Allocating C-S-R plant functional types: a soft approach to a hard problem

J. G. Hodgson, P. J. Wilson, R. Hunt, J. P. Grime and K. Thompson

Hodgson, J. G., Wilson, P. J., Hunt, R., Grime, J. P. and Thompson, K. 1999. Allocating C-S-R plant functional types: a soft approach to a hard problem. – Oikos 85: 282–294.

A long-term research programme, conducted mainly in northern England, has involved field surveys (1965–77), laboratory screening (1974–96), monitoring of permanent plots (1958 to date) and manipulative experiments (1987 to date). The so-called C-S-R classification of plant functional types developed from all this activity. Patterns of covariation among the traits used in the classification have recently been validated in this journal.

The C-S-R classification appears to be applicable to vegetation in general. It thus has considerable potential for interpreting and predicting vegetation and ecosystem properties on a world-wide scale. However, to realize this potential we need to develop simplified procedures to extrapolate the C-S-R system to the many species which have not been the subject of previous ecological investigation.

Here we describe a rapid method for attribution of C-S-R type and we test its accuracy in Britain by comparing it with an independent classification based upon more laborious procedures. The new method allocates a functional type to an unknown herbaceous subject using few, simple predictor variables. We have developed spreadsheets to perform all of the necessary calculations. These may be downloaded from the UCPE website at http://www.shef.ac.uk/uni/academic/N-Q/nuocpe, or obtained by direct application to the E-mail address ucpe@sheffield.ac.uk

J. G. Hodgson, P. J. Wilson, R. Hunt (correspondence), J. P. Grime and K. Thompson, Dept of Animal & Plant Sciences, Univ. of Sheffield, Sheffield, UK S10 2TN (r.hunt@sheffield.ac.uk).

Two key prerequisites for understanding the ecological processes causing floristic change are a simple, robust rulebase concerning vegetation processes and a large, functional, autecological database. C-S-R plant strategy theory can aid the diagnosis of many common instances of climate- or land-use impacts. For example, functional classifications of plants can help interpret and predict current trends in 'winners' and 'losers' (Thompson 1994) and can assist in understanding ecosystem properties such as resistance and resilience (Leps et al. 1982, MacGillivray et al. 1995).

Because the predicted pattern of covariation between the traits used in C-S-R classification has recently been validated by a fully objective method (Grime et al. 1997), the system is now ready for wider application. However, before we can explore the full potential of the C-S-R system we need to identify plant functional type objectively in many individual species. This paper aims to assist in this process by identifying a set of 'soft' predictor variables which can indicate the position of an individual species in relation to each axis of the C-S-R system. We have validated the predictions made by this new procedure against a specimen 'hard' database derived from the laboratory and from the field.

The C-S-R system (Grime 1974, 1977, 1979) involves the established (or adult) phase of plant life-histories. At the origin of the system lie two groups of external environmental factors (Fig. 1a), both of which vitally affect the performance of plants in the field. The first group, stress, consists of factors that place prior restric-

Accepted 10 August 1998 Copyright © OIKOS 1999 ISSN 0030-1299 Printed in Ireland – all rights reserved tions on plant production, such as shortages of light, water and (particularly) mineral nutrients, or sub-optimal temperatures. Factors in the second group, disturbance, cause the partial or total destruction of plant biomass after it has been formed. This group includes intensification factors such as grazing, trampling, mowing and ploughing, and also extreme climatic events such as wind-damage, frosting, drought, soil erosion and fire. Eutrophication takes the form of release from nutrient stress, and dereliction takes the form of release from disturbance.

Grime et al. (1997) established objectively that the species typically associated with three of the four possible permutations of environmental extremes possess distinct sets of traits which confer characteristic ecological behaviour: competitiveness in the case of low stress and low disturbance, stress-tolerance in the case of high stress and low disturbance, and ruderality in the case of low stress and high disturbance. The initials of these three 'primary' types give the C-S-R system its name. The fourth environmental contingency, that of high stress and high disturbance, does not support plant life at all.

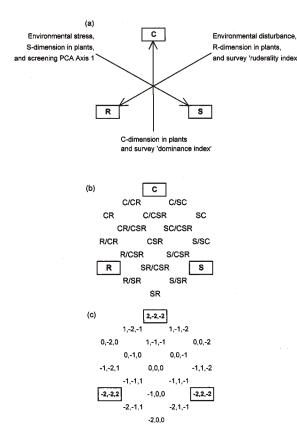


Fig. 1. The basic geography of the C-S-R classification of plant functional types. (a) The orientation of environmental factors, plant dimensions and screening and survey axes in C-S-R space; (b) the location of 19 plant functional types in C-S-R space; (c) the C,S,R triple coordinates for the locations shown in part (b) with ranges -2 to +2.

Intermediate types also exist within the C-S-R system, each exploiting a different intermediate combination of stress and disturbance. The position of any species (or, by aggregating and weighting its component species, that of any vegetation) can be displayed upon a triangular diagram (Fig. 1b). To facilitate quantitative work within C-S-R space, each position on this diagram can be given a three-part C,S,R coordinate on a radiating pattern (Fig. 1c). This coordinate represents the relative importance within the plant (or vegetation) of the three attributes competitiveness, stress-tolerance and ruderality. Because of its particular dimensionality, the C-S-R system is highly relevant to studies of land use (Hodgson 1989, 1991): for example, the C-dimension relates to abandonment (i.e. discontinued management), the S-dimension relates to eutrophication (i.e. release from nutrient stress) and the R-dimension relates to disturbance.

Against such a background, the aims of this paper are (1) to review the theory and practice of C-S-R classification; (2) to develop a simple new procedure for allocating C-S-R functional type from readily available predictor variables; (3) to validate the procedure using independent datasets and criteria; and (4) to present customized spreadsheets so that the procedure may be used by others to interpret and predict vegetation and ecosystem properties in other geographical regions.

Allocating C-S-R

General outline of the procedure

To achieve our aims we first define the methods that currently provide the best estimate, or 'gold standard' definition, of C-S-R type; second, we assemble a set of simple predictor variables derived from 'soft' (i.e. relatively undemanding) plant measurements; third, we build regression models which will predict C-S-R positional data from these simple variables; and fourth, we develop an allocation procedure in which positional data can be identified in terms of a known C-S-R functional type.

Defining the 'gold standard' for C-S-R functional type

In the original formulation of the C-S-R system (Grime 1974), species were ordinated in the C-dimension according to a 'competitive' or 'morphology' index. This was a composite of canopy height, lateral spread and litter accumulation. In the S-dimension, species were similarly ordinated according to maximum relative growth rate in the seedling phase (from work by Grime and Hunt 1975). No explicit ordination of the R-dimension was attempted, and this led to lack of definition in certain areas.

The C-S-R scheme then developed along more sophisticated lines, resulting in the publication of a dichotomous key (Grime 1986) by means of which species could be allocated to one of the three primary types (C, S, R), or to one of four secondary intermediates (CR, SR, SC or CSR). The resulting seven-type classification rested not only on the original criteria but also on others, such as flowering behaviour and leaf turnover. With the recognition of twelve further intermediate types at a tertiary level, this classification was then applied to 502 common British species by Grime et al. (1988). The resulting total of nineteen C-S-R functional types (specified here in Fig. 1b) has not since been subdivided further. The 1988 list of species, types and attributes was published in electronic form by Hodgson et al. (1995).

Though a simple method for allocation of functional type in the R-corner of C-S-R space was devised by Bogaard et al. (1998), no fully worked scheme for C-S-R allocation has been proposed since 1988. However, the results of the extensive, laboratory-based Integrated Screening Programme (Hendry and Grime 1993) have since become available (Grime et al. 1997) and certain of the simpler screening data have also been collected and incorporated into an extensive, but as vet unpublished, 'FIBS' database (Functional Interpretation of Botanical Surveys, J. G. Hodgson and others, in prep.). As a result, it is now possible to define 'gold standard' C-S-R criteria for large samples of British species by using field survey data to ordinate species in the C- and R-dimensions and by using screening data to ordinate in the S-dimension (see Fig. 1a). With this much hard information now to hand, it has also become feasible to develop simpler predictors of C-S-R dimensionality which do not involve complex and laborious investigations such as field surveys or laboratory screening.

The derivation of a 'gold standard' assay for the position of a species in the C-dimension is illustrated in Fig. 2 using UCPE field survey data from undisturbed, unshaded, non-skeletal habitats. The unit of recording in UCPE fieldwork is rooted frequency, the number of standardized quadrat subdivisions which are occupied by rooted specimens of the subject species. If the rooted frequency of subject x in any quadrat is defined as r_x and the total rooted frequency of all n species in the same quadrat is $\Sigma_1^n r_x$, then the relative rooted frequency of the subject species may be defined as $r_x/\Sigma_1^n r_x$. When relative rooted frequency is plotted against rooted frequency itself, using mean values obtained across many samples, the slope of the resulting relation is a valid indicator of the capacity of the subject species to dominate. For example, even when the weakly dominant grass Briza media occupies virtually all sectors in its 1-m² field record, it commonly accounts for only one-tenth of the total abundance of all species present. (All botanical nomenclature used in this paper follows

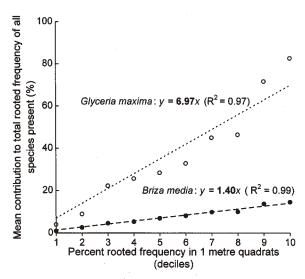


Fig. 2. The 'gold standard' for estimating the C-dimension within C-S-R space: the greater the 'Dominance Index' (= slope), the greater the competitiveness of the species (but see text for assumption).

Stace 1991.) The stand-forming grass Glyceria maxima, on the other hand, can dominate strongly, comprising about three-quarters of the total abundance in situations in which it appears throughout the quadrat. The slopes of the appropriate contribution/frequency relations can be computed for all species and used as a 'Dominance Index', the magnitude of which indicates the closeness of any given subject to the pure C-type.

However, we emphasize that competitiveness (the attribute to which the C-dimension refers) and dominance (the attribute which we have measured) do not always coincide. Competitiveness involves a high rate of resource capture and a rapid re-investment of these resources into organs for further resource-capture, such as leaves and roots. It also requires delayed flowering, which would otherwise sequester resources to the disadvantage of vegetative growth. Dominance simply indicates that a high biomass has been achieved relative to other species. No mechanism for dominance is implied by this term, and no distinction is drawn between large, fast-growing species and large, slow-growing but long-lived ones.

What we have quantified here is 'competitive dominance', in the sense of high resource capture allied to high allocation to large structures (Grime 1979). Competitiveness and dominance are thus assumed to be roughly equivalent within our predictor dataset. This concordance has been achieved by omitting vegetation types where species may become dominant without being highly competitive. On these grounds we excluded woody vegetation (which may be dominated by slow-growing, long-lived species) and disturbed sites, where ruderals, annuals and other ephemeral species may attain dominance 'by default' because potential peren-

nial competitors are removed by disturbance. Woody species (which are outside the scope of this paper) and monocarpic species were both omitted from the predictor regressions for the C-dimension: subsequent estimates of the C-dimension for ruderal species thus involved an extrapolation.

The 'gold standard' for the S-dimension comes from the laboratory and statistical work of Grime et al. (1997). Standardized procedures were used to measure 67 traits in 43 common species of the British flora and the most consistent patterns in the resulting data matrix were uncovered by means of multivariate analyses. These included principal components analyses (PCA), in which a strong 'Axis 1' emerged (loc. cit., Fig. 6) and coincided exactly with the alleged S-dimension of the C-S-R scheme (in cases where this was defined by nutrient stress). Degree of closeness of any species to the pure S-type can thus be defined by its score on this leading PCA axis. Grime et al. (1997) found that the composition of this axis indicated that the pure S-type is associated with low leaf palatability, relatively sustained yields under low nutrient regimes, high leaf tensile strength, low concentrations of major nutrients in leaves, and low decomposition rates and specific leaf areas. This suite of attributes occurred, in varying degree, in both monocotyledons and dicotyledons and appeared to reflect a trade-off between attributes conferring an ability for high rates of resource acquisition in productive habitats and those responsible for retention of resource capital in unproductive conditions.

For a 'gold standard' in the R-dimension we returned to the UCPE survey data. R-type species are classically associated with high levels of temporal and/or spatial disturbance, and also with low amounts of environmental stress (Fig. 1a). Monocarpic species, and (in temperate zones) ephemeroids with a vernal phenology, are particularly good indicators of this combination of conditions, so the association of such species with an unknown subject is a measure of that subject's own closeness to the pure R-type. For example, in the case of the upland sedge *Eriophorum vaginatum*, the mean rooted frequency of monocarpic species and vernals found co-occurring in the same 1-m² quadrat is only 0.6%, but in the grass *Alopecurus myosuroides*, an arable weed, this figure rises to 74.1%.

Selecting predictor variables from 'soft' tests

Productive, undisturbed habitats are colonized by C-type species, or 'competitors'. These are robust perennials of high potential growth rate, having a dense, rapidly expanding biomass above and below ground (Grime 1979). For this type, as with each of the others, our policy for selecting predictor variables was to accept only the most easily measured of the relevant attributes. Thus we ignored root characteristics as esti-

mators of plant size and used only shoot attributes (canopy height, lateral spread and leaf size). Data on canopy height were abstracted from Hodgson et al. (1995) and from the authors' unpublished databases. Classes indicative of degree of lateral spread were constructed by means of a modification of the method used by Hodgson et al. (1995) and involved some field measurements.

Protection against herbivory, and its consequent loss of captured mineral nutrients, is a consistent feature of the S-type. These are slow-growing, stress-tolerant species of chronically unproductive habitats (Grime 1979). As yet, we have no simple test that identifies palatability to vertebrates or invertebrates, or even one which represents degree of chemical or physical deterrence of potential predators. So we concentrated on attributes that measure growth rate and leaf longevity indirectly (and these are also helpful when predicting the C-dimension, see above). The 'soft' traits that can easily be measured and are specifically relevant to the S-dimension are leaf weight, specific leaf area and leaf dry matter content. All of these have been shown to vary along gradients of productivity (Al-Mufti et al. 1977, Givnish 1987, Reich et al. 1992, Garnier and Laurent 1994, Hunt and Cornelissen 1997a, b). The leaf measurements were carried out on material collected in the field from robust, well-grown plants. Where possible, the measurements represent an average result from three populations in separate, climatically contrasted locations within the 3000 km² surrounding Sheffield. In this region the climatic gradient extends from dry, warm, lowland areas (altitude 5 m; mean maximum July temperature 21°C; mean minimum January temperature 1°C; annual rainfall 565 mm) to cool, wet uplands (300 m; 18°C; 0°C; 1290 mm).

Species of type R, or ruderals, are characterized by the early onset of an often prolonged period of reproduction (Grime 1979). Accordingly, among the 'soft' attributes particularly suitable for identifying the R-dimension are the timing of the onset of the flowering period and its subsequent duration. This information was abstracted from Clapham et al. (1987). It would have been advantageous to have included some measurement of reproductive effort (e.g. seed or flower dry weight as a proportion of total shoot dry weight), but such data do not exist on the scale required. Another key attribute of many ruderals is a short life-span. However, as 10% of our surveyed species were annuals, 90% were polycarpic perennials and only 0.1% were monocarpic perennials (biennials), our dataset was insufficiently balanced for us to include this variable. Lateral spread, which includes life-span as a component, could not usefully be included in this context because plants of disturbed habitats occur at both ends of the range (e.g. Poa annua is short-lived and would score a low value in any classification, whereas Elytrigia repens is a perennial with extensive lateral spread and would score highly).

The seven 'soft' variables which we finally selected as predictors of the C-, S- and R-dimensions are listed and defined in Table 1. Four of these variables (Canopy-Height, FloweringPeriod, FloweringStart and Lateral-Spread) can normally be obtained from published sources or by direct observation. The remaining three variables (DryMatterContent, LeafDryWeight and SpecificLeafArea) require simple measurements to be made on well-grown leaf material that has been collected from the field. No growth experiments or chemical analyses are required anywhere in the procedure.

Three of the seven variables are categorical: Canopy-Height, FloweringStart and LateralSpread. Their categories were defined after extensive trial-and-error tests which also extended into the next (regression) stage of the procedure. The three criteria adopted were (a) to secure reasonable numbers of examples within each category; (b) to make each category as distinct from the others as possible; (c) to obtain credible predictions when the whole variable was used in the ensuing regressions.

Building predictor regressions

Having chosen a balanced and accessible set of predictor variables, their statistical properties were checked for suitability for multiple regression analysis. In conse-

quence, it was found necessary to transform LeafDry-Weight to natural logarithms (with the addition of a constant value to avoid negative values), and to square-root both DryMatterContent and SpecificLeafArea.

Multiple regression analysis was performed by SPSS for WindowsTM (Version 6.0). The 'gold standard' indices for the C-, S-, and R-dimensions, defined in the previous section, formed the dependent variables, as appropriate. In the case of each of the three dimensions, species were analysed in two groups. These consisted of grasses, sedges and rushes on the one hand, and the remaining (principally dicotyledonous) species on the other. The analyses were, where possible, carried out using a dataset which was broadly representative of the whole Sheffield flora (i.e. it aimed to include the commonest species in each major habitat). Large samples of species (46 to 167 in number) were thus available for the regressions in the C- and R-dimensions. In the case of the S-dimension, however, the number of species in each group was restricted to the 19 or 20 which were present in the balanced sample studied by Grime et al. (1997).

All seven predictor variables were offered as independent variables to all regressions. All predictor terms were also offered in the form of squares, but no interaction terms were constructed. Throughout, the *P*-value for inclusion of a term in a regression was 0.15, and for elimination, 0.10. This combination of probabilities was

Table 1. Definitions of the predictor variables used in Step 1 of the C-S-R allocation procedure (see the displayed box in the text). These variables are the inputs to the predictor regressions which are shown in Table 2.

Variable	Definition
CanopyHeight	Six-point classification: 1 1-49 mm 2 50-99 mm 3 100-299 mm 4 300-599 mm 5 600-999 mm 6 > 999 mm
DryMatterCon- tent	Mean of percent dry matter content in the largest, fully hydrated, fully expanded leaves (%)
Flowering Period	Normal duration of flowering period (months)
FloweringStart	Six-point classification: 1 First flowering in March or earlier 2 in April 3 in May 4 in June 5 in July 6 in August or later, or before leaves in spring
LateralSpread	Six-point classification: (in graminoids) (in non-graminoids) (in graminoids) (in non-graminoids) (in graminoids) (in non-graminoids) (in non-graminoids) (in non-graminoids) (in non-graminoids) 3 Compactly tufted about a single axis, no thickened rootstock Compactly tufted ramets appressed to each other at base Compactly tufted about a single axis, thickened rootstock present Shortly creeping, <40 mm between ramets Creeping, 40–79 mm between ramets Widely creeping, >79 mm between ramets
LeafDry Weight	Natural logarithm of mean dry weight in the largest, fully hydrated, fully expanded leaves (mg), plus 3
SpecificLeaf Area	Mean of area/dry weight quotient in the largest, fully hydrated, fully expanded leaves (mm ² /mg)

Table 2. The predictor regressions used in Step 2 of the C-S-R allocation procedure (see the displayed box in the text). These regressions accept as inputs the predictor variables shown in Table 1; their outputs, in turn, supply the transformation equations shown in Table 3.

Statistic	(a) Predictions for g	rasses, sedges and rusl	nes	(b) Predictions for other herbaceous species			
	C-dimension	S-dimension	R-dimension	C-dimension	S-dimension	R-dimension	
Total number of cases (species)	46	19	61	71	20	167	
Range of predicted variable	11.7 (dominance units)	112 (PCA axis units)	54 (ruderality units)	10.1 (dominance units)	110 (PCA axis units)	71.9 (ruderality units)	
Predictor variables selected (see Table 1)	(no constant) (constant) (CanopyHeight)^2 (CanopyHeight) (LateralSpread)^2 DryMatterConte SpecificLeafArea (LateralSpread)^2		(no constant) FloweringPeriod LeafDryWeight sqrt(SpecifcLeafArea)	(no constant) (CanopyHeight)^2 (LateralSoread)^2 (LeafDryWeight)^2	(constant) CanopyHeight DryMatterContent (LeafDryWeight)^2	(no constant) Dry Matter Content Flowering Period Flowerinc Start (Leaf Dry Weight) ^ 2 Specific Leaf Area	
Predictor coefficients (sequence as above)	0.1410 0.09061 -	54.6 -1.666 1.069 -2.732 1.722	2.518 -2.748 -5.370	- 0.09245 0.05631 0.01595	-39.52 -7.581 2.633 -0.3510	sqrt(SpecifcLeafArea)1.158 3.137 3.145 -0.0849 -1.193 11.40	
Predictor t-talues (sequence as above)	6.86 5.58	2.22 -4.63 1.83 -7.35 5.67	- 2.94 4.00 5.06	6.68 5.62 4.38	-3.04 -3.65 4.90 -2.64		
Percent variance in regression	92.5	94.5	72.9	91.7	71.9	73.4	
Regression F-value	286	82.3	55.7	263	18.1	79.2	
tandard error for prediction	1.11	8.76	10.0	1.01	15.5	15.1	
5% limit for prediction	2.23	18.7	20.0	2.03	32.8	29.8	
Range/95% limit	5.24	6.00	2.69	4.97	3.36	2.41	
Confidence in quintile	96.0%	97.8%	70.0%	94.7%	83.2%	64.6%	

Table 3. The transformation equations used in Step 3 of the C-S-R allocation procedure (see the displayed box in the text). These equations accept as inputs the outputs of the predictor regressions shown in Table 2. The outputs of the transformation equations supply Steps 4 and 5 of the C-S-R allocation procedure.

(a) For grasses, sedges and rushes C-coordinate = 0.839(raw C-dimension) - 2.5 S-coordinate = 0.0474(raw S-dimension) - 1.103 R-coordinate = 0.119(raw R-dimension) - 2.5

(b) For other herbaceous species C-coordinate = 0.839(raw C-dimension) - 2.5 S-coordinate = 0.0531(raw S-dimension) - 1.249 R-coordinate = 0.119(raw R-dimension) - 2.5

the result of extensive trial-and-error regressions. These were designed to secure a reasonable balance between parsimony and predictive power. The presence or otherwise of constant (intercept) terms was decided a priori: constants were excluded in the case of regressions in the C- and R-dimensions (because of the nature of the relationships being modelled, see the previous text and also Fig. 2), but were admitted (if significant) in the case of the S-dimension regressions.

Table 2 specifies in full the predictor regressions that were finally accepted. Both of the S-dimension regressions

It is assumed that the unknown subject exists within C-S-R space (Fig. 1a) at, or near to, one of the positions displayed in Fig. 1b. The procedure predicts the most likely C,S,R coordinates for the unknown subject. By matching these to the 'known' coordinates of previously recognized types (Fig. 1c), the procedure allocates a C-S-R identity to the unknown subject.

For each new subject the procedure involves gathering (and sometimes pre-processing) certain simple input data, solving three predictor regressions (one for each element within the C-S-R system), and post-processing the regression outputs. The steps to be followed are similar in principle for grasses, sedges and rushes on the one hand, and for the remaining herbaceous species on the other, though details differ between these two groups.

Step 1 (Data assembly). For all unknown subjects, obtain values of the six predictor variables, Canopy-Height, DryMatterContent, FloweringPeriod, Lateral-Spread, LeafDryWeight and SpecificLeafArea according to the definitions given in Table 1. For non-grasses, etc., obtain also a value for the predictor variable FloweringStart.

Step 2 (Regression). Use the predictor variables from Step 1 to solve predictor regressions for the C-, S- and R- dimensions. The regression equations are specified in Table 2a (for grasses, etc.) or in Table 2b (for other species). Be careful to comply with the squaring or

returned significant intercept terms. The simplest model (that for the C-dimension in grasses, etc.) had just two independent terms; the most complex (that for the R-dimension in 'other' species), had six terms. From a knowledge of the absolute range of each dependent variable, and of the 95% limit for its prediction by regression, it was possible to calculate the confidence with which quintile subdivisions of the dependent variable could be predicted (this being the level of precision necessary for the C-S-R allocation method to succeed, see Fig. 1c). These levels of confidence lay in the range 83–98% for the C- and S-dimensions, and 67–70% in the case of the R-dimension.

Allocating a C-S-R functional type to an unknown subject

From the predictor regressions it was possible to obtain the apparent C-, S- and R- coordinates of an unknown species and thence proceed to an allocation of C-S-R type. The method by which we did this is set out formally in the following box. This describes the whole procedure, from beginning to end, for obtaining an unknown C-S-R classification from an herbaceous species starting from values of six or seven predictor variables.

square-rooting of certain terms and note that one of the predictor variables (*LeafDryWeight*) requires ln-transformation and the addition of a constant.

Step 3 (Transformation). Take the regression outputs for the C-, S- and R-dimensions from Step 2 and transform them to raw predictions of decimal C,S,R coordinates by means of the appropriate function selected from Table 3a (for grasses, etc.) or in Table 3b (for other species).

Step. 4 (Adjustment). Take the raw C,S,R predictions from Step 3 and truncate any outliers such that no coordinate takes a value which is below -2.5 or above +2.5. Then round-off all coordinate values to one decimal place in the direction of zero.

Step 5 (Identification of C-S-R type). Take the adjusted C,S,R coordinates from Step 4 and compare them in turn with those of all of the recognized functional types. For each of the nineteen possible matches shown in Fig. 1c calculate a sum of squares (SS) by means of the formula $SS = (C_u - C_m)^2 + (S_u - S_m)^2 + (R_u - R_m)^2$. The C, S and R values used in this formula are either the adjusted predictions from Step 4 (for the unknown subject, with suffix u), or the integer coordinates for one of the possible matches (with suffix m). The minimum among the resulting nineteen values of SS points to the closest match between the unknown subject and a recognized functional type.

The procedure given in the preceding box is suitable for a manual solution of the C-S-R problem when taken in conjunction with information supplied in Tables 1–3. However, we have also developed two customized spreadsheets to carry out the whole procedure (written in Microsoft[®] Excel for WindowsTM, Version 5.0). One spreadsheet is for grasses, etc., and the other

is for other herbaceous species. Together, they incorporate all of the methods and information necessary for an automated prediction of an unknown functional type from inputted predictor variables. Fig. 3 illustrates one of these spreadsheets in use in the case of an unknown dicotyledonous subject. It also serves as a worked example of the whole procedure. Working

For NON-GRASSE		Fallopia japonica (test subject)
Fill in the red boxe		ier (optional, above) and predictor values (required, below)
CanopyHeight	2000	(millimetres maximum)
DryMatterContent	26	(percent in fully-expanded leaves)
FloweringPeriod	2	(months in duration)
LateralSpread	6	(special six-point classification, see text)
LeafDryWeight	915	(mg per fully-expanded leaf)
SpecificLeafArea	15	(square mm per mg dry weight in fully-expanded leaves)
FloweringStart	6	(special six-point classification, see text)
Predicted type is:	С	based upon the above information
	C	S R
with coordinates:	2	-2 -2
		shows the intermediate calculations that led to this prediction
Processed input of		(do (4 4)
CanopyHeight	6	(now classified as 1-6)
DryMatterContent	5.1	(now square root of the original value)
FloweringPeriod	2	(as original value)
LateralSpread	6	(as original classification)
LeafDryWeight	9.82	(now natural log of original value, plus 3)
SpecificLeafArea	3.87	(now square root of the original value)
FloweringStart	6	(as original classification)
Regression predic	ctions of	raw C-S-R dimensions using processed input data
Raw C-dimension	6.893	('dominance index' units)
Raw S-dimension	-50.388	(PCA axis units)
Raw R-dimension	13.108	('ruderality index' units)
Raw C-S-R dimer	nsions co	nverted to raw decimal C-S-R coordinates
С	3.283	decimal coordinate)
S	-1.427	(decimal coordinate)
R	-0.940	(decimal coordinate)
Correction of raw	decimal	C-S-R coordinates
(a) Adjusted for h	igh outlie	ers
C	2.500	(decimal coordinate)
S	-1.427	(decimal coordinate)
R	-0.940	(decimal coordinate)
(b) Adjusted for I		_
С	2.500	(decimal coordinate)
S	-1.427	(decimal coordinate)
R	-0.940	(decimal coordinate)
	L	owards zero, with one decimal place
С	2.5	(decimal coordinate)
S	-1.4	(decimal coordinate)
R	-0.9	(decimal coordinate)
••		

Fig. 3. A facsimile of one of the spreadsheets for allocating C-S-R plant functional type. The functional type of the tall herb *Fallopia japonica* has been correctly identified as being pure type C.

Fig. 3. (Continued)

Identification of closest valid combination of coordinates

lugiiuiicau	OII OI CI	OSEST Val	iu comi	ination of cool	ulliates
Type	С	S	R	Variance	
C	2	-2	-2	1.82	Minimum variance = 1.82
C/CR	1	-2	-1	2.62	at position in list = 1
C/SC	1	-1	-2	3.62	
CR	0	-2	0	7.42	
C/CSR	1	-1	-1	2.42	
SC	0	0	-2	9.42	
CR/CSR	0	-1	0	7.22	
SC/CSR	0	.0	-1	8.22	
R/CR	-1	-2	1	16.22	Mean departure 0.832
CSR	0	0	0	9.02	
S/SC	-1	1	-2	19.22	
R/CSR	-1	-1	1	16.02	Predicted functional type
S/CSR	-1	1	-1	18.02	
R	-2	-2	2	29.02	С
SR/CSR	-1	0	0	15.02	
S .	-2	2	-2	33.02	coordinates
R/SR	-2	-1	1	24.02	C S R
S/SR	-2	1	-1	26.02	2 -2 -2
SR	-2	0	0	23.02	

copies of both spreadsheets may be downloaded from the UCPE website at http://www.shef.ac.uk/uni/academic/N-Q/nuocpe, or may be obtained by direct application to the authors at their E-mail address, ucpe@sheffield.ac.uk

Validation

Results

The new allocation procedure has been tested on two large, independent datasets. Further samples, comprising 110 monocotyledonous and 383 dicotyledonous species, were obtained from the FIBS database (as outlined earlier). None of these species had been used in the construction of the predictor regressions, but all of them had a 'known' or 'correct' allocation to C-S-R type by the method of by Grime et al. (1988). Values of the necessary predictor variables were also known.

When the procedure was applied to the validating datasets (Fig. 4), the great majority of allocations (76% over-all) either fell exactly upon the correct type or upon one of its immediate neighbours in C-S-R space, as shown in Fig. 1b (i.e. the 'nodal error' referred to in Fig. 4 was either zero or one). Instances of progressively more serious errors in allocation declined sharply in frequency, both for monocotyledonous and for dicotyledonous species.

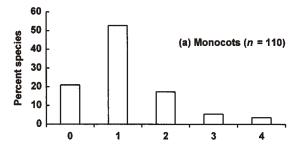
Validation criteria

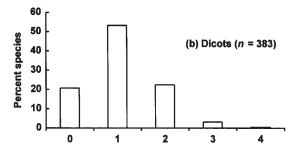
To gauge the efficacy of the new procedure we turned to the field of medicine. There, research workers and clinicians frequently face situations in which exact, but laborious, 'hard' tests for a particular condition co-exist with 'soft' tests, which are (possibly) less exact but generally simpler to perform. For example, Greenhalgh (1997) contrasted the exact, blood-based tests that are commonly required for the definitive diagnosis of diabetes with the simple dipstick urine test which is commonly used in preliminary investigations. She went on to summarize the validation criteria (shown here in Table 4) by which simple tests may be compared with their relevant 'gold standards', thereby indicating the confidence that both physicians and patients may place in the former.

In the case of C-S-R allocation, the diagnostic problem is very similar. We may either allocate or not allocate a particular functional type to a particular unknown subject, and we may either be right or wrong in doing so. The four contingencies which thus arise are displayed in Table 4 in the form of instances a to d. By variously manipulating the observed frequencies of such instances in a comparison of 'gold standard' and 'soft' tests, the validation criteria defined in the footnote to Table 4 may be derived.

We used these medical criteria for interpreting our validation procedure, both for whole groups of species (Table 5) and for individual functional types (Table 6). Throughout, we took the term a to be the frequency of allocation to the 'correct' type (defined as being either to the exact type or to an immediate neighbour); the term b to be the frequency of allocation to an 'incorrect' type (i.e. neither to the exact type nor to an immediate neighbour); the term c to be the frequency of non-allocation to a correct type; and the term d to be the frequency of non-allocation to an incorrect type.

The diagnostic criteria presented in Table 5 suggest that our allocation procedure has been broadly successful. Across nearly five hundred species, 96% of allocations were 'correct' ones (within the definition adopted).





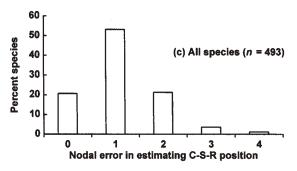


Fig. 4. Frequency diagrams showing the accuracy with which the new procedure predicts C-S-R functional type in two large, independent samples of herbaceous species. The nodal error is the deviation between the correct position of the species in C-S-R space (Fig. 1b) and the position given by the new procedure.

The procedure appears to be equally successful with monocotyledonous as with dicotyledonous species. It is better at excluding incorrect types than it is at allocating correct ones, but even in the latter respect its over-all sensitivity is 74%. The likelihood ratios are heavily biased in the desired directions.

With respect to the individual functional types (Table 6), the allocation procedure was particularly successful in the region of the pure C-type, where accuracies at or near to 100% were common. However, the distribution across the whole of C-S-R space is also acceptably uniform. The few places where lower accuracies appeared (85–95% in types SR/CSR and R/CSR) are those areas which are known to contain relatively few species, so this result may be an artefact. Again, the procedure appears to be equally applicable to mono-

Table 4. Definitions of certain diagnostic criteria adapted from medical research (see Greenhalgh 1997). These criteria are used to validate the new C-S-R allocation procedure in Tables 5–7.

	Result of 'gold standard' test					
Result of 'soft' test	Positive identification $(a+c)$	Negative identification $(b+d)$				
Test true $(a+b)$ Test false $(c+d)$	True positive (a) False negative (c)	False positive (b) True negative (d)				

Sensitivity (true positive rate) = a/(a+c); Specificity (true negative rate) = d/(b+d); Positive predictive value = a/(a+b); Negative predictive value = d/(c+d); Accuracy = (a+d)/(a+b+c+d); Likelihood ratio of positive test = (sensitivity)/(1-specificity); likelihood rate of negative test = (1-sensitivity)/((1-specificity)); (specificity).

cotyledonous and to dicotyledonous subjects, and also provides outcomes of similar sensitivity across a complete range of inland habitats in northern Britain (Table 7).

Usefulness of the procedure

Caveats

Despite the apparent success of the new procedure we could improve its ecological accuracy in several ways. For example, there are two major problems relating to the calculation of gold standards.

Table 5. Validation of the C-S-R allocation procedure: a general summary of diagnostic criteria for two large, independent samples of herbaceous species (see Table 4 for definitions of criteria).

Monocots

Dicots

All spp

	(n = 110)	(n = 383)	(n = 493)
Sensitivity How good is the test a	74% t picking up th	74% he correct ty	74% pe?
Specificity How good is the test a	98% t correctly exc	98% cluding types	98%
Positive predictive value	78%	75%	76%
<i>If a type is identified w</i> correct?	nai is ine pro	υαυππη τηατ	inis is
Negative predictive value If a type is excluded w. correct?	98% hat is the prob	98% bability that	98% this is
Accuracy What proportion of all	96% tests have giv	96% en a correct	96% result?
Likelihood ratio of positive test How much more likely untrue?	34 is an identific	66 ation to be	59 true than
Likelihood ratio of negative test How much less likely i. true?	0.34 s an exclusion	0.25 to be true to	0.27

Table 6. Validation of the C-S-R allocation procedure: diagnostic criteria for individual functional types. Abbreviations used: Sens(itivity), Spec(ificity), PV+ (Positive predictive value), PV- (Negative predictive value), and Acc(uracy). (See Table 4 for definitions of criteria.)

C-S-R type	Grass	es, sedg	es and r	ushes (n	= 110)	Other species $(n = 383)$				All species $(n = 493)$					
	Sens	Spec	PV+	PV-	Acc	Sens	Spec	PV+	PV-	Acc	Sens	Spec	PV+	PV-	Acc
C	0.50	1.00	1.00	0.99	0.99	0.67	1.00	1.00	0.99	0.99	0.63	1.00	1.00	0.99	0.99
C/CR	0.50	1.00	1.00	0.99	0.99	0.92	1.00	0.92	1.00	0.99	0.83	1.00	0.94	0.99	0.99
C/CSR	n/a	n/a	n/a	n/a	n/a	0.72	1.00	0.95	0.96	0.96	n/a	n/a	n/a	n/a	n/a
C/SC	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
CR	0.67	0.93	0.40	0.98	0.92	0.54	0.96	0.66	0.93	0.90	0.57	0.95	0.60	0.94	0.90
CR/CSR	0.33	0.99	0.50	0.97	0.96	0.67	0.97	0.67	0.97	0.95	0.59	0.98	0.63	0.97	0.95
CSR \	0.92	0.98	0.92	0.98	0.97	0.85	0.96	0.83	0.97	0.94	0.86	0.97	0.85	0.97	0.95
R	0.50	1.00	1.00	0.99	0.99	0.58	1.00	0.95	0.95	0.95	0.56	1.00	0.96	0.96	0.96
R/CR	0.50	1.00	1.00	0.98	0.98	0.70	0.99	0.91	0.96	0.95	0.66	0.99	0.93	0.96	0.96
R/CSR	0.50	1.00	1.00	0.99	0.99	0.78	0.92	0.53	0.97	0.90	0.72	0.94	0.63	0.98	0.92
R/SR	1.00	0.99	0.75	1.00	0.99	0.88	0.98	0.79	0.99	0.98	0.91	0.98	0.78	0.99	0.98
S	0.77	0.98	0.91	0.94	0.94	0.62	0.99	0.87	0.97	0.97	0.65	0.99	0.88	0.97	0.96
S/CSR	0.73	1.00	1.00	0.89	0.92	0.79	0.99	0.88	0.97	0.96	0.78	0.99	0.91	0.96	0.95
S/SC	0.57	1.00	1.00	0.95	0.96	0.67	0.98	0.29	1.00	0.98	0.65	0.99	0.45	0.99	0.97
S/SR	0.90	1.00	1.00	0.98	0.98	0.67	0.99	0.67	0.99	0.98	0.72	0.99	0.74	0.99	0.98
SC	1.00	0.93	0.50	1.00	0.93	0.75	0.99	0.60	0.99	0.98	0.81	0.97	0.58	1.00	0.97
SC/CSR	1.00	0.97	0.50	1.00	0.97	0.80	0.96	0.57	0.99	0.95	0.84	0.96	0.56	0.99	0.96
SR	0.73	0.93	0.62	0.96	0.90	0.87	0.97	0.80	0.98	0.96	0.84	0.96	0.76	0.98	0.95
SR/CSR	0.00	0.89	0.00	0.96	0.85	0.89	0.91	0.31	0.99	0.91	0.69	0.91	0.24	0.99	0.90
Means	0.67	0.98	0.78	0.98	0.96	0.76	0.98	0.75	0.98	0.96	0.74	0.98	0.75	0.98	0.96

First, the method described for the C-dimension in Fig. 2 is sensitive to species density. Subordinates in species-poor, tall-herb communities (e.g. *Poa trivialis*), and plants of acidic habitats which also contain few species (e.g. *Deschampsia flexuosa*), have higher-than-expected estimates of the C-coordinate. Because such species represent only a small minority in our sample, our values are close to expectation over-all when the predictor regressions are applied. However, this may not be the case elsewhere.

The second problem, which relates to the gold standard for the R-dimension, is more serious. The prediction of the R-coordinate is dominated by the functional attributes, and particularly the flowering period, of the largest group of ruderals, comprising annual species of arable and waste land. Other R-type species, often associated with habitats of temporally predictable disturbance, may include vernal geophytes and monocarpic perennials, and these have a much shorter flowering period. This often leads to their predicted R-coordinates being misleadingly small.

In addition, the leaf characters we used as predictor variables need a thorough re-evaluation. Interspecific comparisons of specific leaf area and dry matter content are not strictly valid unless all comparisons are made between functionally equivalent structures. Unfortunately this is not always so here: there is an unequal allocation of tissue to mechanical support. Many species have separate supporting structures in the form of stems and petioles and these are rightly excluded from estimates of SLA and dry matter content. In some species, however, the leaf is a wholly self-supporting structure. In such cases its necessarily greater structural component

reduces the apparent value of its SLA and increases that of its dry matter content. Hence, species with narrow leaves tend to have low SLA and high dry matter content irrespective of their habitat or true functional type.

Other problems exist within particular groups of species: (1) those for which the stem is the major photosynthetic organ (e.g. Juncus effusus); (2) succulents (e.g. Sedum acre) which, using the present predictor regressions, are classified as ruderals because of the very high water content of their thick, fleshy leaves; (3) slow-growing species of very shaded habitats with thin, watery leaves (e.g. Oxalis acetosella), and (4) halophytes, where a substantial correction for ash content is needed when calculating leaf attributes.

The majority of the 'mismatches' of functional type that became evident during the validation (i.e. instances having nodal errors of two or more units in Fig. 4) appeared to be caused by one or more of the problems outlined above. It is notable that in woodland, the habitat in which the C-S-R prediction performed with least sensitivity (Table 7), two of these major problems exist simultaneously (those relating to vernal geophytes and to extreme shade-tolerance). The resolution of these methodological difficulties across the whole dataset would increase the sensitivity of the C-S-R allocation procedure to a value well above the current level of 74%.

Conclusions

This paper provides both an examination of an ecological theory and a 'finished' ecological product. In the

Table 7. Validation of the C-S-R allocation procedure: the sensitivity criterion (see Table 4 for definition) within the seven major habitat types recognised by Grime et al. (1988).

	Monocots		Dicots		All species		
	Sensitivity	n	Sensitivity	n	Sensitivity	n	
(a) Among species	s used in one or more	predictor regre	essions				
All habitats	79 %	71	76%	190	77%	261	
Arable	n/a	2	80%	30	80%	32	
Grassland	84%	25	75%	40	78%	65 22 32	
Skeletal	n/a	2	85%	20	85%	22	
Spoil	n/a	2	77%	30 .	77%	32	
Wasteland	71%	7	81%	27	79 %	34	
Wetland	75%	24	70%	27	73%	51	
Woodland	78%	9	63%	16	68%	25	
(b) Among specie	s used in the validation	n					
All habitats	74%	110	74%	383	74%	493	
Arable	n/a	2	75%	48	75%	50	
Grassland	71%	17	69%	39	70%	56	
Skeletal	n/a	3	72%	46	72%	49	
Spoil	73%	11	78%	46	77%	57	
Ŵasteland	71%	14	76%	105	76%	119	
Wetland	82%	38	74%	54	77%	92	
Woodland	72%	25	69%	45	70%	70	
(c) Among all spe	cies combined						
All habitats	76%	181	75%	573	75%	754	
Arable	75%	4	77%	78	77%	82	
Grassland	79%	42	72%	79	74%	121	
Skeletal	60%	5	76%	66	75%	71	
Spoil	69%	13	78%	76	76%	89	
W asteland	71%	21	77%	132	76%	153	
Wetland	79%	62	73%	81	76%	143	
Woodland	74%	34	67%	61	69%	95	

n/a not available.

light of our opening review of the theory and practice of C-S-R classification it is valuable merely to have developed a procedure for allocating C-S-R functional type, even before any predictions are made, because it reveals why some of the functional characteristics used in the regression equations may be less useful than expected (e.g. specific leaf area in the estimation of the S-dimension, as outlined above).

Although the validation of our C-S-R allocation procedure involved a very large proportion of the herbaceous species of inland Britain, the customized spreadsheets that we have created promise to deliver still more when the approach is extended to other herbaceous systems world-wide. There is also no theoretical reason why the approach may not later be extended to deal with woody species.

Large datasets of the type pioneered by Ellenberg et al. (1992) are required if we are to interpret the impacts of changing land use and climate on ecosystems and landscapes. We hope that our example will encourage others to generate data appropriate for these objectives. The practical measurements demanded here do not require great field experience or ecological insight: technical competence and modest laboratory facilities are all that are necessary.

Though clearly in need of refinement, our procedure holds out the prospect of a global functional database being assembled for plant functional types. With collaborators in Argentina, Spain, Jordan and Greece, we are already heading towards this goal and we hope that the appearance of this new method of allocating C-S-R functional type will encourage others to join us.

Acknowledgements – This research was funded primarily by the Natural Environment Research Council. We gladly acknowledge the long-term assistance of S. R. Band and the collaboration of other members of the Unit of Comparative Plant Ecology.

References

Al-Mufti, M. M., Sydes, C. L., Furness, S. B., Grime, J. P. and Band, S. R. 1977. A quantitative analysis of shoot phenology and dominance in herbaceous vegetation. – J. Ecol. 65: 759–791.

Bogaard, A., Hodgson, J. G., Wilson, P. J. and Band, S. R. 1998. An index of weed size for assessing the soil productivity of ancient crop fields. – Veg. Hist. Archaeobot. 7: 17–22.

Clapham, A. R., Tutin, T. G. and Moore, D. M. 1987. Flora of the British Isles. – Cambridge Univ. Press, Cambridge.
Ellenberg, H., Weber, H. E., Düll, R., Wirth, V., Werner, W. and Paulißen, D. 1992. Zeigerwerte von Pflanzen in Mitteleuropa. – Scripta Geobot. 18: 1–258.

- Garnier, E. and Laurent, G. 1994. Leaf anatomy, specific mass and water content in congeneric annual and perennial grass species. – New Phytol. 128: 725-736.
- Givnish, T. J. 1987. Comparative studies of leaf form: assessing the relative roles of selective pressures and phylogenetic constraints. New Phytol. 106: 131-160.
- Greenhalgh, T. 1997. How to read a paper: papers that report diagnostic or screening tests. Br. Med. J. 315: 540-543.
- Grime, J. P. 1974. Vegetation classification by reference to strategies. Nature 250: 26-31.
- Grime, J. P. 1977. Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. Am. Nat. 111: 1169–1194.
- Grime, J. P. 1979. Plant strategies and vegetation processes. Wiley, Chichester.
- Grime, J. P. 1986. Manipulation of plant species and communities. In: Bradshaw, A. D., Goode, D. A. and Thorpe, E. (eds), Ecology and design in landscape. Blackwell Scientific, Oxford, pp. 175–194.
- Grime, J. P. and Hunt, R. 1975. Relative growth-rate: its range and adaptive significance in a local flora. J. Ecol. 63: 393-422.
- Grime, J. P., Hodgson, J. G. and Hunt, R. 1988. Comparative plant ecology: a functional approach to common British species. Unwin Hyman, London.
- Grime, J. P., Thompson, K., Hunt, R., Hodgson, J. G., Cornelissen, J. H. C. et al. 1997. Integrated screening validates primary axes of specialisation in plants. – Oikos 79: 259–281.
- Hendry, G. A. F. and Grime, J. P. (eds) 1993. Methods in comparative plant ecology: a laboratory manual. Chapman & Hall, London.
- Hodgson, J. G. 1989. The use of autecological information for selecting and managing plant materials used in habitat construction and the creation of species-rich vegeta-

- tion. In: Buckley G. P. (ed.), Habitat reconstruction, transplantation and repair. Belhaven Press, London, pp. 45–67.
- Hodgson, J. G. 1991. The use of ecological theory and autecological datasets in studies of endangered plant and animal species and communities. – Pirineos 138: 3–28.
- Hodgson, J. G., Grime, J. P., Hunt, R. and Thompson, K. 1995. The electronic comparative plant ecology. – Chapman & Hall, London.
- Hunt, R. and Cornelissen, J. H. C. 1997a. Components of relative growth rate and their interrelations in 59 temperate species. New Phytol. 135: 395-417.
- Hunt, R. and Cornelissen, J. H. C. 1997b. Physiology, allocation and growth rate: a re-examination of the Tilman model. Am. Nat. 150: 122–130.
- Leps, J., Osbornova-Kosinova, J. and Rejmanek, M. 1982.
 Community stability, complexity and species life-history.
 Vegetatio 50: 53-63.
- MacGillivray, C. W., Grime, J. P. and the ISP team. 1995. Testing predictions of resistance and resilience of vegetation subjected to extreme events. – Funct. Ecol. 9: 640– 649.
- Reich, P. B., Walters, M. B. and Ellsworth, D. S. 1992. Leaf life-span in relation to leaf, plant and stand characteristics among diverse ecosystems. – Ecol. Monogr. 62: 365–392.
- Stace, C. 1991. New flora of the British Isles. Cambridge Univ. Press, Cambridge.
- Thompson, K. 1994. Predicting the fate of temperate species in response to human disturbance and global change. –
 In: Boyle, T. J. B. and Boyle, C. E. B. (eds), NATO Advanced Research Workshop on biodiversity, temperate ecosystems and global change. Springer-Verlag, Berlin, pp. 61-76.