Europ. J. Agronomy 52 (2014) 210–217

Contents lists available at ScienceDirect

European Journal of Agronomy

journal homepage: www.elsevier.com/locate/eja

Prediction of winter wheat cultivar performance in Germany: At national, regional and location scale

Donghui Ma*, Hartmut Stützel

Institute of Biological Production Systems, Leibniz Universität Hannover, Herrenhäuser Str. 2, Hannover, Germany

ARTICLE INFO

Article history:
Received 5 November 2012
Received in revised form 8 September 2013
Accepted 9 September 2013

Keywords:
Cultivar recommendation
Value testing
tBLUP
Markov chain
State probability
Transition probability

ABSTRACT

Winter wheat cultivar recommendation is usually based on the cultivar performance observed in post-registration trials. In Germany, official recommendations are based on state cultivar trials, which are conducted individually by the federal states, usually over a period of three years. In each predefined winter wheat cultivation region a subset of registered cultivars is tested. The recommendation in a particular region is mainly based on the yields from trials on several locations in this region. Practically, the farmer's interest is a prediction of the yielding ability of cultivars on his own farm in the following growing season. This prediction can be made based on data from different scales, and with one year or multiple-year data. Here, we evaluated the prediction ability with the data from national, regional and location scales per se, and tried to find the optimal information source (scale and number of years) to predict the relative yield of a specific cultivar for a specific location. For this purpose, data from the country wide value testing trials from 1991 to 2001 carried out by the Federal Office of Plant Varieties (Bundessortenamt) were used. Winter wheat cultivation regions were adopted according to the German convention which gives the chance of further dividing the data into regional subgroups. The results of the analyses indicate that for a given location, the two years regional data have the highest predictive power for superior cultivars. Two years' data from that specific location give the highest predictive power for intermediate and inferior cultivars. In general, the predictive power of single year data is much lower than of two years data. The results confirm the merit of the definition of different cultivation regions. By proper definition of regions, the multiple year data collected within the region have high predictive power for the cultivar performance for the locations within that region.

1. Introduction

National and regional cultivar trials are conducted to help growers select cultivars for the next growing season. In the current German winter wheat recommendation trial system, the whole production area is subdivided into 23 different soil-climate-regions (SCR, Roßberg et al., 2007; Table 1). The clustering criteria for the growing regions are based on soil texture as well as the mean temperature and precipitation from March to August. Post-registration trials for sets of relevant cultivars are carried out in 3–7 locations yearly within each SCR.

The assumption underlying this testing system is that the mean yield estimated by sampling the test locations within one soil-climate-region is a good measure for the specific adaptedness of a cultivar to this region. The variance of yields of this specific cultivar across multiple years is considered an appropriate estimate of the temporal stability. A common approach of this adaptation analysis is the genotype × environment interaction (GEI) concept (Gauch and Zobel, 1996; Piepho, 1996, 1998). The existence of GEI complicates the decision-making and makes the further dissection of GEI for the multiple-location and multiple-years trials necessary. The composition of phenotypic variation has been discussed and the significant effect of genotype × environment interaction is empathized for plant breeding and plant adaptation (Comstock and Moll, 1963). Annicchiarico and Perenzin (1994) reported significant variances of both the genotype × location and genotype × year interactions, the former being almost 80% greater than the latter. This implies that a closer look should be taken at the location level. However, these interactions give no straightforward insight into how a particular cultivar behaves in relation to other cultivars at a given location from year to year. Yan and Rajcan (2003) studied the predictive power of single-year as well as multiple-year, multiple-location trial results for soybean cultivar selection in Canada, and were able to illustrate the performance change of each specific cultivar in the multiple-year trials. It was shown that a single-year, multiple-location trial had sufficient power for identifying genotypes that would perform well or poorly in the next year. Two to four years' data gave only slightly better predictions of next-year

* Corresponding author. Tel.: +49 511 7622635; fax: +49 511 7623606.
E-mail addresses: ma@gem.uni-hannover.de, ma@ingerland.de (D. Ma).

1161-0301/$ – see front matter
http://dx.doi.org/10.1016/j.eja.2013.09.005

Table 1
Soil-climate-regions of winter wheat in Germany and number of test locations for value testing from 1991 to 2001.

Region ID	Region name	Number of test locations
1	Marsch	4
2	Geest (Sand)	2
3	Östliches Hügelland SH (Lehm)	1
4	Diluviale Standorte nördl. Ostdeutschland	8
5	Diluviale Standorte südl. Ostdeutschland	2
6	Sandböden Nordwest	2
7	Köln-Aachener Bucht	3
8	Sandböden Nordhannover	1
10	Oderbruch	1
12	Lehmböden Südhannover	2
13	Lößstandorte der Ackerebene Mittel-/Ostd.	4
14	Lehmstandorte Nordwest	8
16	Höhenlagen Mitte/West	10
17	Verwitterungsstandorte Südost	2
18	Lößstandorte der Überganslagen Mittel-/Ostd.	7
19	Höhenlagen Südwest	2
20	Wärmelagen Südwest	7
21	Fränkische Platten	4
22	Tertiärhügelland, Bayerisches Gäu	11
23	Jura/Hügelland	3

performances than single-year data. This is similar to the result derived from the analysis of Gellner (1989), who concluded that one or two years of data were equally effective for predicting top-yielding cultivars as 3 years of yield data. All these previous analyses focused on the prediction of the cultivar performance for the test location based on the data from the same location. Although there have been several efforts to analyze the predictive power of location specific data, especially under controlled experimental conditions, in practice only partial information or even no information may be available for a target site, and in turn no hint exists about the predictive power of data from other locations for a specific farm location.

In this paper, we will systematically evaluate the predictive power of variety testing strategies stretching over spatial and temporal scales.

2. Materials and methods

2.1. Description of datasets

Yield data from the 1991 to 2001 winter wheat registration trails conducted by the German Federal Office of Plant Varieties (Bundessortenamt, Hannover) were used in this study. The registration trials included 84 test locations across 20 growing regions in Germany. There were on average 4 locations in each region. The minimum number of test locations was 1 which occurred in three regions; the maximum number of test locations was 11 in one region (Table 1). Geographically, the test locations were spread across the whole German winter wheat growing region and were distributed uniformly across the whole country.

In each year approximately 100 cultivars or breeding lines from different breeding companies were tested. To ensure the reliability of the analysis, this study included only cultivars which passed the registration trial and were later listed in the national list. A total of 76 cultivars were considered during this 11-year period. These cultivars belong to four quality groups, A, B, E, and C. There were 24 A cultivars, 33 B cultivars, 7 E cultivars and 12 C cultivars. A, B and E cultivars were grouped together in our analysis since their grain yield performances were similar. Group C cultivars were analyzed separately since they were superior in yield performance compared to other groups. The composition of genotypes tested varied between locations and years (Table 2). Only 4 cultivars were tested in all 11 years. For each pair of years, 6–60 common cultivars were tested (Table 3).

In each trial there were two cultivation intensity levels differing in N fertilization and chemical crop protection. Level 1 was a combination of extensive management practices, i.e. reduced nitrogen fertilization and crop protection; level 2 was a combination of intensive management practices, i.e. optimal N fertilization and crop protection to obtain disease-free crops. In this study, we concentrated mainly on the intensive level. The results from the extensive cultivation will be discussed briefly.

2.2. Statistical analysis of the datasets

Before 1996, lattice designs with 2 replications for each location were used. From 1996 on, an alpha design was adapted in the testing system (Meyer, 2000). Due to the different designs within this 11-year period, a two-stage analysis was used. Mean grain yield was first computed in accordance with the experimental design for each genotype at each location-year combination. The subsequent meta-analysis was based on the mean values from the first stage analysis.

For the analysis of the whole dataset, the following model was used: $Y_{ijk} = \mu + g_i + h_j + l_k + (hl)_{jk} + e_{ijk}$ where Y_{ijk} is the mean yield of cultivar i in year j at location k, μ is the grand mean, g_i is the main effect of cultivar i, h_j is the main effect of year j, l_k is the main effect of location k, and $(hl)_{jk}$ is the interaction between year j and location k. The term e_{ijk} is the residual associated with cultivar i in year j at location k. This error term also includes GEI (including genotype by location, genotype by year and genotype by location by year). The variance components were estimated using Restricted Maximum Likelihood (REML). All main effects were assumed random, best linear unbiased predictors (BLUP) were calculated for each main effect. The SAS Proc Mixed Procedure (SAS Institute, 1996) code from Yan et al. (2002) was used to obtain the variance components and the random effects.

The selection of locations for this analysis was done at three scales. The largest scale included all test locations from the whole of Germany. The second scale included locations within each growing region. At the lowest level the analysis was made for each location. There, the factor location was dropped from the model.

The multiple-years data were first analyzed across all years. For the purpose of prediction the analysis was confined to single-year and two-year analyses. When only annual data were analyzed, the year factor was dropped from the model. Usually, a cultivar was tested for only 3 years in this testing system. The design of the registration trial allowed the use of one and two-year data for the prediction in the 3rd year. To predict the next year's performance of the cultivars, only genotypes tested in the year in which the cultivar performance was to be predicted were included in the analysis. For example, if the prediction for year 1997 was based on data from 1995 to 1996, only genotypes present in 1997 would be included in the analysis.

We estimated the variance components for each random factor, which gives information on the relative importance of the source of variation as well as the fixed and random effects including grand mean, genotypic (cultivar) effects, year effects and location effects. To compare the BLUP for a given cultivar to the grand mean, the t-statistics of BLUP, in short as tBLUP (Yan et al., 2002), which is the ratio of BLUP to the associated prediction error, was calculated. Two one-sided tests were carried out for BLUP of each cultivar against the grand mean. At the 0.95 confidence level, a threshold of 1.67 was used for tBLUP. Based on this threshold, the cultivar was classified as superior ($t \geq 1.67$), inferior ($t \leq -1.67$), and intermediate ($-1.67 < t < 1.67$). After testing the BLUP for each cultivar against the grand mean using one-year or two-year data, the prediction was made using the results from the tBLUP test. First, each cultivar