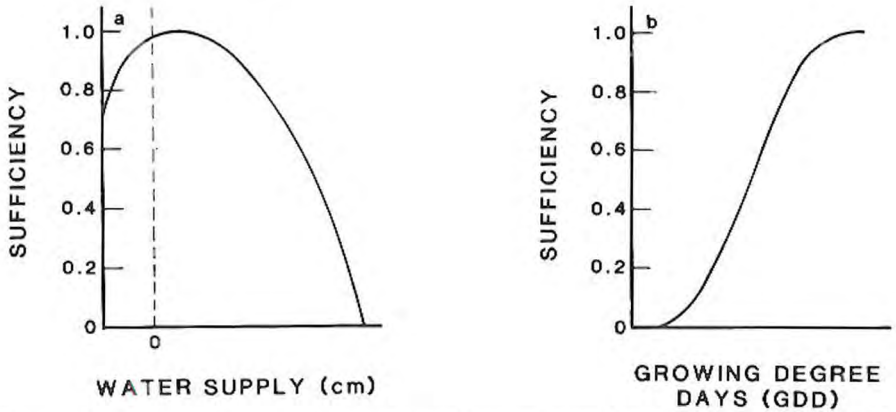

Figure 5. Schematic response curves for water supply and growing degree days (after Pierce *et al.*) 1984c)

in Fig. 5). Similarly, a scale representing growing degree days must be assigned to the *X*-axis of the GDD graph in Fig. 5, although differences in water use and maturing time for spring and winter small grains will probably require separate curves for these crops.

The regression models in Table 3 imply that slope has a direct effect on crop growth and that every increase in slope equates to a specified decrease in yield. However, the 'slope' parameters in these models probably represent microclimate and soil properties that are inversely related to crop production and not considered elsewhere in the model. Slope probably represents the effects of runon/runoff on crop yields in semi-arid environments with the result that the effects of sloping terrain would be best represented by a factor which combines information about slope aspect, gradient, length and curvature, and the location of specific soils on a slope (Larson *et al.* 1983; Daniels *et al.* 1987; Ulmer and Patterson 1988a, 1988b). The slope sufficiency curve built during the next phase of this study will be limited to slope gradient because current soil survey reports and the SOILS-5 database do not include information about these additional terrain attributes.

There are two major challenges to overcome in building a sufficiency curve for $CaCO_3$ content. The first is the conversion of the effervescence classes to some continuous scale that can be represented by a sufficiency curve. The second challenge is to sort out whether this factor needs to consider both the depth to, and concentration of, $CaCO_3$ within the top 20 cm of soil or simply the presence/absence of $CaCO_3$ in this zone. The variety of significant interaction terms in Table 3 illustrates the difficulties involved in trying to interpret the relationships between the locations and concentrations of $CaCO_3$ and the other significant independent variables with yield.

Conclusions

The comparison of the PI values obtained with the original model and crop yields in Cascade County illustrates the dangers of taking a model developed and verified for one set of environmental conditions and applying it without further testing in a vastly different environment. The PI model was designed and tested to evaluate the productivity effects of accelerated soil erosion with the USDA–SCS SOILS-5

database in the US Corn Belt. However, its design is such that it is relatively easy to apply to other regions, so long as the appropriate inputs are used.

The original PI values explained less than 35 per cent of the variability in crop yield. This indicates why the original model should not be used to evaluate the productivity effects of erosion in Cascade County. However, the multiple regression outcomes suggest a new and expanded model that probably could be used for this assessment. R^2 increased an average of 77 per cent and accounted for 54, 59 and 58 per cent of the variation in barley, spring wheat and winter wheat yield, respectively, when new factors representing water balance, slope, growing degree days and $CaCO_3$ content were added to the available water-holding capacity, bulk density and pH factors in the original model. Such a model would require computerized soil and climate databases for input; however, the explosive growth in geographic information systems technologies and databases means that data of this kind are becoming more readily available in many parts of the world. The challenge now is to build sufficiency curves for the appropriate factors in different parts of the world and to evaluate their performance against the field and laboratory data reported in modern soil surveys.

Acknowledgements

We thank Mary Blackwood, David Gerhart, Matt Miller, John Pings, Robert Snyder and Loretta Thomas for their assistance with data compilation and computer programming. Jeff Jacobsen, Gordon Decker and two anonymous reviewers provided valuable comments on earlier drafts of the paper. Contribution no. J-2635 of the Montana Agricultural Experiment Station and Grant no. SES-87212881 of the National Science Foundation Grant through the Geography and Regional Science Program enabled the research reported in this paper to be undertaken.

References

Aubertin, G. M. and Kardos, L. T. (1965) Root growth through porous media under controlled conditions: 1. Effect of pore size and rigidity. *Soil Science Society of America Proceedings*, 29, 290–293.

Brown, L. R. and Wolf, E. C. (1984) *Soil erosion: quiet crisis in the world economy*. Washington, DC: Worldwatch Institute, Paper No. 60.

Burke, T. H. (1984) *Evaluating selected soil morphological, classification, climatic, and site variables that influence dryland small grain yield on Montana soils*. Unpublished MSc thesis, Montana State University, Department of Plant and Soil Science.

Clark, C. O., Farnsworth, D. H., Miller, F. T. and Wright, B. N. (1982) *Soil survey of Cascade County area, Montana*. Washington, DC: US Department of Agriculture, Soil Conservation Service.

Crosson, P. R. with Stout, A. T. (1983) *Productivity effects of cropland erosion in the United States*. Washington, DC: Resources for the Future, Inc.

Daniels, R. B., Gilliam, J. W., Cassell, D. K. and Nelson, L. A. (1987) Quantifying the effects of past soil erosion on present soil productivity. *Journal of Soil and Water Conservation*, 42(3), 183–187.

Freund, R. J. and Littell, R. C. (1986) *SAS system for regression*. Cary: SAS Institute.

Gantzer, C. J. and McCarty, T. R. (1987) Predicting corn yields on a claypan soil using a soil productivity index. *Transactions of the American Society of Agricultural Engineers*, 30(5), 1347–1352.

Gerhart, K. E. S. (1989) *Evaluating the performance of the Soil Productivity Index (PI) model in Cascade County, Montana*. Unpublished MSc thesis, Montana State University, Department of Earth Sciences.

Karathanasis, A. D., Johnson, V. A., Peterson, G. A., Sander, D. H. and Olson, R. A. (1980) Relation of soil properties and other environmental factors to grain yield and quality of winter wheat grown at international sites. *Agronomy Journal*, 72, 329–336.

Kiniry, L. N., Scrivner, C. L. and Keener, M. E. (1983) *A soil productivity index based upon predicted water depletion and root growth*. Columbia: University of Missouri-Columbia, College of Agriculture, Agricultural Experiment Station, Research Bulletin No. 1051.

Larson, M. H. (1986) *The influence of soil series on cereal grain yield*. Unpublished MSc thesis, Montana State University, Department of Plant and Soil Science.

Larson, W. E., Pierce, F. J. and Dowdy, R. H. (1983) The threat of soil erosion to long-term crop production. *Science*, 219, 458–465.

Mausbach, M. J. and Gamble, E. E. (1984) Bulk density summaries for midwestern soils. *Journal of Soil and Water Conservation*, 39(3), 203–205.

Montagne, C., Munn, L. C., Nielsen, G. A., Rogers, J. W. and Hunter, H. E. (1982) *Soils of Montana*. Bozeman: Montana Agricultural Experimental Station, Bulletin No. 744.

Munn, L. C., Schweitzer, B. D., Lund, R. E. and Nielsen, G. A. (1982) Use of paired sampling to quantify soil productivity. *Agronomy Journal*, 74, 148–151.

Neill, L. L. (1979) *An evaluation of soil productivity based on root growth and water depletion*. Unpublished MSc thesis, University of Missouri, Department of Agronomy.

Neter, J., Wasserman, W. and Kutner, M. H. (1985) *Applied linear statistical models* (2nd ed.). Homewood: Irwin.

Nielsen, G. A., Caprio, J. M., McDaniel, P. A., Synder, R. D. and Montagne, C. (1990) MAPS: a GIS for land resource management in Montana. *Journal of Soil and Water Conservation*, 45(4), 450–453.

Onstad, C. A., Pierce, F. J., Dowdy, R. H. and Larson, W. E. (1985) Erosion and productivity interrelations on a soil landscape. In *Erosion and soil productivity: proceedings of the national symposium on erosion and productivity*, pp. 193–200. St Joseph: American Society of Agricultural Engineers.

Pierce, F. J., Larson, W. E., Dowdy, R. H. and Graham, W. A. P. (1983) Productivity of soils: assessing long-term change due to erosion. *Journal of Soil and Water Conservation*, 38(1), 39–44.

Pierce, F. J., Dowdy, R. H., Larson, W. E. and Graham, W. A. P. (1984a) Soil productivity in the Corn Belt: an assessment of erosion's long-term effects. *Journal of Soil and Water Conservation*, 39(2), 131–138.

Pierce, F. J., Larson, W. E. and Dowdy, R. H. (1984b) Soil loss tolerance: maintenance of long-term productivity. *Journal of Soil and Water Conservation*, 39(2), 136–138.

Pierce, F. J., Larson, W. E. and Dowdy, R. H. (1984c) Evaluating soil productivity in relation to soil erosion. In *Quantification of the effect of erosion on soil productivity in the international context* (F. R. Rijsberman and M. G. Wolman, eds), pp. 53–69. Delft, The Netherlands: Delft Hydraulics Laboratory.

Pimental, D., Allen, J., Beers, A., Guirand, L., Linder, R., McLaughlin, P., Meer, B., Mussonds, D., Perdue, D., Poisson, S., Siebert, S., Stoner, K., Salazar, R. and Hawkins, A. (1987) World agriculture and soil erosion: erosion threatens world food production. *BioScience*, 37(4), 277–283.

Proffitt, A. P. B., Berliner, P. R. and Oosterhuis, D. M. (1985) A comparative study of root distribution and water extraction efficiency by wheat grown under high and low frequency irrigation. *Agronomy Journal*, 77(5), 655–662.

Rickman, R. W., Allmaras, R. R. and Ramig, R. E. (1977) Root-sink descriptions of water supply to dryland wheat. *Agronomy Journal*, 70(5), 723–727.

Rijsberman, R. F. and Wolman, M. G. (1985) Effect of erosion on soil productivity: an international comparison. *Journal of Soil and Water Conservation*, 40(4), 349–354.

Rust, R. H. and Hanson, L. D. (1975) *Crop equivalent rating guide for soils of Minnesota*. St Paul: University of Minnesota and Agricultural Experiment Station, Miscellaneous Report No. 132.

Sandor, S. P. (1989) *A productivity index model for Montana soils*. Unpublished MSc thesis, Montana State University, Department of Plant and Soil Science.

Schweitzer, B. D. (1990) *Spring wheat yields on two contrasting aridic argiborolls in northcentral Montana*. Unpublished MSc thesis, Montana State University, Department of Plant and Soil Science.

Spratt, E. D. and McIver, R. N. (1972) Effects of topographical positions, soil test values, and fertilizer use on yields of wheat in a complex of black chernozemic and gleysolic soils. *Canadian Journal of Soil Science*, 52, 53–58.

Ulmer, M. G. and Patterson, D. D. (1988a) Crop yield interpretation for North Dakota: 1. Sequential sampling. *Soil Survey Horizons*, 29, 116–122.

Ulmer, M. G. and Patterson, D. D. (1988b) Crop yield interpretation for North Dakota: 3. Long-term empirical models. *Soil Survey Horizons*, 29, 132–141.

Williams, J. R. (1981) Soil erosion effects on soil productivity: a research perspective. *Journal of Soil and Water Conservation*, 36, 82–90.

Williams, J. R., Jones, C. A. and Dyke, P. T. (1984) A modeling approach to determining the relationship between erosion and soil productivity. *Transactions of the American Society of Agricultural Engineers*, 27(1), 129–144.

Wilson, J. P., Sandor, S. P. and Nielsen, G. A. (1991) Productivity Index model modified to estimate field variability of Montana small grain yields. *Soil Science Society of America Journal*, 55 (1), 228–234.

Winkelman, L. J., Graham, W. A. P., Larson, W. E., Pierce, F. J. and Dowdy, R. H. (1984) *A program to display the effects of erosion on soil productivity* (Revision 3). St Paul: University of Minnesota Department of Soil Science and USDA Agricultural Research Service.

(*Revised manuscript received 6 February 1992*)