The Journal of Agricultural Science

cambridge.org/ags

Organic system vs. conventional – a Bayesian analysis of Polish potato post-registration trials*

M. Przystalski and T. Lenartowicz

Research Centre for Cultivar Testing, 63-022 Słupia Wielka 34, Poland

Abstract

Interest in organic agriculture worldwide is growing and is mainly supported by a strong consumer interest. In the literature, a lot of attention has been paid to comparing organic and conventional systems, on studying the yield gap between the two systems and, how to reduce it. In the present work, based on the results from Polish organic and conventional series of field trials carried out in 2019–2021, organic and conventional systems were compared in terms of potato tuber yield. Moreover, we propose a Bayesian approach to the variety × environment × system data set and describe Bayesian counterparts of two stability measures. Using this methodology, we identify the most stable and highest tuber yielding varieties in the Polish potato organic and conventional series of field trials. It is shown that the tuber yield in the organic system was approx. 44% lower than the tuber yield in the conventional system. Moreover, varieties Tajfun and Otolia were the most stable and highest yielding varieties in the organic system, whereas in the conventional system, the variety Jurek was the most stable and highest yielding variety among the tested varieties. In the present work, the use of the Bayesian approach allowed us to calculate the probability that the mean of a given variety in given system exceeds the mean of control varieties in that system.

Crops and Soils Research Paper

*Mention of trade names or commercial products in this article is solely for the purpose of providing scientific information and does not imply recommendation or endorsement by the Research Centre for Cultivar Testing.

Cite this article: Przystalski M, Lenartowicz T (2023). Organic system vs. conventional – a Bayesian analysis of Polish potato post-registration trials. *The Journal of Agricultural Science* **161**, 97–108. https://doi.org/10.1017/S0021859623000084

Received: 6 September 2022
Revised: 25 November 2022
Accepted: 27 December 2022
First published online: 25 January 2023

Key words:
Comparison of two agronomic systems; hierarchical model; stability

Author for correspondence:
M. Przystalski, E-mail: marprzyst@gmail.com

Introduction

Interest in organic agriculture worldwide is growing, mainly supported by a strong consumer interest. This can measured by the value of retail sales of organic products, which grew to an impressive 44.2 billion euros in the European Union (EU) (Willer *et al.*, 2022). The market trend is still growing and grows faster than the organic farmland area. In 2020, in most of the EU-member states the organic farmland area has increased (Willer *et al.*, 2022). The biggest shares of the organic farmland area in the total utilized area are in Austria (26.5%), Estonia (22.4%), Sweden (20.4%) and Italy (16%) (Willer *et al.*, 2022). In Poland, the share is equal to 3.5%, which is 507 637 ha.

For this reason, a lot of attention has been paid to comparison of the organic and conventional systems. At the early stages, these comparisons were mainly made for cereals (Murphy *et al.*, 2007; Przystalski *et al.*, 2008; Hoagland, 2009; Reid *et al.*, 2011; Kirk *et al.*, 2012). Recently, the studies are focused on studying the yield gap between the two systems (de Ponti *et al.*, 2012; Ponisio *et al.*, 2015; Lesur-Dumoulin *et al.*, 2017) and how to reduce it (de Ponti *et al.*, 2012; Shah *et al.*, 2017; Schrama *et al.*, 2018). To compare the two agronomic systems Przystalski *et al.* (2008) used linear mixed models: they treated the effects of system, of environments and their interaction as fixed, while the genotype × system interaction effect was treated as random. Moreover, they only assumed that the genotypic values in the two systems can have different variances, and that they can be correlated. In the present work, we modified the model used in Przystalski *et al.* (2008), assuming that the effects of environments × system interaction are random. Furthermore, in our model, we assumed that the environmental values and the variety × environment interaction values in the two systems can have different variances and can be correlated.

In parallel to that study, in many EU countries, variety offices (the authorities in charge of variety testing) started an organic value-for-cultivation-and-use (VCU) variety testing for national listing of cereals (Pedersen, 2012). In Austria a VCU testing system of varieties for the organic system has existed since 2002. For example, in that country, it is possible to have varieties of winter wheat VCU-tested under organic conditions. In these trials the new varieties are compared to those varieties which are used in organic farming. Moreover, winter wheat varieties can be registered without conventional results. In Denmark, since 2006 it is possible to have a variety VCU tested with supplementary organic trials. In Poland, new varieties of important species are assessed prior to the registration in value-for-cultivation-and use (VCU) trials, and then in post-registration trials. Based on the results of post-registration trials a recommendation to the farmers is given. Since 2014, COBORU runs organic post-

© The Author(s), 2023. Published by Cambridge University Press

registration trials with potatoes, cereals and peas, first only in Węgrzce experimental station and from 2019 in nine additional locations. As in conventional post-registration trials for potatoes, the main characteristic in organic trials is tuber yield. Only stable and high-yielding varieties are recommended for cultivation.

It is believed that cultivation of crops in Poland under organic conditions, including potatoes, will grow in the coming years. For this reason, it is important to introduce stable and high yielding varieties to the cultivation. Usually, stability of agronomic traits is assessed in multi-environment trials. A common approach is to analyse these trials a using a two-stage approach (see, e.g., Flis *et al.*, 2014; Caliński *et al.*, 2017; Damesa *et al.*, 2017), where each combination of year and site is treated as an environment. In the literature, several stability measures have been described, for example, Shukla's stability variance (Shukla, 1972), regression on the environmental mean approach to assessments of stability (Finlay and Wilkinson, 1963; Eberhart and Russell, 1966; Digby, 1979), superiority stability coefficient (Lin and Binns, 1988) or additive main effects and multiplicative model (Gauch, 1992). Whereas the stability analysis is routinely performed for traits assessed in the conventional system, little attention has been paid in the literature to stability of traits in the organic system. Recently, Kucek *et al.* (2019) performed stability analysis for yield, protein, falling number and test weight of organically managed spring and winter wheat using the framework outlined by Annicchiarico (2002). In a different study, Rakszegi *et al.* (2016) used a genotype main effect and genotype-by-environment interaction (GGE) model (Yan and Kang, 2003) to assess the stability of organic and low-input wheat varieties. However, there is little work on how to assess stability of varieties in organic and conventional systems simultaneously. In the present study, we propose a solution to this problem.

In plant breeding programmes at the early stages (see e.g. Löschenberger *et al.*, 2008), the lines are often selected for both organic and conventional low input systems. In the present study, the three-way variety × environment × system data was analysed using a Bayesian approach, which may be useful for plant breeders and agronomists. One of the most useful characteristics of Bayesian statistics is making use of previous information, e.g. from previous agricultural trials. A subjective approach involves defining priors for unknown parameters according to personal experience and impression, recognizing that expert opinion is better than no knowledge. For example, one can use the knowledge from previous studies. For variance components such procedures have been described in Silva *et al.* (2013) or in Azevedo *et al.* (2022). Further, Bayesian analysis offers a possibility of calculating posterior distributions of new quantities which are functions of model parameters. Moreover, the Bayesian approach can alleviate problems associated with estimating complex models such as zero estimates of variance components (Theobald *et al.*, 2002). The use of the Bayesian approach in the context of agricultural field experiments is rare (Theobald *et al.*, 2002, 2006; Theobald and Talbot, 2002, 2004; Edwards and Jannink, 2006; Crossa *et al.*, 2011; Josse *et al.*, 2014; Orellana *et al.*, 2014; Edwards and Orellana, 2015; de Oliveira *et al.*, 2016; Bernardo *et al.*, 2018; Nascimento *et al.*, 2020; Przystalski and Lenartowicz, 2020) and is mainly focused on AMMI models (Crossa *et al.*, 2011; Josse *et al.*, 2014; Bernardo *et al.*, 2018) and GGE biplots (de Oliveira *et al.*, 2016). The Bayesian counterpart of the Finlay–Wilkinson model has been described by Lian and de los Campos (2016). Recently, de Oliveira *et al.* (2018) and Nascimento *et al.* (2020) described the Bayesian counterpart of the Eberhart and Russell model and its modifications. Przystalski and Lenartowicz (2020) obtained Bayesian counterparts of two stability measures described in Piepho (1999) by assuming different covariance matrices for the random vector of environment × variety interaction and modifying the list of random effects in the model. All these methods can be used to assess the stability of varieties in each individual agronomic system (either organic or conventional).

The aim of the current study was to compare organic and conventional systems in terms of potato tuber yield based on the results from Polish organic and conventional series of field trials carried out in 2019–2021. In addition, by using the predicted means for environment × system interaction from our model, we investigated the difference in yields between the two systems in the Polish potato post-registration trial system. Further, based on the posterior estimates of genotypic values from the Bayesian linear mixed model, the stability of varieties in both agronomic systems was assessed using the harmonic mean of relative performance of genotypic values (HMRPVG) method (see e.g. Resende, 2007; Colombari Filho *et al.*, 2013; Dias *et al.*, 2018; Bocianowski and Liersch, 2021) and the superiority stability coefficient (Lin and Binns, 1988). To the best our knowledge, this is the first study in which the stability of varieties was assessed simultaneously in both agronomic systems. Additionally, we calculated for each variety in a given system a probability that the variety mean in that system exceeds the mean of control varieties in that system. In general, in the frequentist approach, the probabilities that the yields of varieties fall below a certain critical level were first considered by Mead *et al.* (1986), and later modified by Eskridge (1990), Eskridge and Mumm (1992), and Piepho (1996, 1998, 2000). Eskridge and Mumm (1992) calculated the reliability of a test variety by taking into account normally distributed differences between the effects of the test variety and the check variety to leverage the probabilities from the cumulative function of a normal distribution, or by considering a nonparametric model. In the present study, following Dias *et al.* (2022), we calculated the probabilities by using MCMC samples from the Bayesian analysis. Using that methodology, we identify the most stable and highest yielding potato varieties for a Polish organic and conventional series of field trials conducted in the years 2019–2021. As in Edwards and Jannink (2006), Przystalski *et al.* (2008), Lenartowicz *et al.* (2020) and Przystalski and Lenartowicz (2020), the Bayesian linear mixed model was fitted under the assumption of heterogeneity of error variance.

Materials and methods

Data

The data set consists of potato field trials performed in the years 2019–2021 in two cropping systems: organic and conventional. In both agronomic systems, the trials were laid out in a randomized complete block design with three replicates. In each plot there were 60 potato seeds planted. The organic and conventional trials were conducted at experimental stations (sites) belonging to the Research Centre for Cultivar Testing (COBORU) (Table 1), which were located in different parts of Poland (Fig. 1).

However, only the Węgrzce station has a certificated organic field, where COBORU performs small ecological post-registration trials with potato, cereals and peas. For the remainder of the sites, the fields were organically managed.

https://doi.org/10.1017/S0021859623000084 Published online by Cambridge University Press

Table 1. Sites used in the 3-year organic and conventional variety trials conducted from 2019 to 2021

Site	Geographic co-ordinates		
	Latitude	Longitude	m a.s.l.
Krzyżewo	53°01′ N	22°46′ E	135
Lućmierz	52°12′ N	19°08′ E	132
Tarnów	50°35′ N	16°47′ E	300
Węgrzce	50°07′ N	19°59′ E	285

During the three years of study, in both agronomic cropping systems, there were in total 11 mid-early varieties tested. This maturity group is the largest group among the tested varieties in the Polish post-registration trials, and varieties belonging to this maturity group are most often planted by farmers in Poland. Since we were interested in assessing yielding stability, we used results for varieties which were tested for three years. The list of varieties used in the present study with their country of origin and registration year is given in Table 2.

In Polish potato post-registration trials, one of the analysed characteristics is tuber yield, which is observed on plots. The observed tuber yield is expressed in decitonnes per hectare (dt/ha).

Statistical estimation

To compare the potato tuber yield in organic and conventional systems a Bayesian hierarchical model was used. For clarity, throughout the paper by 'environment' we mean a combination of year and location, and by 'trial' we mean a combination of environment and system. For identification, environments (Table 3) and trials were numbered. In the data set, the organic trials were numbered from 1 to 12, whereas the conventional trials were numbered from 13 to 24.



Fig. 1. Map of Poland showing the locations of the experimental sites.

Table 2. List of varieties with their country of origin and registration year

No.	Variety	Country	Registration year
1	Jelly	Germany	2004
2	Jurek	Poland	2012
3	Laskara	Poland	2013
4	Mazur	Poland	2014
5	Otolia	Germany	2014
6	Satina	Germany	2000
7	Tajfun	Poland	2004

Let y_{jklr} be the potato tuber yield for the jth variety ($j = 1, \ldots, J$) in the kth environment ($k = 1, \ldots, K$) in the lth system ($l = 1, 2$ for organic and conventional systems, respectively) in the rth replicate ($r = 1, \ldots, R$). Then the model can be written as:

$$y_{jklr} = u_{kl} + v_{jl} + w_{jkl} + z_{klr} + e_{jklr}, \quad (1)$$

where u_{kl}, v_{jl}, w_{jkl}, z_{klr} and e_{jklr} denote the random effects of environment × system interaction, of variety × system interaction, of variety × environment × system interaction, of replicates nested within system and environment, and of errors, respectively. In using model (1) it may appear that we are fitting the model without the general means for both agronomic systems, but they are included in the hierarchical definition of the parameters and their joint distribution.

The model has three hierarchies or stages. In the first, it is assumed that observations are exchangeable samples from a normal distribution

$$y_{jklr}|u_{kl}, v_{jl}, w_{jkl}, z_{klr}, \sigma^2_{e(lk)} \sim N(u_{kl} + v_{jl} + w_{jkl} + z_{klr}, \sigma^2_{e(lk)}) \quad (2)$$

where $\sigma^2_{e(lk)}$ is the error variance of the trial in the lth system and kth environment, while '∼' means 'distributed as' or 'distributed independently as', according to the context.

In the second stage, the prior distributions are assigned to u_{kl}, v_{jl}, w_{jkl} and z_{klr}. Let $\boldsymbol{u}_k = [u_{k1}, u_{k2}]'$ ($k = 1, \ldots, K$) be the vector of environment × system interaction effects. We assume that $\boldsymbol{u}_k$ follow a two-dimensional normal distribution with mean vector $\boldsymbol{a}$ and covariance matrix Σ_u as

$$\boldsymbol{u}_k|\boldsymbol{a}, \Sigma_u \sim N_2(\boldsymbol{a}, \Sigma_u) \quad (3)$$

where $\boldsymbol{a} = [a_1, a_2]'$ is a vector of general means for both agronomic systems.

Table 3. Identification numbers for environments included in the analysis

Location	2019	2020	2021
Krzyżewo	1	5	9
Lućmierz	2	6	10
Tarnów	3	7	11
Węgrzce	4	8	12

https://doi.org/10.1017/S0021859623000084 Published online by Cambridge University Press

For the vector of variety × system interaction effects $\boldsymbol{v}_j = [v_{j1}, v_{j2}]'$ we assigned a two-dimensional normal distribution with mean vector of zeros and covariance matrix Σ_v

$$\boldsymbol{v}_j|\Sigma_v \sim N_2(0, \Sigma_v). \tag{4}$$

For the vector of variety × environment × system interaction effects $\boldsymbol{w}_{jk} = [w_{jk1}, w_{jk2}]'$ we assigned a two-dimensional normal distribution with mean vector of zeros and covariance matrix Σ_w

$$\boldsymbol{w}_{jk}|\Sigma_w \sim N_2(0, \Sigma_w). \tag{5}$$

For the effects of replicates nested within environment and system z_{klr} we assumed a priori that the effects of replicates nested within environments in organic system (z_{k1r}) and the effects of replicates nested within environments in conventional system (z_{k2r}) are independent, and we assigned normal distribution with zero means and as:

$$z_{klr}|\sigma_{z,l} \sim N(0, \sigma^2_{z,l})\, l = 1, 2 \tag{6}$$

For e_{jklr} we assigned a normal distribution with zero mean as:

$$e_{jklr}|\sigma^2_{e(lk)} \sim N(0, \sigma^2_{e(lk)}) \tag{7}$$

In the last stage, prior distributions are assigned for a_l ($l = 1, 2$), Σ_u, Σ_v, Σ_w, $\sigma_{z,l}$ ($l = 1, 2$) and error variances $\sigma^2_{e(lk)}$. For a_l we assigned a normal distribution with mean m_{a_l} and variance 1000 as:

$$a_l \sim N(m_{a_l}, 1000) \quad l = 1, 2 \tag{8}$$

For covariance matrix Σ_u, we assigned an inverse-Wishart distribution (IW) with ν_u degrees of belief and scale matrix S_u. For matrices Σ_v and $\Sigma_w$we assigned hierarchical a half-t prior (Huang and Wand, 2013) as

$$\Sigma_t \sim IW(\nu + 2 - 1,\ 2 \cdot \nu \cdot \Lambda_t)\ \lambda_{l,t} \sim Ga\left(0.5,\ \frac{1}{A^2_{l,t}}\right)$$
$$l = 1,\ 2;\ t = v,\ w \tag{9}$$

where $Ga(\cdot, \cdot)$ denotes Gamma distribution, Λ_t is diagonal matrix with l^{th} element $\lambda_{l,t}$, ν and $A_{l,t}$ ($l = 1, 2$; $t = v, w$) are positive numbers. In the present work we chose $\nu = 2$. For this particular choice, the marginal distribution of each correlation in Σ_t ($t = v, w$) is uniform on (-1, 1).

For $\sigma_{z,l}$ we assigned half-Cauchy (HC) distribution centred at zero with scale hyperparameter equal to 10 as:

$$\sigma_{z,l} \sim\ HC(0, 10)\ l = 1, 2 \tag{10}$$

For modelling variance components using hierarchical models this prior is recommended (see e.g. Gelman *et al.*, 2013).

Finally, for error variances $\sigma^2_{e(lk)}$ we assigned inverse-Gamma (IG) distribution with parameters α and β

$$\sigma^2_{e(lk)} \sim\ IG(\alpha, \beta) \tag{11}$$

We chose $\alpha = 1$ and $\beta = 0.001$ to obtain flat priors. This choice allows the variance components to be shrunk to very nearly zero, if this is warranted by the data.

For the data set used in the present study, we also fitted model (1) and the model described in Przystalski *et al.* (2008) using the REML algorithm implemented in Genstat (see Genstat code in the appendix).

By considering the posterior estimates of environmental means for each system ($\widehat{Env}_{kl}$), of variety ×system interaction effects ($\hat{v}_{jl}$) and of variety × environment × system interaction effects ($\hat{w}_{jkl}$), the predicted variety means for each system and environment were calculated as:

$$V_{jkl} = \widehat{Env}_{kl} + \hat{v}_{jl} + \hat{w}_{jkl} \tag{12}$$

Next, by using the posterior estimates of environmental means ($\widehat{Env}_{kl}$) and the predicted variety means for each system and environment (V_{jkl}), we calculated within a Bayesian framework the harmonic mean of the relative performance of genotypic values (HMRPVG) stability index (see e.g. Resende, 2007; Colombari Filho *et al.*, 2013; Dias *et al.*, 2018) as:

$$HMRPVG_{jl} = \frac{K}{\sum_{k=1}^{K} \frac{\widehat{Env}_{kl}}{V_{jkl}}} \tag{13}$$

Resende (2007) described a method of selecting for yield and stability based on the harmonic mean of genotypic values (HMVG) for each genotype tested in different environments. The lower the standard deviation of genotypic performance across environments, the greater is the harmonic mean of its genotypic values. Therefore, varieties with higher HMGV are those that have high yield and high stability. For adaptability, Resende (2007) used the relative performance of genotypic values (RPVG) across environments. To calculate this index, one first expresses the predicted genotypic values as proportion of the overall mean of each environment and next, the mean of these ratios is calculated for each genotype. The HMRPVG method combines the methods HMVG and RPVG, simultaneously, penalizing genotype instability, similarly to the superiority measure (Lin and Binns, 1988). At the same time, adaptability is expressed in the sense of responding to environmental improvement, by considering the proportions of the means of each genotype in each environment, compared with the overall means in these environments, similar to the method proposed by Annicchiarico (1992). Therefore, varieties with higher HMRPGV are those that have high adaptability and high stability simultaneously for the environments and systems evaluated in the current study.

For comparison, by considering the predicted variety means in each system and environment, we computed within a Bayesian framework the superiority stability coefficient (Lin and Binns, 1988) as:

$$P_{jl} = \frac{\sum_{k=1}^{K} (V_{jkl} - M_{kl})^2}{2 \cdot K} \tag{14}$$

where M_{kl} is the maximum response among all varieties in the kth environment and the lth system. Varieties with the smallest values of the superiority measure tend to have better yields and to be more stable.

https://doi.org/10.1017/S0021859623000084 Published online by Cambridge University Press

Finally, based on variety means for each system we calculated the probabilities that mean tuber yield in that system would be higher than the mean of control varieties, i.e.

$$Probab_{jl} = P(m_{jl} > M_{cl}|y) = \frac{1}{S}\sum_{s=1}^{S} I(m_{jl}^{s} > M_{cl}^{s}|y), \quad j = 1, \ldots, J; l = 1, 2 \tag{15}$$

where m_{jl} mean of the jth variety in lth system, M_{cl} is the mean of the control varieties in the lth system, $I(m_{jl}^{s} > M_{cl}^{s}|y)$ is an indicator function taking value 1 if the posterior sample m_{jl}^{s} is greater than M_{cl}^{s} and 0 otherwise.

Every year, the Research Institute of Organic Agriculture FiBL publishes a list of varieties recommended for organic cultivation in Switzerland. Two varieties from our list (varieties Jelly and Ototlia) were recommended for cultivation by FIBL (Dierauer *et al.*, 2021). For this reason, these two varieties were treated as control varieties in the organic system and we calculated the probability that the mean tuber yield of a given variety in this system would be higher than the mean of those two varieties ($Probab_{1,j1}$). In the conventional system, according to methodology used in COBORU, all varieties are treated as control varieties in the conventional post-registration variety trial system. For this reason, in the conventional system we calculated the probability that the mean tuber yield of a given variety in this system would be higher than the mean of all varieties ($Probab_{2,j2}$). For comparison we also calculated $Probab_2$ for varieties grown in organic system.

The data set was analysed using Markov Chain Monte Carlo as implemented in the R-package NIMBLE (de Valpine *et al.*, 2017, 2022). Three Markov chains were generated with different starting values. To improve convergence and mixing of the chains, the observed tuber yields were expressed in tonnes per hectare (t/ha). Joint analysis of organic and conventional trials is new in COBORU and no prior knowledge was available. For this reason, to apply the model to the data, we specified the prior means m_{a_l} of a_l ($l = 1, 2$) as empirical means for each system. In our model, as scale matrix S_u for the covariance Σ_u as we used a diagonal matrix of order two with elements equal to empirical variances of environmental means in each system divided by two. The few of degrees of belief for Σ_u and non-informative priors for Σ_v and Σ_w indicate lack of experience in jointly analysing organic and conventional field trials. Values defining the prior distributions of general mean and covariance matrices are summarized in Table 4.

Table 4. Values defining the prior distributions of general mean and covariance matrix Σ_u, Σ_v and Σ_w

Parameter	Estimate
m_{a_1}	28.34
m_{a_2}	50.38
$S_{u,11}$	100.175
$S_{u,22}$	81.97
ν_u	4
$A_{l,v}^{-2}$ ($l = 1, 2$)	0.05
$A_{l,w}^{-2}$ ($l = 1, 2$)	0.0005

The convergence of chains was examined by visual inspection of trace plots and using the Gelman and Rubin (1992) convergence diagnostic (potential scale reduction factor, $\hat{R}$) (see also Cowles and Carlin, 1996; Gelman *et al.*, 2013) implemented in the CODA R-package (Plummer *et al.*, 2006). The length of each Markov chain was selected in such a way that the $\hat{R}$ was less than 1.08 for each parameter in the model. For this reason, the length of each chain was set to 2 500 000 iterations with a burn-in period of 1 250 000 iterations and a thinning interval equal to 40. Using the function calculate widely applicable information criterion (WAIC) from the NIMBLE package we calculated the value of the WAIC for the model. For model (1), the estimates of parameters of interest from all three Markov chains were summarized using the summary function from MCMCvis R-package (Youngflesh, 2018).

Results

The data set was analysed using model (1) (see Supplement S1). The value of WAIC for model (1) was equal to 2471.9. Before looking at the estimates from model (1), we first examined the convergence of the random and fixed effects by visually inspecting the trace plots (Supplement S2). From the trace plots, one can observe that all the chains show good convergence. This was confirmed by the Gelman and Rubin tests. The values of point estimates of potential scale reduction factor (Point est.) for all fixed and random effects were equal to one (see Supplement S4). The evolution of Gelman and Rubin's shrink factor is shown in Supplement S3.

Figure 2 shows the highest posterior density intervals for the elements of the covariance matrices Σ_u, Σ_v and Σ_w.

For covariance $\Sigma_{v,12}$ zero belongs to the highest posterior density interval. This means that there is no evidence of association between varieties in organic and conventional system. For this reason, we modified model (1). In the new model, we replaced (2) by

$$v_{jl}|\sigma_{v,l} \sim N(0, \sigma_{v,l}^2)\ l = 1, 2;\ v = 1, \ldots, J \tag{16}$$

For $\sigma_{v,l}$ we assigned half-Cauchy (HC) distribution with parameters 0 and 10

$$\sigma_{v,l} \sim HC(0, 10)\ l = 1, 2 \tag{17}$$

The data set was re-analysed using the modified model (1) with different starting values for each chain (see Supplement S5). We used the same values defining the prior distributions of general mean and covariance matrices as in model (1) (Table 4). The value of WAIC for the new model (1) was equal to 2471.4. The WAIC values from the two models are statistically identical. Hence, following the principle of parsimony, we prefer to choose the simpler model.

Before looking at the estimates from the modified model (1), we first examined the convergence of the random and fixed effects by visually inspecting the trace plots (Supplement S6). From the trace plots, one can observe that all chains show good convergence. This was confirmed by the Gelman and Rubin tests. The values of point estimates of potential scale reduction factor (Point est.) for all fixed and random effects were equal to one (see Supplement S8). The evolution of the Gelman and Rubin's shrink factor is shown in Supplement S7.

https://doi.org/10.1017/S0021859623000084 Published online by Cambridge University Press

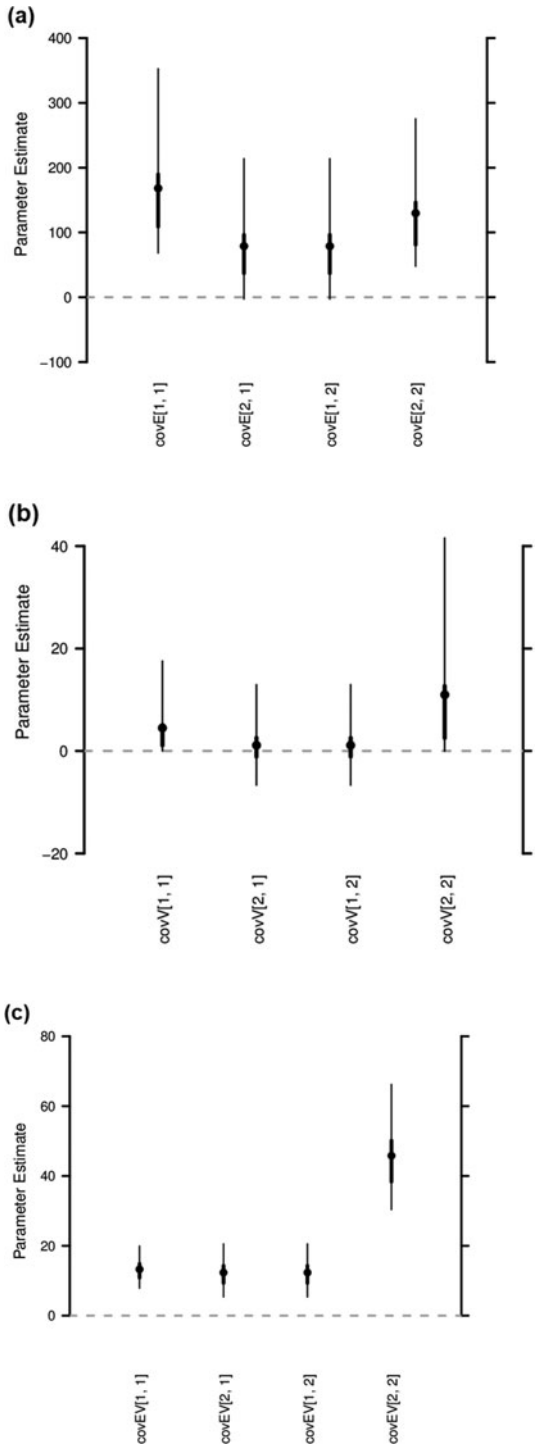


Fig. 2. The highest posterior density intervals for the elements of covariance matrices Σ_u (a), Σ_v (b) and Σ_w (c).

For the series of field trials, the analysis of the new model provided several estimated parameters and statistics (Tables 5 and 6). The posterior estimates of general means in each system, variance components and covariance matrices are reported in Table 5 (see also Supplement S8).

It can be noted that the mean value for the organic system was smaller than the mean value for non-organic system. The mean for the organic system was equal to 28.3 t/ha, which was 56.3% of the mean yield of the non-organic system. Next, one can observe that the environmental variance in organic conditions was higher than in conventional. The calculated environmental correlation between the two systems was equal to 0.53. Further, it can be noticed that for the rest of the variance components, the values in the organic system were always smaller than the values in the non-organic system. Finally, in the last two rows of Table 5, the mean error variances for both agronomic systems are reported. It can be seen that the mean error variance for the organic system was smaller than the mean error for the conventional system. A detailed inspection of the residual variances revealed that the biggest error variance in the organic system occurred in environments four and eleven. In the conventional system, the biggest error variance occurred in environments four, six, seven, eight and twelve.

Table 5. Posterior summaries (mean, 95% confidence interval) for hyperparameters in the modified model (1)

Parameter	Estimate	$CI_{0.95}$
α_1	28.3	(20.3–36.3)
α_2	50.3	(42.8–57.8)
$\Sigma_{u,11}$	186.59	(84.31–398.27)
$\Sigma_{u,12}$	89.50	(8.40–236.94)
$\Sigma_{u,22}$	144.48	(62.22–317.63)
$\sigma^2_{v,1}$	6.13	(0.63–23.01)
$\sigma^2_{v,2}$	15.66	(1.17–57.09)
$\Sigma_{w,11}$	13.52	(8.37–20.55)
$\Sigma_{u,12}$	12.66	(6.08–21.28)
$\Sigma_{u,22}$	46.64	(31.55–67.83)
$\sigma^2_{z,1}$	1.90	(0.78–3.96)
$\sigma^2_{z,2}$	2.30	(0.90–4.88)
$\bar{\sigma}^2_{e,1}$	9.12	–
$\bar{\sigma}^2_{e,2}$	11.58	–

$\bar{\sigma}^2_{e,1}$ – the mean error variance for organic experiments.
$\bar{\sigma}^2_{e,2}$ – the mean error variance for conventional experiments.

In Fig. 3 the estimated posterior environmental means for both agronomic systems is depicted (see also Supplement S8).

Firstly, it can be noticed that the highest values for the organic system were obtained for environments three, four and seven. Moreover, one can observe that in all environments the mean values for the organic system were smaller than the means for the non-organic system. However, a detailed inspection of environmental means revealed that, for environments ten and seven, the yield gap between the two systems was small in comparison to the rest of environments. For those two environments, the yield gaps were equal to 9 and 10.9 t/ha, respectively. On the other hand, the biggest yield gap between the two systems was observed for environment twelve and was equal to 48.3 t/ha.

Table 6 reports the estimated posterior variety means, values of HMRPVG index, superiority stability coefficient and the values of probabilities.

The estimated posterior variety means with their 95% confidence intervals (CI) are reported in columns three and four of Table 6. Among the tested varieties, variety Jurek had the highest tuber yield in the conventional system. In the organic system, variety Tajfun had the highest tuber yield, while in the conventional

https://doi.org/10.1017/S0021859623000084 Published online by Cambridge University Press

https://doi.org/10.1017/S0021859623000084 Published online by Cambridge University Press

Table 6. Posterior variety means, values of HMRPVG index and superiority stability index (*P*), and probabilities that mean tuber yield in a given system would be higher than the mean of control varieties

		Mean[t/ha]		HMRPVG		*P*		Probab	
System	Variety	Estimate	$CI_{0.95}$	Estimate	$CI_{0.95}$	Estimate	$CI_{0.95}$	$Probab_1$	$Probab_2$
Organic	Jelly	*25.2 [7]*	(17.11–33.25)	*0.84 [7]*	(0.756–0.923)	*53.66 [7]*	(38.38–70.06)	0.005	0.002
	Jurek	28.2 [6]	(20.20–36.19)	1.02 [3]	(0.928–1.127)	19.83 [4]	(10.21–29.39)	0.810	0.460
	Laskara	28.9 [3]	(20.95–36.88)	1.01 [4]	(0.912–1.119)	11.77 [3]	(6.43–18.61)	0.929	0.754
	Mazur	28.3 [5]	(20.37–36.28)	0.95 [6]	(0.857–1.056)	17.26 [5]	(12.29–23.14)	0.843	0.504
	Otolia	29.2 [2]	(21.16–37.09)	1.04 [2]	(0.947–1.154)	11.71 [2]	(6.41–18.69)	0.995	0.836
	Satina	28.1 [4]	(20.10–36.04)	0.98 [5]	(0.894–1.084)	32.36 [6]	(18.03–45.48)	0.791	0.406
	Tajfun	**30.3 [1]**	(22.30–38.27)	**1.05 [1]**	(0.950–1.171)	**4.67 [1]**	(2.98–7.79)	0.992	0.984
Conventional	Jelly	49.6 [5]	(41.93–57.22)	0.95 [6]	(0.890–1.026)	92.92 [6]	(72.21–116.22)	–	0.313
	Jurek	**54.7 [1]**	(46.83–62.68)	**1.11 [1]**	(1.041–1.200)	**18.43 [1]**	(10.65–27.82)	–	0.991
	Laskara	49.8 [6]	(42.21–57.38)	0.97 [4]	(0.908–1.049)	73.10 [5]	(54.54–94.48)	–	0.368
	Mazur	50.3 [4]	(42.75–57.93)	0.99 [3]	(0.920–1.063)	65.93 [4]	(49.85–83.68)	–	0.505
	Otolia	50.6 [3]	(43.07–58.27)	0.97 [5]	(0.906–1.049)	54.87 [3]	(44.09–67.74)	–	0.579
	Satina	*46.0 [7]*	(38.03–53.89)	*0.86 [7]*	(0.805–0.930)	*182.88 [7]*	(149.01–218.21)	–	0.009
	Tajfun	51.3 [2]	(43.67–58.90)	1.02 [2]	(0.955–1.102)	48.06 [2]	(33.81–64.33)	–	0.723

The numbers in square brackets denote the ranking of variety. Most favourable scores are highlighted in bold while least favourable scores are shown in italics.

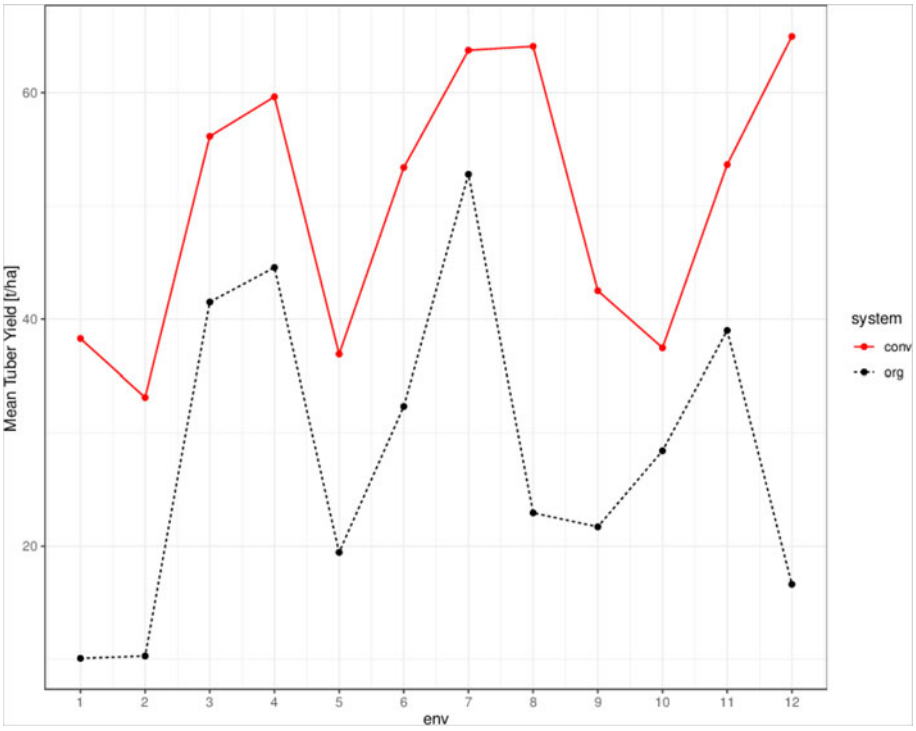


Fig. 3. Adaptive tuber yield response patterns across 12 environments in Poland for organic and conventional system.

system this variety was ranked second. Varieties Jelly and Otolia, recommended by FIBL, were ranked seventh and second, respectively, in the organic system. In the conventional system, these varieties were ranked fifth and third, respectively. The posterior distributions of variety means are shown in Supplement S6.

The posterior estimates of HMRPVG indices and superiority stability coefficients with their 95% CI are reported in columns five to eight of Table 6. The posterior distributions of the two stability measures are shown in Supplement S6. One can observe that, in the organic system, variety Tajfun had the highest value of HMRPVG index and the lowest value of the superiority stability coefficient. This means that this variety was the most stable among the tested varieties in the organic system. The second-best variety in terms of HMRPVG index and superiority stability coefficient was variety Otolia. On the other hand, variety Jelly was the most unstable in terms of both stability measures. In the conventional system, variety Jurek was the most stable variety in terms of both stability coefficients. The best variety in the organic system was the second most stable variety. Furthermore, variety Otolia was the fifth best variety in terms of HMRPVG index, while in the case of superiority stability coefficient it was ranked third. Finally, the Spearman rank correlations between the systems were equal to 0.5 and 0.571, for the HMRPVG index and the superiority stability coefficient, respectively. Moreover, for the organic system, the Spearman rank correlation between the two stability measures was equal to 0.821, whereas, for the conventional system, the value of the correlation coefficient was equal to 0.893.

In the last two columns of Table 6, the probabilities that mean tuber yield in a given system would be higher than the mean of control varieties are reported. It can be noticed that in the organic system the highest values of $Probab_1$ and $Probab_2$ were obtained for varieties Otolia and Tajfun. In the conventional system the highest value of $Probab_2$ was obtained for variety Jurek. Furthermore, it can be seen that variety Tajfun was the second best in the conventional system. On the other hand, variety Jelly was one of the worst varieties in both agronomic systems.

Discussion

In the present study, we simultaneously assessed the yield stability of potato varieties in organic and conventional systems. In the literature, stability of varieties is assessed in multi-environment trials, which are analysed using either a two-stage approach or one-stage approach (see e.g. Flis *et al.*, 2014; Caliński *et al.*, 2017; Damesa *et al.*, 2017; Lenartowicz *et al.*, 2020). In the first approach, one first analyses each trial separately and next, the variety means from all trials are analysed using a linear mixed model or AMMI model (Gauch, 1992). In the latter approach, the plot data from all trials in the series are analysed in a single stage.

In the present work, the stability analysis was performed on plot data and by using Bayesian hierarchical models. This approach was preferred for several reasons. First of all, the Bayesian approach allows one to incorporate the knowledge about likely values of average yields and variance components in a systematic way using proper priors. For example, one can use the knowledge from previous studies. For variance components such a procedure has been described in Silva *et al.* (2013), where the authors initially performed the analyses with non-informative priors for the inverse of variance components. As the same population and phenotypes were analysed by ANOVA and REML in a previous study, they used those results to construct informative priors. However, in order to validate the inclusion of these two studies in the meta-analysis, they used a homogeneity test based on the Q statistics (Hedges and Olkin, 1985). As variety offices or plant breeding companies have long-term trial data at their disposal, it is possible to update the hyperparameters and consequently update the knowledge regarding the variance parameters. Considering ten years of collection, Azevedo *et al.* (2022) used two types of prior information

https://doi.org/10.1017/S0021859623000084 Published online by Cambridge University Press

in the analysis in the *i*th year of study: (a) the hyperparameters were calculated by analysing the $(i - 1)$-th year, (b) the hyperparameters were calculated by joint analysis of the years 1, …, $i - 1$. However, according to Silva *et al.* (2013), given a trait, using estimates of additive and error variances from distinct experiments as prior information can provide biased estimates of the scale parameters' mean and variance. Azevedo *et al.* (2022) came to the same conclusion regarding their procedure II. To overcome this problem, Silva *et al.* (2013) proposed to re-parameterize the model using heritability and phenotypic variance. Although heritability is also a population parameter under a given experimental condition, heritability values are more consistent. However, the breeders know that for some traits the heritability is low, but for other traits the heritability is consistently high, and they can easily incorporate their knowledge about likely values of heritability. A similar concept is used in the FW package (Lian and de los Campos, 2016), in which prior values of variance components in the model are expressed as a proportion of phenotypic variance. In the present study, as a prior knowledge about environments, the empirical variances of environmental means in each system were used. A similar approach was used in Mathew *et al.* (2016): the authors applied a multi-trait animal model in the analysis of plant breeding trials and as the prior knowledge about parameters of interest they used phenotypic variances for each analysed trait. In a different study, Dias *et al.* (2022) used weakly-informative priors for standard deviations in their model and defined the known global hyperparameter as $\max(y) \times 10$. Moreover, we assumed *a priori* that the effects of replication nested in environments in an organic system and the effects of replication nested in environments in a conventional system are independent. The same assumption was used for e_{jlkr} errors. The main reason for this approach was that organic and conventional trials are managed differently and that experimental designs for organic and conventional trials in the same environment were generated independently. Further, Bayesian analysis offers a possibility of calculating posterior distributions of new quantities, which are functions of model parameters. In the present study, the posterior distributions for HMRPVG indices and superiority stability coefficients were calculated. Finally, by using the MCMC samples we were able to calculate the probability that the mean of a given variety for a given system exceeds the mean of control varieties in that system. A similar approach was used by Dias *et al.* (2022), where the authors used the probability methods of stability analysis in a Bayesian framework for unravelling genotype × environment interaction.

Based on the results from the current study, it can be concluded that the varieties Tajfun and Otolia were the highest yielding varieties among the tested varieties in the organic system. The lowest tuber yield in the organic system was found for variety Jelly. According to the HMRPVG index and superiority stability coefficient, the highest yielding varieties in the organic system were also the most stable. Thus, one can draw the conclusion that varieties Tajfun and Otolia should be promoted, and variety Jelly should not be recommended for cultivation in the organic system.

In the present paper, we also compared tuber yields in organic and conventional systems. The tuber yield in the organic system was approx. 44% lower than the tuber yield in the conventional system. This is agreement with the results obtained by Kazimierczak *et al.* (2019). However, for environments seven and ten the yield gap between the two systems was smaller. For these two environments, the tuber yields in organic trials were 9 and 11%, respectively, lower than the tuber yields in conventional trials. On the other hand, the biggest difference between the two systems was observed in environment twelve. In general, such big differences can be partly explained by the fact that the tested varieties were bred and selected for their ability to produce under the conventional system. Till now, no modern variety that is developed to produce high yields under organic conditions has been registered in the Polish National List. In the literature, depending on the crop, different yield gaps between the two systems are reported. In a meta-analytic study, Lesur-Dumoulin *et al.* (2017) showed that the yield in the organic system was on average 10 to 32% lower than the yields in the conventional system. For potato, the tuber yield in the organic system was on average 30% lower than the tuber yield in the conventional system (de Ponti *et al.*, 2012; Ponisio *et al.*, 2015). However, the yield gap can be still reduced. Schrama *et al.* (2018) have shown that the yield gap between organic and conventional systems declines with progressing time since conversion, which coincides, according to the authors, with enhanced N-input efficiency of organic compared to the other systems.

We conclude from the current study that there is no association between varieties in organic and conventional systems, since zero belonged to the highest posterior density interval for $\Sigma_{v,12}$. The estimated posterior genetic correlation between the two systems was approx. equal to 0.2. This result was surprising. However, for the data set used in the present study, a similar behaviour has been observed for the REML estimates in the model (1) and the model in Przystalski *et al.* (2008). The REML estimates of the genetic correlations were equal: 0.41 (S.E. = 0.45) and 0.44 (S.E. = 0.49), respectively. On the other hand, for the original data set with eleven varieties, the estimated genetic correlations in the two models were equal, 0.80 (S.E. = 0.16) and 0.83 (S.E. = 0.16), and were within the range provided by Przystalski *et al.* (2008) (0.79–1) for the yield. Therefore, it can be concluded that the value of the genetic correlation between the two systems depends on the number of varieties in the data set. Moreover, the value of the genetic correlation between the two systems depends on the aim of the study. In the case of interest in a joint stability assessment in both agronomic systems, a data set with a reduced number of varieties should be used and a low genetic correlation value should be expected. On the other hand, if the main goal is estimation of the genetic correlation between the two systems, then one should use the whole data set. In the present study we were interested in assessing yield stability, for this reason we used results for varieties which were tested for three years.

The models described in the present work can be easily programmed in NIMBLE R-package. In this package the syntax used to specify the model is similar to that used by WinBUGS or JAGS, so scientists familiar with WinBUGS or JAGS will not have problems with specifying their models. Moreover, for large data sets, to speed up the computations one can fit the model in NIMBLE using parallel computing and run each chain on separate core. However, in this case one obtains the WAIC value for each Markov chain and not a single value of the WAIC for all chains. One can also fit the models described in the present study by using the MCMCglmm package (Hadfield, 2012). However, in this case the models should be compared in terms of the deviance information criterion (Spiegelhalter *et al.*, 2002; DIC). Further, the NIMBLE package has an extended list of available probability distributions in comparison to WinBUGS, JAGS or MCMCglmm. For example, one can directly assign the

https://doi.org/10.1017/S0021859623000084 Published online by Cambridge University Press

inverse-Wishart distribution prior for a covariance matrix. In Lunn *et al.* (2013), it was pointed out that the least informative proper inverse-Wishart is obtained by setting degrees of belief equal to $\dim \Sigma_u$. In the present study to estimate the covariance matrix Σ_u an inverse-Wishart distribution with $\dim \Sigma_u + 2$ degrees of belief was used. This distribution can be considered as weakly-informative. In the literature concerning Bayesian statistics, it is argued against using the inverse-Wishart distribution priors for covariance matrices because they impose a degree of informativity and are sensitive to the choice of hyperparameters. Instead, it was proposed to use the half-Cauchy distribution or inverse-Gamma distribution with parameters $\alpha = 2$ and $\beta = 2/A$ (where A is large number) for standard deviations, and scaled-inverse-Wishart (SIW) distribution or LKJ distribution (Lewandowski *et al.*, 2009) for covariance matrices (Gelman *et al.*, 2013). In the MCMCglmm package, by default, an inverse-Wishart distribution is used as prior for variance components and covariance matrices. In the present work, by using the IW distribution in NIMBLE, we assigned hierarchical half-t priors (Huang and Wand, 2013) for the covariance matrices Σ_v and Σ_w. This distribution can be easily programmed within the NIMBLE framework. Alternatively, one can try to program in NIMBLE the SIW or LKJ distributions using the piece of code given in the nimble manual (de Valpine *et al.*, 2022).

In the current study, for each variety, we calculated the probability that the mean of a given variety in a given system exceeds the mean of the controls in that system. A similar approach was used by Eskridge and Mumm (1992), who were also interested in the selection of cultivars based on the probability of dominance of the control cultivar. They calculated the probabilities assuming normality and using a nonparametric model. The normality assumption was also used by Eskridge (1990), Lenartowicz *et al.* (2020) and Przystalski and Lenartowicz (2020) in their risk analysis. However, the distribution of difference between the means of the test variety and the control variety, or the distribution of yield across the environments is usually unknown. For this reason, as in Dias *et al.* (2022), the probabilities were calculated using MCMC samples from Bayesian analysis. In a frequentist approach, a similar procedure was proposed by Piepho (1998) to calculate the probability that one system outperforms another system. In that study, to calculate the probability that System 1 outperforms System 2, he proposed to test both systems in a large number of environments, calculate the difference $D_j = y_{1j} - y_{2j}$ in each, and determine the relative frequency with which $D_j > 0$.

In the present study we used the HMRPVG index and the superiority stability coefficient to simultaneously assess stability of varieties in both agronomic systems. The two stability measures were preferred for several reasons. First of all, just like animal breeders, plant breeders also use best linear unbiased predictions (BLUP's) to select the best lines and to evaluate their stability. For example, Colombari Filho *et al.* (2013) used BLUP's to assess adaptability and stability of lines of upland rice in Brazil. Recently, Bocianowski and Liersch (2021) used the HMPRVG index to select the best oilseed rape lines. In a different study, Derejko *et al.* (2020) evaluated 12 varieties using the superiority stability coefficient separately for each of six agro-zones in Poland, and based on the values of this stability measure they chose the best variety for each region. Next, these two stability measures can be easily estimated within R-NIMBLE framework. Finally, these measures have recently been implemented in the metan package (Olivoto and Lúcio, 2020). However, we also estimated Shukla's stability variances for varieties in both systems (results not shown) by extending the linear mixed model methodology described in Piepho (1999). For this purpose, we assumed in Eqn (5) that

$$\boldsymbol{w}_{jk} | \Sigma_{w,j} \sim N(0, \Sigma_{w,j})$$

For a fixed agronomic system, model (1) with the above assumption reduces to the Shukla's stability variance model described in Przystalski and Lenartowicz (2020). From this perspective, model (1) with the above assumption can be considered as an extension of the Shukla's stability variance model described in that paper to the case of three-way variety × environment × system data set and cannot be implemented in the MCMCglmm package. However, from the practical point of view, model (1) with the above assumption is computationally demanding and time consuming, especially for large data. In our case, we had to generate seven covariance matrices. From this perspective, a better method of estimation Shukla's stability variance is a method proposed by Dias *et al.* (2022). In that paper, the authors assessed the stability of a given genotype across k environments based on the variance of GEI effects ($var(ge)^s_{jk}$). Finally, further improvement of the analysis can be achieved by analysing a variety × year × site × system data set. On the other hand, in VCU and post-registration trials, apart from yield, other traits are observed (e.g. plant height, disease resistance). By slightly modifying the proposed methodology, it is possible to analyse multi-trait multi-environment data (either variety × environment × trait data set or variety × year × site × trait data set) and at the same time assess the stability of the analysed traits. Further, since the problem of recommending varieties for cultivation can be treated as a Bayesian decision-theoretical problem (see e.g. Theobald and Talbot, 2002, 2004; Theobald *et al.*, 2006), one can try to build a multi-trait utility function and calculate a posteriori the expected utility for each variety. We plan to explore these topics in a future work.

Conclusion

The tuber yield in the organic system was approx.44% lower than the tuber yield in the conventional system. For organic system, varieties Tajfun and Otolia were the highest yielding varieties among the tested varieties. For conventional system, the variety Jurek was the highest yielding variety. According to HMPRGV stability index and the superiority index, variety Tajfun was the most stable variety in organic system, whereas variety Jurek was the most stable in conventional system. Additionally, varieties Tajfun and Otolia were the most stable and highest yielding varieties in the organic system, whereas in the conventional system, the variety Jurek was the most stable and highest yielding variety among the tested varieties.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0021859623000084

Acknowledgements. The authors would like to thank Prof Rosemary Bailey for helpful comments and corrections, which led to the improvement of the paper.

Authorship. MP perceived the idea of investigating the topic. TL designed the experiments and was responsible for data curation. MP proposed the statistical methodology and performed the analysis. MP wrote the early version of the manuscript and revised it. MP and TL amended it. Both authors read the manuscript.

Financial support. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflict of interest. The authors declare that they have no conflict of interest.

Ethical standard. Not applicable

References

Annicchiarico P (1992) Cultivar adaptation and recommendation from alfalfa trials in northern Italy. *Journal of Genetics and Breeding* **46**, 269–278.

Annicchiarico P (2002) *Genotype × Environment Interactions: Challenges and Opportunities for Plant Breeding and Cultivar Recommendations*, vol. **1**. Rome: FAO.

Azevedo CF, Nascimento M, Carvalho IR, Nascimento ACC, Ferreira de Almeida HC, Cruz CD and Gonzalez da Silva JA (2022) Updated knowledge in the estimation of genetics parameters: a Bayesian approach in white oat (*Avena sativa L.*). *Euphytica* **218**, 1–13.

Bernardo Júnior LAY, da Silva CP, de Oliveira LA, Nuvunga JJ, Pires LPM, Von Pinho RG and Balestre M (2018) AMMI Bayesian models to study stability and adaptability in maize. *Agronomy Journal* **110**, 1765–1776.

Bocianowski J and Liersch A (2021) Multi-environmental evaluation of winter oilseed rape genotypic performance using mixed models. *Euphytica* **217**, 1–9.

Caliński T, Czajka S, Kaczmarek Z, Krajewski P, Pilarczyk W, Siatkowski I and Siatkowski M (2017) On mixed model analysis of multi-environment variety trials: a reconsideration of the one-stage and the two-stage models and analyses. *Statistical Papers* **58**, 433–465.

Colombari Filho JM, de Resende MDV, de Morais OP, de Castro AP, Guimaraes EP, Pereira JA, Utumi MM and Breseghello F (2013) Upland rice breeding in Brazil: a simultaneous genotypic evaluation of stability, adaptability and grain yield. *Euphytica* **192**, 117–129.

Cowles MK and Carlin BP (1996) Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association* **91**, 833–904.

Crossa J, Perez-Elizalde S, Jarquin D, Cotes JM, Viele K, Liu G and Cornelius PL (2011) Bayesian Estimation of additive main effects and multiplicative interaction model. *Crop Science* **51**, 1458–1469.

Damesa TM, Möhring J, Worku M and Piepho HP (2017) One step at a time: stage-wise analysis of series of experiments. *Agronomy Journal* **109**, 845–857.

de Oliveira LA, da Silva CP, Nuvunga JJ, da Silva AQ and Balestre M (2016) Bayesian GGE biplot models applied to maize multi-environment trials. *Genetics and Molecular Research* **15**, 1–21.

de Oliveira TRA, de Carvalho HWL, Nascimento M, Costa EFN, do Amaral Junior AT, Gravina GDA and de Carvalho Filho JLS (2018) The Eberhart and Russell's Bayesian method used as an instrument to select maize hybrids. *Euphytica* **214**, 1–9.

De Ponti T, Rijk B and Van Ittersum MK (2012) The crop yield gap between organic and conventional agriculture. *Agricultural Systems* **108**, 1–9.

de Valpine P, Turek D, Paciorek CJ, Anderson-Bergman C, Lang DT and Bodik R (2017) Programming with models: writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics* **26**, 403–413.

de Valpine P, Paciorek C, Turek D, Michaud N, Anderson-Bergman C, Obermeyer F, Wehrhahn Cortes C, Rodrìguez A, Lang DT and Paganin S (2022) NIMBLE User Manual. R package manual version 0.12.2. Available at https://r-nimble.org (Assessed 09.06.2022).

Derejko A, Studnicki M, Wójcik-Gront E and Gacek E (2020) Adaptive grain yield patterns of Triticale (*Triticosecale* Wittmack) cultivars in six regions of Poland. *Agronomy* **10**, 1–14.

Dias PC, Xavier A, de Resende MDV, Barbosa MHP, Biernaski FA and Estopa RA (2018) Genetic evaluation of *Pinus taeda* clones from somatic embryogenesis and their genotype × environment interaction. *Crop Breeding and Applied Biotechnology* **18**, 55–64.

Dias KOG, dos Santos JPR, Krause MD, Piepho HP, Guimarães LJM, Pastina MM and Garcia AAF (2022) Leveraging probability concepts for cultivar recommendation in multi-environment trials. *Theoretical and Applied Genetics* **135**, 1385–1399.

Dierauer H, Gelencsér T and Klaiss M (2021) Sortenliste Biokartoffeln. Available at https://www.fibl.\newline.org/en/shop-en/1041-biokartoffeln (Assessed 16.02.2022).

Digby PGN (1979) Modified joint regression analysis for incomplete variety × environment data. *The Journal of Agricultural Sciences* **93**, 81–86.

Eberhart SA and Russell WA (1966) Stability parameters for comparing varieties. *Crop Science* **6**, 36–40.

Edwards JW and Jannink JL (2006) Bayesian modeling of heterogeneous error and genotype × environment interaction variances. *Crop Science* **46**, 820–833.

Edwards JW and Orellana M (2015) Increasing selection response by Bayesian modeling of heterogeneous environmental variances. *Crop Science* **55**, 556–563.

Eskridge KM (1990) Selection of stable cultivars using a safety-first rule. *Crop Science* **30**, 369–374.

Eskridge KM and Mumm RF (1992) Choosing plant cultivars based on the probability of outperforming a check. *Theoretical and Applied Genetics* **84**, 494–500.

Finlay K and Wilkinson G (1963) The analysis of adaptation in a plant breeding programme. *Australian Journal of Agricultural Research* **14**, 742–754.

Flis B, Domański L, Zimoch-Guzowska E, Polgar Z, Pousa SÁ and Pawlak A (2014) Stability analysis of agronomic traits in potato cultivars of different origin. *American Journal of Potato Research* **91**, 404–413.

Gauch HG (1992) *Statistical Analysis of Regional Yield Trials. AMMI Analysis of Factorial Designs*. New York: Elsevier.

Gelman A and Rubin DB (1992) Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science* **7**, 457–511.

Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A and Rubin DB (2013) *Bayesian Data Analysis*, 3rd Edn. Boca Raton: Chapman & Hall/CRC.

Hadfield JD (2012) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software* **33**, 1–22.

Hedges LV and Olkin I (1985) *Statistical Methods for Meta-Analysis*. Orlando: Academic Press.

Hoagland C (2009) Impact of conventional and organic production on agronomic and end-use quality traits of winter wheat (MS thesis). Univ. of Nebraska, Lincoln.

Huang A and Wand MP (2013) Simple marginally noninformative prior distributions for covariance matrices. *Bayesian Analysis* **8**, 439–452.

Josse J, van Eeuwijk F, Piepho HP and Denis JB (2014) Another look at Bayesian analysis of AMMI models for genotype-environment data. *Journal of Agricultural, Biological and Environmental Statistics* **19**, 240–257.

Kazimierczak R, Średnicka-Tober D, Hallmann E, Kopczyńska K and Zarzyńska K (2019) The impact of organic vs. conventional agricultural practices on selected quality features of eight potato cultivars. *Agronomy* **9**, 799.

Kirk AP, Fox SL and Entz MH (2012) Comparison of organic and conventional selection environments for spring wheat. *Plant Breeding* **131**, 687–694.

Kucek LK, Santantonio N, Gauch HG, Dawson JC, Mallory EB, Darby HM and Sorrells ME (2019) Genotype×environment interactions and stability in organic wheat. *Crop Science* **58**, 1–8.

Lenartowicz T, Piepho HP and Przystalski M (2020) Stability analysis of tuber yield and starch yield in mid-late and late maturing starch cultivars of potato (*Solanum tuberosum*). *Potato Research* **63**, 179–197.

Lesur-Dumoulin C, Malézieux E, Ben-Ari T, Langlais C and Makowski D (2017) Lower average yields but similar yield stability in organic versus conventional horticulture. A meta-analysis. *Agronomy for Sustainable Development* **37**, 45.

Lewandowski D, Kurowicka D and Joe H (2009) Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis* **100**, 1989–2001.

Lian L and de los Campos G (2016) FW: an R package for Finlay–Wilkinson regression that incorporates genomic/pedigree information and covariance structures between environments. *G3 Genes, Genomes, Genetics* **6**, 586–597.

https://doi.org/10.1017/S0021859623000084 Published online by Cambridge University Press

Lin CS and Binns MR (1988) A superiority measure of cultivar performance for cultivar × location data. *Canadian Journal of Plant Sciences* **68**, 193–198.

Löschenberger F, Fleck A, Grausgruber H, Hetzendorfer H, Hof G, Lafferty J, Marn M, Neumayer A, Pfaffinger G and Birschitzky J (2008) Breeding for organic agriculture: the example of winter wheat in Austria. *Euphytica* **163**, 469–480.

Lunn D, Jackson C, Best N, Thomas A and Spiegelhalter D (2013) *The BUGS Book. A Practical Introduction to Bayesian Analysis*. Boca Raton: Chapman& Hall/CRC.

Mathew B, Holand AM, Koistinen P, Léon J and Sillanpää MJ (2016) Reparametrization-based estimation of genetic parameters in multi-trait animal model using integrated nested Laplace approximation. *Theoretical and Applied Genetics* **129**, 215–225.

Mead R, Riley J, Dear K and Singh SP (1986) Stability comparison of intercropping and monocropping systems. *Biometrics* **42**, 253–266.

Murphy KM, Campbell KG, Lyon SR and Jones SS (2007) Evidence of varietal adaptation to organic farming systems. *Field Crops Research* **102**, 172–177.

Nascimento M, Nascimento ACC, Fabyano Fonseca e Silva FF, Teodoro PE, Azevedo CF, de Oliveira TRA, do Amaral Junior AT, Cruz CD, Farias FJC and de Carvalho LP (2020) Bayesian Segmented regression model for adaptability and stability evaluation of cotton genotypes. *Euphytica* **216**, article 30.

Olivoto T and Lúcio ADC (2020) metan: an R package for multi-environment trial analysis. *Methods in Ecology and Evolution* **11**, 783–789.

Orellana M, Edwards J and Carriquiry A (2014) Heterogeneous variances in multi-environment yield trials for corn hybrids. *Crop Science* **54**, 1048–1056.

Pedersen TM (2012) *Organic VCU Testing. Current status in 16 European Countries*. Aarhus, Denmark: Knowledge Centre for Agriculture (SEGES).

Piepho HP (1996) A simplified procedure for comparing the stability of cropping systems. *Biometrics* **52**, 315–320.

Piepho HP (1998) Methods for comparing the yield stability of cropping systems. *Journal of Agronomy and Crop Science* **180**, 193–213.

Piepho HP (1999) Stability analysis using the SAS system. *Agronomy Journal* **91**, 154–160.

Piepho HP (2000) Exact confidence limits for covariate-dependent risk in cultivar trials. *Journal of Agricultural, Biological and Environmental Statistics* **5**, 202–213.

Plummer M, Best N, Cowles K and Vines K (2006) CODA: convergence diagnosis and output analysis for MCMC. *RNews* **6**, 7–11.

Ponisio LC, M'Gonigle LK, Mace KC, Palomino J, de Valpine P and Kremen C (2015) Diversification practices reduce organic to conventional yield gap. *Proceedings of the Royal Society B* **282**, 20141396.

Przystalski M and Lenartowicz L (2020) Yielding stability of early maturing potato varieties: Bayesian analysis. *The Journal of Agricultural Science* **158**, 564–573.

Przystalski M, Osman A, Thiemt EM, Rolland B, Ericson L, Østergård H, Levy L, Wolfe M, Büchse A, Piepho HP and Krajewski P (2008) Comparing the performance of cereal winter wheat varieties in organic and non-organic systems in different European countries. *Euphytica* **163**, 417–433.

Rakszegi M, Mikó P, Löschenberger F, Hiltbrunner J, Aebi R, Knapp S, Tremmel-Bede K, Megyeri M, Kovács G, Molnár-Láng M, Vida G, Láng L and Bedő Z (2016) Comparison of quality parameters of wheat varieties with different breeding origin under organic and low-input conventional conditions. *Journal of Cereal Science* **69**, 297–305.

Reid TA, Yang R-C, Salmon DF, Navabi A and Spaner D (2011) Realized gains from selection for spring wheat grain yield are different in conventional and organically managed systems. *Euphytica* **177**, 253–266.

Resende MDV (2007) *Matematica e estatistica na Analise de Experimentos e no Melhoramento Genetico*. Colombo: Embrapa Florestas.

Schrama M, de Haan JJ, Kroonen M, Verstegen H and Van der Putten WH (2018) Crop yield gap and stability in organic and conventional farming systems. *Agriculture, Ecosystems and Environment* **256**, 123–130.

Shah A, Askegaard M, Rasmussen IA, Jimenez EMC and Olesen JE (2017) Productivity of organic and conventional arable cropping systems in long-term experiments in Denmark. *European Journal of Agronomy* **90**, 12–22.

Shukla GK (1972) Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity* **29**, 237–245.

Silva FF, Viana JMS, Faria VR and de Resende M (2013) Bayesian inference of mixed models in quantitative genetics of crop species. *Theoretical and Applied Genetics* **126**, 1749–1761.

Spiegelhalter DJ, Best NG, Carlin BP and Van der Linde A (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B* **64**, 583–689.

Theobald CM and Talbot M (2002) The Bayesian choice of crop variety and fertilizer dose. *Journal of the Royal Statistical Society Series C – Applied Statistics* **51**, 23–36.

Theobald CM and Talbot M (2004) Bayesian selection of fertilizer level when crop price depends on quality. *Computational Statistics and Data Analysis* **47**, 867–880.

Theobald CM, Talbot M and Nabugoomu F (2002) Bayesian approach to regional and local-area prediction from crop variety trials. *Journal of Agricultural, Biological and Environmental Statistics* **7**, 403–419.

Theobald CM, Roberts AMI, Talbot M and Spink JM (2006) Estimation of economically optimum seed rates for winter wheat from series of trials. *The Journal of Agricultural Science* **144**, 303–316.

Willer H, Trávníček J, Meier C and Schlatter B (eds) (2022) The World of Organic Agriculture. Statistics and Emerging Trends 2022. Research Institute of Organic Agriculture FiBL, Frick, and IFOAM – Organics International, Bonn. (Assessed 16.02.2022).

Yan W and Kang MS (2003) *GGE Biplot Analysis: A Graphical Tool for Breeders, Genetists and Agronomists*. Boca Raton: CRC Press.

Youngflesh C (2018) MCMCvis: tools to visualize, manipulate, and summarize MCMC output. *Journal of Open Source Software* **3**, 640.

Appendix

This Appendix presents the Genstat code used to fit the model described in Przystalski *et al.* (2008) and model (1).

```
job
import 'orgconv.xlsx';sheet = 'Arkusz10'
list
pointer [values = sys, env] d
facproduct d;ss
'Przystalski et al. (2008)'
vcomponents [fixed = system + env + env.system;factorial = 9;experiments = ss]\
random = variety.system + env.variety + variety.system.env +
env.system.rep
vstructure [terms = system.variety] model = correlation, iden;heterogeneity
= outside; factor = system, variety
reml [print = model, components, wald, deviance, means; pse = differ-
ences; mvinclude = *;\
method = AI;workspace = 30]yield/10
'Model (1)'
vcomponents [fixed = system;factorial = 9;experiments = ss]\
random = env.system + variety.system + variety.system.env +
env.system.rep
vstructure [terms = system.env] model = correlation, iden;heterogeneity =
outside; factor = system, env
vstructure [terms = system.variety] model = correlation, iden;heterogeneity
= outside; factor = system, variety
vstructure [terms = system.variety.env] model = correlation, iden, iden;het-
erogeneity = outside; factor = system, variety, env
vstructure [terms = env.system.rep] model = iden, diag, iden; factor = env,
system, rep
reml [print = model, components, wald, deviance, means; pse = differ-
ences; mvinclude = *;\
method = AI;workspace = 30]yield/10
```

https://doi.org/10.1017/S0021859623000084 Published online by Cambridge University Press