

Article

Genomic Prediction of Kernel Water Content in a Hybrid Population for Mechanized Harvesting in Maize in Northern China

Ping Luo ^{1,2,3,†} , Ruisi Yang ^{1,2,†}, Lin Zhang ⁴, Jie Yang ⁵, Houwen Wang ², Hongjun Yong ² , Runze Zhang ², Wenzhe Li ², Fei Wang ², Mingshun Li ², Jianfeng Weng ², Degui Zhang ², Zhiqiang Zhou ², Jienan Han ², Wenwei Gao ¹, Xinlong Xu ², Ke Yang ², Xuecai Zhang ³, Junjie Fu ² , Xinhai Li ², Zhuanfang Hao ^{2,*} and Zhiyong Ni ^{1,*} 

- ¹ College of Agriculture, Xinjiang Agricultural University, Urumqi 830052, China; luoping987@126.com (P.L.); yangruisixj@163.com (R.Y.)
- ² State Key Laboratory of Crop Gene Resources and Breeding, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
- ³ International Maize and Wheat Improvement Center (CIMMYT), Texcoco 56237, Mexico
- ⁴ College of Agronomy, Northeast Agricultural University, Harbin 150030, China
- ⁵ Food Crops Research Institute, Xinjiang Academy of Agricultural Science, Urumqi 830052, China
- * Correspondence: haozhuanfang@163.com (Z.H.); nizhiyong@126.com (Z.N.)
- † These authors contributed equally to this work.

Abstract: Genomic prediction enables rapid selection of maize varieties with low kernel water content (KWC), facilitating the development of mechanized maize harvesting and reducing costs. This study evaluated and characterized the KWC and grain yield (GY) of hybrid maize in northern China and used genomic prediction to identify superior hybrid combinations with low kernel water content at maturity (MKWC) and high GY adapted to northern China. A total of 285 hybrids obtained from single crosses of 34 inbred lines from Stiff Stalk and Non-Stiff Stalk heterotic groups were used for genomic prediction of KWC and GY. We tested 20 different statistical prediction models considering additive effects and evaluating the impact of dominance and epistasis on prediction accuracy. Employing 10-fold cross-validation, it showed that the average prediction accuracy ranged drastically from 0.386 to 0.874 across traits and models. Eight linear statistical methods displayed a very similar prediction accuracy for each trait. The average prediction accuracy of machine learning methods was lower than that of linear statistical methods for KWC-related traits, but the random forest model had a high prediction accuracy of 0.510 for GY. When genetic effects were incorporated into the prediction model, the prediction accuracy for each trait was improved. Overall, the model with dominant and epistatic effects (G:AD(AA)) performed best. For the same number of markers, predictions using trait-specific markers resulted in higher prediction accuracy than randomly selected markers. When the number of trait-specific SNPs was set to 100, the prediction accuracy of GY increased by 33.27%, from 0.406 to 0.541. Out of all the 561 potential hybrids, the TOP 30 hybrids selected by genomic prediction would lead to a 1.44% decrease in MKWC compared with Xianyu335, a hybrid with a fast kernel water dry-down, and these hybrids also had higher GY simultaneously. Our results confirm the value of genomic prediction for hybrid breeding low MKWC suitable for maize mechanized harvesting in northern China. In conclusion, this study highlights the potential of genomic prediction to optimize maize hybrid breeding, enhancing efficiency and providing insights into genotype-accuracy relationships. The findings offer new strategies for hybrid design and advancing mechanized harvesting in northern China.

Keywords: maize; genomic prediction; kernel water content; genetic effects; mechanized harvesting



Citation: Luo, P.; Yang, R.; Zhang, L.; Yang, J.; Wang, H.; Yong, H.; Zhang, R.; Li, W.; Wang, F.; Li, M.; et al. Genomic Prediction of Kernel Water Content in a Hybrid Population for Mechanized Harvesting in Maize in Northern China. *Agronomy* **2024**, *14*, 2795. <https://doi.org/10.3390/agronomy14122795>

Academic Editor: Hongwei Cai

Received: 6 November 2024

Revised: 22 November 2024

Accepted: 23 November 2024

Published: 25 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Maize is in great demand worldwide as food and feed, and with global warming, it has the potential to be an excellent cereal crop in the future as a C4 plant. As human labor costs continue to rise, mechanized maize harvesting is an inevitable trend in the modernization of agriculture. Low kernel water content at harvest (HKWC) is a key factor for efficient mechanical harvesting, as it significantly reduces the need for post-harvest drying. When the HKWC was low, the grain required less drying time and energy, thereby reducing drying costs and processing time. This was particularly crucial in large-scale farming operations, where rapid processing and cost efficiency are essential [1]. The HKWC, in turn, depends on MKWC (kernel water content at maturity) at the physiological maturity stage and kernel water dry-down rate (DR). Therefore, accelerating the selection of maize varieties suitable for mechanization is an important goal for breeders and is also of great significance for the development of agricultural modernization.

HKWC is determined by two stages: the first stage is pre-physiological maturity, which is the rapid filling stage of the maize kernel. During this stage, dry matter increases rapidly and reaches a maximum value. Grain filling continues with a gradual increase in kernel water content (KWC), which peaks and then begins to decline gradually to a state of net water loss. This stage is primarily a developmental dehydration stage, where water dynamics vary from variety to variety and are regulated mainly by genetic factors. The second stage is the dehydration process after physiological maturity, marked by the formation of a black layer at the base of the kernel, which cuts off the material connection from the plant; the dry weight of the KWC remains stable, and the water content gradually decreases. At this time, the KWC is not solely determined by genetics but is also affected by external environmental variables [2,3]. The dehydration rate after the physiological maturity of maize kernel is closely related to HKWC. Fast kernel filling and kernel water dry-down in the field are distinctive characteristics of maize varieties with low KWC while maintaining high yield potential [4,5].

KWC is strongly associated with kernel composition, such as sugars and certain hydrophobic compounds [6,7], as well as a range of agronomic characteristics. These traits encompass kernel row number, ear length and diameter, and husk tightness and length [8–13]. KWC, as a quantitative trait, can be affected by multiple factors at varying points in time, thereby complicating the process of trait estimation. Previous studies have revealed that KWC was controlled by numerous quantitative trait loci (QTL), and hundreds of QTL for maize KWC have been identified [14–24]. The use of marker-assisted selection (MAS) for KWC in maize breeding programs was not ideal due to the lack of large effect QTL, and the incorporation of additional markers was necessary. Unlike MAS, genomic prediction (GP) used the collective contribution of all genetic markers [25], offering a more comprehensive and accurate method for selecting hybrids with low KWC, regardless of the small effect sizes of individual QTL. Therefore, GP can overcome the challenges posed by QTL-based approaches, providing a more efficient and precise solution for improving KWC in maize breeding programs.

GP is a highly effective method of utilizing markers across the entire genome for forecasting crop trait outcomes, even in the absence of information regarding genes associated with the traits [26]. To date, a number of statistical methods have been developed for GP, including genomic best linear unbiased prediction (GBLUP) [27,28], Bayesian approach (Bayes A, B, C) [29,30], ridge regression (rrBLUP) [31,32], and least absolute shrinkage and selector operator (LASSO) [33,34]. In addition, nonparametric methods for genomic prediction based on machine learning (ML) techniques such as random forest (RF) [35–37], support vector machine (SVM) [38,39], and feed-forward artificial neural networks (ANN) were proposed [40]. Some progress has been made in GP related to maize breeding. By optimizing parameters including single-nucleotide polymorphism (SNP) marker density and training population size, a high-accuracy GP model was developed for evaluating traits such as plant height [41], husk length [42], and GY [43,44]. Modeling using the training population is a key step in GP. Prediction models, such as LASSO, rrBLUP, and

ML methods, can be used for GP [45]. In maize breeding, GP is used to predict and select target traits in breeding lines [26,46]. It has also been utilized to predict the performance of hybrids and accurately assess the general and specific combining abilities (GCA and SCA) of the parental lines [47].

Maize mechanical harvesting places high demands on controlling grain water content [48]. Excessive grain water affects threshing efficiency, increases drying and storage costs, and reduces grain quality and market value. Globally, controlling KWC is a critical factor in enhancing maize production efficiency. In regions facing similar challenges, such as the US corn belt, genetic improvements and agronomic practices have been implemented to accelerate kernel dry-down [49]. In Europe, efforts have also focused on breeding maize hybrids with lower KWC to meet the demands of mechanized systems [50]. In northern China, due to rapid temperature drops in autumn, the natural drying process of maize kernels during maturity is relatively slow, making high water content a major challenge for mechanical harvesting [51]. Recent studies have explored the genetic basis of kernel dry-down traits and their application in hybrid breeding [24,52]. However, most research has relied on traditional phenotypic selection, which is time-consuming and less precise. Northern China's unique climate further amplifies the challenge, necessitating innovative breeding strategies to address the slow natural drying process. Ensuring that maize is harvested at an optimal KWC is essential for achieving efficient, low-cost production, boosting farmer profits, and advancing agricultural mechanization.

GP has been performed on maize, and all have yielded promising results; however, few studies have used GP for KWC-related traits in maize. This study evaluated 285 maize hybrids, with phenotypic data collected across four environments. These data were used to (1) understand the distribution of KWC in maize hybrids in northern China, (2) estimate the prediction accuracy of different prediction models for KWC, (3) appraise the influence of genetic effects and specific markers on prediction accuracy, and (4) explore the potential of GP for KWC in maize breeding.

2. Materials and Methods

2.1. Plant Material and Field Trials

In this study, 34 diverse inbred lines were chosen, of which 18 inbred lines belonged to the Stiff Stalk (SS) heterotic group, and the remaining 16 lines belonged to the Non-Stiff Stalk (NSS, Sipingtou/Iodent) heterotic group. Two hundred and eighty-five single-cross hybrids were produced from a partial diallel cross design between the two groups of lines [53].

The field experiments were performed in four geographically distinct locations in 2022: Harbin, Heilongjiang (HLJ, 130° E, 46° N), Gongzhuling, Jilin (JL, 125° E, 44° N), Tieling, Liaoning (LN, 118° E, 38° N), and Urumqi, Xinjiang (XJ, 86° E, 45° N), respectively. In every location, hybrids were planted in double-row plots utilizing an incomplete block design comprising two replicates. The length of each row was 4.0 m with a row spacing of 0.6 m. The number of plants sown in each double-row plot for each hybrid was fourteen. Field management followed the best local practices.

2.2. Phenotypic Data Collection and Analysis

The two hundred and eighty-five maize hybrids were sown on four dates based on the growing season of maize hybrids to achieve similar physiological maturity in each environment. All hybrids were under open pollination. The time of tassel anthesis and ear silking was recorded for each hybrid in each environment, and kernel water content at grain physiological maturity (MKWC) collection was performed based on the time of maturity (55–65 days after anthesis). The MKWC for 5 uniformly growing plants in each row was measured using a hand-held moisture meter (TimberMaster). MKWC was calibrated according to the temperature and humidity at the time of measurement. In general, the effect of temperature on the MKWC calibration can be approximately compensated as follows: for every 5 °C above 20 °C, subtract 0.5 from the value displayed. For every 5 °C below 20 °C, add 0.5mc from the value displayed. The effect of humidity on the

MKWC calibration can be approximately compensated as follows: for every 10% above 65%, subtract 0.5 from the value displayed. For every 10% below 65%, add 0.5 from the value displayed. Kernel water content at harvest (HKWC), kernel water dry-down rate (DR), and grain yield (GY) were also estimated. GY was adjusted to a 14% standard water content. HKWC was measured using a grain moisture meter after harvest at each location. DR was calculated using the following formula: $DR = MKWC - HKWC$.

Phenotypic data were evaluated with R version 4.3.1 on Windows (<https://www.r-project.org/>, accessed on 3 March 2024). To reduce environmental impacts, maize phenotype BLUP (best linear unbiased prediction) was computed across four locations through the lme4 package's lmer function [54]. The ggplot2 and patchwork package were used to analyze the correlation coefficients and make a box plot. Generalized heritability was determined from variance components:

$$H^2 = \left(\delta_g^2 \right) / \left(\delta_g^2 + \delta_g e^2 / y + \delta_e^2 / yr \right)$$

where y and r represent the number of growing locations and the number of field replicates per location, respectively; δ_g^2 represents genetic variance; $\delta_g e^2$ represents genotype-by-location interaction variance; and δ_e^2 is residual variance [55].

2.3. Genotyping

The 34 diverse parental inbred lines, including 18 from the SS and 16 from the NSS heterotic group, were genotyped using the GenoBaits Maize 45K Panel [56]. Briefly, the final 23,692 SNPs were used for the subsequent analysis after quality control [53]. Finally, both genotypic and phenotypic data were available for 285 crosses, including crosses between the 18 SS lines and the 16 NSS lines. Based on the mating design, genomic profiles of the 285 single-cross hybrids were derived using genotypes of the parental lines.

2.4. Genomic Prediction with Different Models

We compared twenty different statistical prediction models considering the genetic effects of the four traits. Based on the characterization of the Reproducing Kernel Hilbert Space (RKHS) with a user-arbitrarily defined covariance structure [30], we performed genomic prediction using the RKHS, considering additive, dominance, and epistatic effects. The difference with previous studies was that we considered the effect of gamete origin and used a genomic relationship matrix completely developed from theory [57] (rather than transplants based on the concept of pedigree).

2.4.1. G Model

Studies have indicated that this model was defined as grounded in hybrid genotypes and overlooked the origin of gametes (Stiff Stalk and Non-Stiff Stalk parental origins) [58]. The G-model can be written as follows:

$$y = 1\mu + g_{A(H)} + g_{D(H)} + g_{AA(H)} + E$$

where $g_{A(H)}$ are the additive genetic effects of hybrids distributed as $g_{A(H)} \sim \text{MVN}\left(0, G_{A(H)}\sigma_{A(H)}^2\right)$, the dominance genetic effects are $g_{D(H)} \sim \text{MVN}\left(\mathbf{0}, D_H\sigma_{D(H)}^2\right)$, and the additive-by-additive epistatic. For hybrids, the details of computing these various genomic relationship matrices were described in the study by González-Diéguez and colleagues [57].

We also tested the $g_{A(H)}$ under different statistical models, such as linear models and ML models. Linear models used include GBLUP, RR, LASSO, Bayes A, B, C, partial least squares (PLS), generalized linear models (GLM); ML models include RF, SVM, gradient boosted machine (GBM), and ANN. GBLUP was implemented using the R package "sommer" [59]. RR and LASSO models were implemented using the R package "glmnet", and the lambda parameters were optimized by 10-fold cross-validation [33]. RKHS employs

a Gaussian process with an arbitrary user-defined covariance structure to realize various genomic relationships in regression. Bayes (A, B, C) and RKHS were implemented using the R package “BGLR” [31]. PLS, GLM, RF, SVM, GBM, and ANN were implemented using the Sparse Kernel Methods (SKM) R library [60].

2.4.2. GCA Model

The GCA model defines the genetic effect based on gamete origin [57]. Thus, the GCA model for the hybrid (S; N) resulting from crossing the parental lines S (from the Stiff Stalk population) and N (from the Non-Stiff Stalk population) can be formulated as follows:

$$y = 1\mu + Z_d g_{A_S} + Z_f g_{A_N} + Z g_D + Z_d g_{AA_S} + Z_f g_{AA_N} + Z g_{AA_{SN}} + e$$

where y is the vector of the original phenotype of the n phenotyped hybrids, and μ is the intercept. g_{A_S} and g_{A_N} are the vectors of the random additive effect (A) from the stiff stalk and Non-stiff stalk parental lines with $g_{A_S} \sim \mathcal{N}(0, K_{A_S} \sigma_{A_S}^2)$ and $g_{A_N} \sim \mathcal{N}(0, K_{A_N} \sigma_{A_N}^2)$, respectively. g_D is the vector of random dominance effect (D) with $g_D \sim \mathcal{N}(0, K_D \sigma_D^2)$. g_{AA_S} is the vector of random A-by-A epistatic effect (E) within the stiff stalk (resp. Non-stiff stalk) A-by-A E across the stiff stalk and Non-stiff stalk populations. $g_{AA_{SN}} \sim \mathcal{N}(0, K_{AA_{SN}} \sigma_{AA_{SN}}^2) \cdot K_{A_S}, K_{A_N}$, epistasis within the stiff stalk population, A-by-A E within the Non-Stiff Stalk population, and the A-by-A E across populations genomic kinship matrices calculated according to González-Diéguez et al. [57]. e is the vector of error terms, with $e \sim \mathcal{N}(0, I\sigma_e^2)$. The various random effects are considered to be independent. Refer to Table 1 for a summary of the models that were tested.

Table 1. Definition of tested genomic prediction models.

Models	Model Code	Random Genetic Effects	Reference		
G	G: A (GBLUP)	gA	VanRaden (2008) [28] Endelman (2011) [32] Friedman et al. (2010) [33] Karkkainen et al. (2012) [29] Montesinos-López et al. (2023) [60]		
	G: A (rrBLUP)				
	G: A (LASSO)				
	G: A (Bayes A)				
	G: A (Bayes B)				
	G: A (Bayes C)				
	G: A (PLS)				
	G: A (GLM)				
	G: A (RF)				
	G: A (SVM)				
	G: A (GBM)				
	G: A (ANN)				
	G: AD			$gA + gD$	Vitezica et al. (2017) [58]
	G: A(AA _H)				
G: AD(AA _H)					
GCA	GCA: A	$gA_S + gA_N + r$	González-Diéguez et al. (2021) [57]		
	GCA: AD	$gA_S + gA_N + gD$			
	GCA: A(AA _{SN})	$gA_S + gA_N + gAA_{SN}$			
	GCA: AD(AA _{SN})	$gA_S + gA_N + gD + gAA_{SN}$			
	GCA: AD(AA _S)(AA _N)(AA _H)	$gA_S + gA_N + gD + gAA_S + gAA_N + gAA_{SN}$			

G and GCA models are models that successively added additive effects (A), dominance effects (AD), and additive-by-additive genetic effects (AD(AA)). The list of the random genetic effects considered in the GCA models corresponds to Stiff Stalk (g_{A_S}), Non-Stiff Stalk (g_{A_N}), intragroup additive-by-additive epistasis for the Stiff Stalk (g_{AA_S}) and Non-Stiff Stalk group (g_{AA_N}), and intergroup additive-by-additive epistasis ($g_{AA_{SN}}$) effects. In the G models, random genetic effects within hybrids(H) correspond to additive (gA), dominance (gD), and additive-by-additive epistasis (gAA) effects.

2.5. The Design of SNP Marker Dataset and Prediction of Specific Markers as Fixed Effects

The study chose specific markers through genome-wide association analysis (GWAS) within the hybrid population and then utilized them for genomic prediction through

two methods. Two strategies, GWAS-assistant selection and random selection, were used to build the SNP marker dataset. In the GWAS-assistant selection strategy, the MLM (mixed linear model) algorithm was used for GWAS. Furthermore, we investigated the effect of the number of ts_SNPs on the prediction accuracy. Six random-SNP marker datasets with 50, 100, 500, 1000, 5000, and 10,000 SNPs were established for GP. Then, GP analysis was performed with two SNP marker datasets. The markers specific to the trait were regarded as fixed effects in the analysis [61]. The trait-specific SNP sets (ts_SNPs) were selected based on the score of SNPs ($-\log^{10}(P)$). For the ts_SNPs dataset, the top 50, 100, 500, 1000, 5000, and 10,000 SNPs with the highest $-\log^{10}(P)$ value were selected as six ts_SNPs marker datasets, respectively. For the random-SNP dataset, the same number of SNPs were retrieved using the biomaRt package and the script retrieve_SNP_info.R randomly. In addition, we tested the effect of changes in the top SNP dataset on prediction accuracy when the top SNPs were modeled as a fixed effect (fixed).

We used 10-fold cross-validation to evaluate the accuracy of GP in this study. In each cross-validation, 90% of the hybrids were randomly selected as the training set and the remaining 10% as a test set. The cross-validation was repeated 100 times, and the prediction accuracy was assessed using Pearson's correlation coefficient between the observed and the predicted phenotypic values.

2.6. Superior Hybrids Selected by Genomic Prediction

We used all 285 hybrids as a training set and predicted the hybrid performance of all possible 561 single-cross combinations between the 34 parent inbred lines based on the GBLUP (A + D) model. We selected the 30 hybrids with the lowest kernel water content at maturity (5%, 30/561, MKWC_TOP 30) and compared the predicted phenotypic values of these hybrids with all other single-cross combinations (Others). Based on the predicted breeding values, we calculated the selection difference, that is, the difference between the mean performance of the selected hybrids and the performance of the control "Xianyu 335" (a commercial hybrid with fast dehydration in China). Based on the 30 hybrids with the lowest MKWC of predicted value, predicted values of HKWC, DR, and GY based on these hybrids were used to compare gains with all remaining hybrids.

3. Results

3.1. Extensive Phenotypic Variation About KWC-Related Traits Observed in Field Trials

The mean values of MKWC in the four locations, HLJ, JL, LN, and XJ, were 37.3%, 37.7%, 39.3%, and 38.7%, respectively (Figure 1A). The MKWC for all hybrids in the four environments ranged from 31.2% to 43.0%, reflecting considerable phenotypic variation across environments. This variation was critical because it underscored the influence of environmental factors, such as temperature and humidity, on MKWC. The MKWC at these four environments showed a good correlation with an average Pearson correlation of 0.328 ($p < 0.001$) (Figure 1B). MKWC was influenced by the genotypic effect (G), environmental effect (E), and genotype-environment interaction ($G \times E$). Significant genotypic effect (G) was observed ($p < 0.001$), and the broad-sense heritability (H^2) reached 0.75 (Table 2), indicating that genetic factors play a significant role in determining MKWC traits.

Given the high correlation of MKWC in multiple environments, the BLUP method was used to integrate MKWC data from the four environments to obtain stable individual genetic MKWC_BLUP values. The distribution of MKWC_BLUP in the maize hybrid population was in line with the normal distribution (Figure 1C). Analysis of variance (ANOVA) showed that different maize hybrids differed significantly ($p < 0.01$) in MKWC_BLUP, with a mean coefficient of variation (CV) over 4.91% (Figure 1C).

We also collected three traits, HKWC, DR, and GY, but the phenotypic data of these three traits were not collected in XJ due to weather reasons. The distribution of these three traits in the population conformed to a normal distribution (Figure S1). Heritabilities were 0.85, 0.76, and 0.53, respectively. For all three traits, the genotypic effect and genotype-

environment interaction effect were highly significant ($p < 0.001$). However, the correlations between the traits were low and even negative for some of the traits (Table S1).

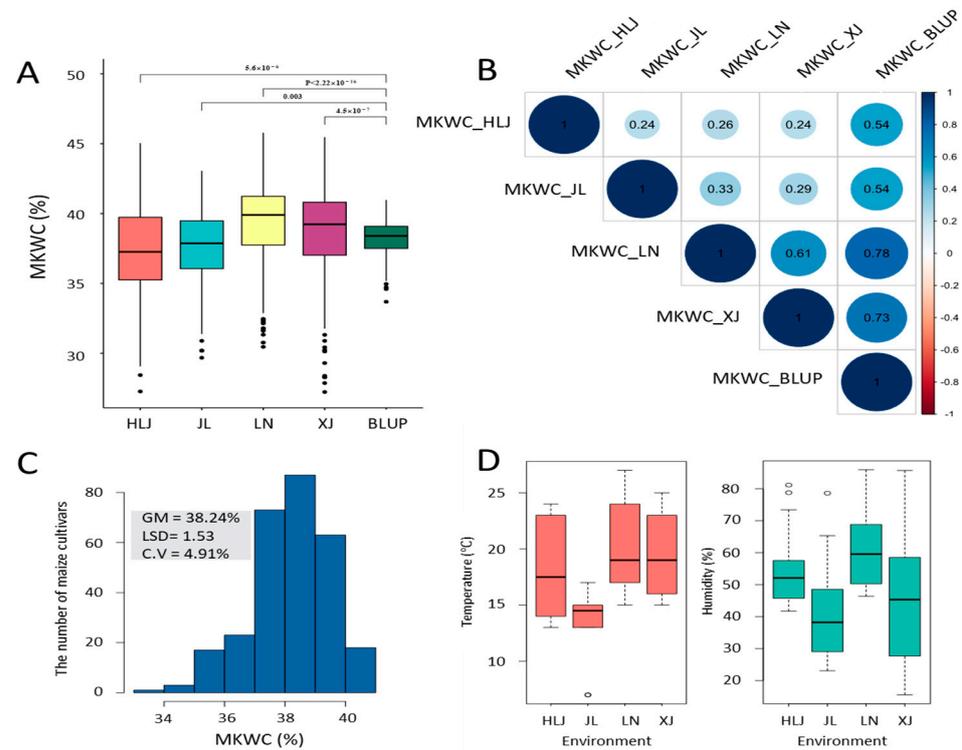


Figure 1. Phenotypic variation distribution, phenotypic analysis, and temperature and humidity calibration chart. Distribution of MKWC in four environments (A). Correlation analysis of the MKWC among HLJ, JL, LN, XJ, and MKWC_BLUP (B). The correlation coefficient between each pair of environments was calculated. The bar shows the value of the correlation coefficient. The frequency distribution of MKWC_BLUP value in 285 hybrid maize (C). Temperature and humidity distribution of sampling environment (D).

Table 2. Variance analysis of maize kernel water content-related traits and yield.

Source	Traits	Df	Sum Sq	Mean Sq	F Value	p-Value	H ²
G	MKWC	284	0.674	0.002	6.221	$< 2 \times 10^{-16}$ ***	0.75
	HKWC	284	0.573	0.002	8.189	$< 2 \times 10^{-16}$ ***	0.85
	DR	284	0.561	0.002	3.439	$< 2 \times 10^{-16}$ ***	0.76
	GY	284	1110.060	3.909	3.476	$< 2 \times 10^{-16}$ ***	0.53
E	MKWC	3	0.145	0.048	127.068	$< 2 \times 10^{-16}$ ***	
	HKWC	2	0.358	0.179	725.118	$< 2 \times 10^{-16}$ ***	
	DR	2	0.068	0.034	58.908	$< 2 \times 10^{-16}$ ***	
	GY	2	309.340	154.669	137.557	$< 2 \times 10^{-16}$ ***	
G × E	MKWC	847	0.512	0.001	1.587	2.463×10^{-13} ***	
	HKWC	563	0.178	0.000	1.280	6.233×10^{-4} ***	
	DR	562	0.462	0.001	1.430	1.519×10^{-6} ***	
	GY	563	800.420	1.422	1.264	1.091×10^{-3} ***	

*** Significant at $p = 0.001$; G, genotype; E, environment; G × E, genotype-by-environment interaction.

3.2. Performance of Prediction Models Based on Hybrid Genotypes

The accuracies of 12 different genomic prediction models relying on additive effects were compared by using 10-fold cross-validation (Table 3). The average prediction accuracy ranged between 0.386 and 0.874 across traits and models. For each trait, we observed that the prediction accuracies of the eight linear methods were similar, while the prediction

accuracies of these ML models RF and SVM were similar to that of the linear methods, with the RF model performing the best in predicting GY at 0.510. ML models GBM and ANN had lower prediction accuracies. Prediction accuracies of the three traits MKWC, HKWC, and DR were not significantly different among the twelve models. The prediction accuracy of GY was significantly different among the 12 models. For example, the linear model Bayes C had a higher prediction accuracy of 0.491 for GY. In contrast, the ML model ANN had the lowest prediction accuracy of 0.386 for GY, and the prediction accuracy of ANN for MKWC was also the lowest, 0.684.

Table 3. Prediction accuracy of maize kernel water content-related traits and yield.

Genetic Model	Prediction Method	MKWC	HKWC	DR	GY
G: A	GBLUP	0.788 ± 0.065	0.846 ± 0.073	0.693 ± 0.115	0.463 ± 0.123
	rrBLUP	0.797 ± 0.088	0.861 ± 0.052	0.700 ± 0.098	0.445 ± 0.121
	LASSO	0.791 ± 0.075	0.848 ± 0.062	0.702 ± 0.123	0.433 ± 0.145
	Bayes A	0.816 ± 0.065	0.850 ± 0.045	0.714 ± 0.109	0.424 ± 0.143
	Bayes B	0.793 ± 0.067	0.853 ± 0.071	0.703 ± 0.107	0.455 ± 0.158
	Bayes C	0.799 ± 0.072	0.850 ± 0.065	0.682 ± 0.103	0.491 ± 0.134
	PLS	0.767 ± 0.041	0.839 ± 0.031	0.673 ± 0.094	0.423 ± 0.140
	GLM	0.787 ± 0.075	0.844 ± 0.058	0.716 ± 0.081	0.435 ± 0.095
	RF	0.739 ± 0.064	0.801 ± 0.067	0.708 ± 0.071	0.510 ± 0.098
	SVM	0.781 ± 0.069	0.863 ± 0.065	0.709 ± 0.061	0.473 ± 0.099
	GBM	0.758 ± 0.088	0.811 ± 0.067	0.676 ± 0.093	0.431 ± 0.096
	ANN	0.684 ± 0.121	0.811 ± 0.048	0.685 ± 0.105	0.386 ± 0.098
G: AD	RHKS	0.793 ± 0.074	0.843 ± 0.061	0.730 ± 0.103	0.432 ± 0.138
G: A(AA)	RHKS	0.811 ± 0.084	0.874 ± 0.051	0.751 ± 0.075	0.486 ± 0.151
G: AD(AA)	RHKS	0.810 ± 0.055	0.873 ± 0.049	0.747 ± 0.111	0.520 ± 0.116
GCA: A		0.801 ± 0.062	0.851 ± 0.052	0.721 ± 0.086	0.481 ± 0.114
GCA: AD		0.754 ± 0.074	0.864 ± 0.047	0.664 ± 0.120	0.409 ± 0.113
GCA: A(AA _{SN})		0.794 ± 0.072	0.867 ± 0.048	0.728 ± 0.086	0.421 ± 0.165
GCA: AD(AA _{SN})		0.773 ± 0.072	0.851 ± 0.068	0.695 ± 0.111	0.415 ± 0.127
GCA: AD(AA _S)(AA _N)(AA _H)		0.778 ± 0.084	0.849 ± 0.059	0.701 ± 0.098	0.465 ± 0.113

The model was extended based on hybrid genotypes to include dominance and epistatic effects based on the RHKS model. Based on the results in Table 3, it can be found that the prediction accuracy of MKWC and DR was not improved in the prediction model, considering dominance and epistatic effects. In contrast, the prediction accuracy of the G: A(AA) and G:AD(AA) models for HKWC were improved to 0.874 and 0.873, respectively. The prediction accuracies of the two traits, MKWC and DR, in the models with the addition of dominance and epistatic effects were similar to those of the linear and ML models. The average prediction accuracy for GY improved to 0.479 after considering both dominance and epistatic effects compared to the model with only additive effects. The prediction accuracy of GY in the model with both dominance and epistatic effects (G:AD(AA)) reached 0.520, which was higher than that of the linear and ML models.

3.3. Prediction Accuracy of Prediction Models Based on Parental Genotypes

Several nested models were tested by successfully incorporating dominance and epistatic effects into the additive effects based on the genotypes of the two parents. According to the results in Table 3, it was found that the prediction models based on parental genotypes did not improve the prediction accuracy when incorporating dominance and epistatic effects. In the GCA model, the average prediction accuracy of MKWC was 0.781. The average prediction accuracies of the remaining three traits, HKWC, DR, and GY, were 0.858, 0.702, and 0.436, respectively. The prediction accuracies of the GCA model were all reduced compared to the G model.

3.4. The Potential of Trait-Specific SNPs in Genomic Prediction

Figure 2A,B shows the p -values of the markers obtained by the MLM GWAS method and the p -value comparison of SNPs (MKWC) in the *ts_SNPs* dataset and the *random_SNPs* dataset. For other traits, see Figures S2 and S3. In our study, for all four traits, prediction accuracies improved when using the top 100 SNPs with the largest effect on trait markers, with accuracies of 0.811, 0.867, 0.737, and 0.541, respectively (Figure 2C–F). When predicting with randomly selected SNPs (*random_SNPs*), prediction accuracies of MKWC, HKWC, DR, and GY were lower than that of *ts_SNPs*. When the *ts_SNPs* marker set was modeled as a fixed effect, prediction accuracy decreased slightly as the number of SNPs increased but was mostly better than *random_SNPs*.

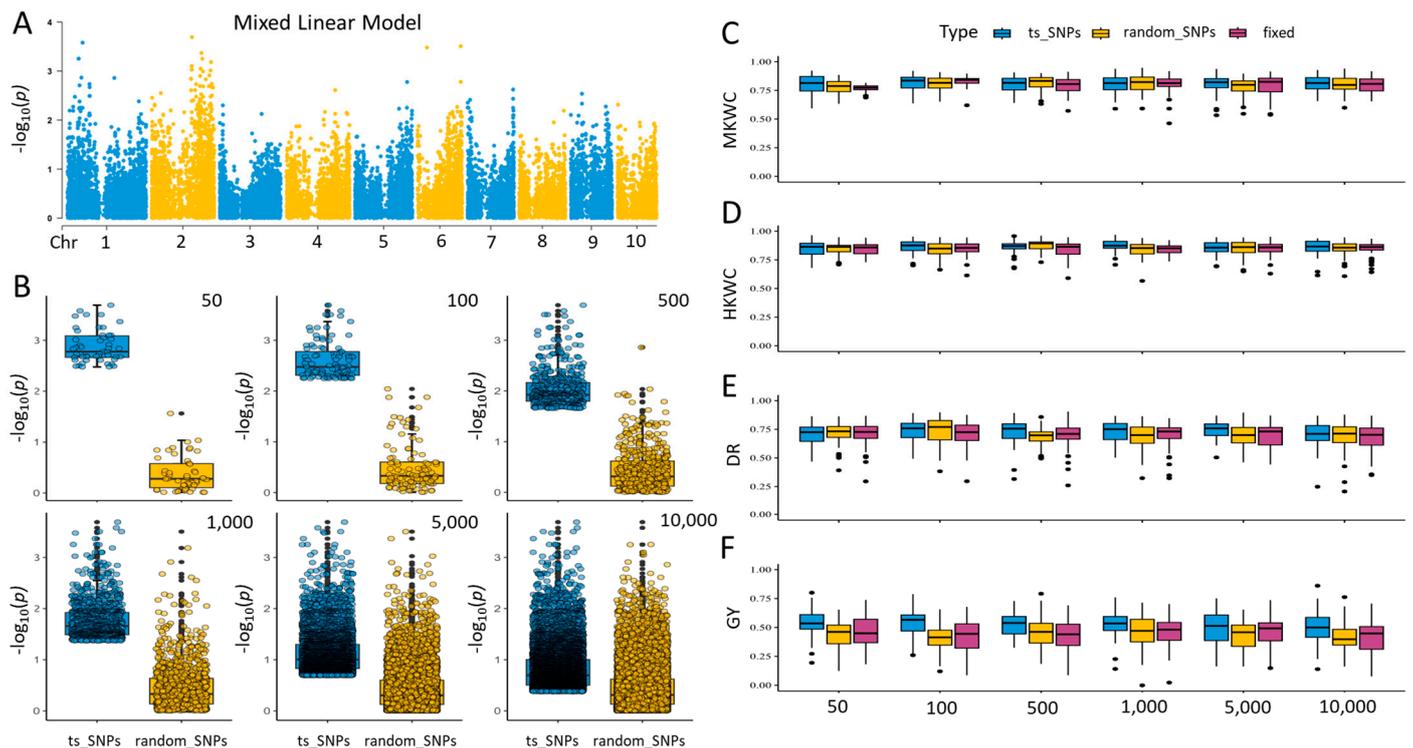


Figure 2. The SNP marker datasets and prediction accuracy. Manhattan plots for MKWC were obtained from MLM GWAS methods (A). The comparison of $-\log_{10}(P)$ value of SNPs (MKWC) from *ts_SNPs* dataset and *random_SNPs* dataset (B). Prediction accuracy was obtained using different SNP datasets to predict MKWC, HKWC, DR, and GY (C–F).

In addition, where GWAS was performed for four traits in each environment, and using environmentally specific SNPs (*ts_SNPs* = 100) as genotypes improved prediction accuracy for all of them compared to the *random_SNPs* model (Figure S4).

3.5. Identification of 19 Single-Cross Combinations with Low MKWC and Higher GY

Based on the predicted performance of 561 single-cross combinations, we selected 30 promising combinations with low MKWC (Table S2, Figure 3). These data were used to compare the selection gain of MKWC_TOP30 with the remaining combinations. The mean value of MKWC_TOP30 was 36.28%, while the mean value of MKWC in others was 38.34%, which was significantly lower than MKWC with all the remaining combinations. Interestingly, based on the yield performance of MKWC_TOP30, the yields of these combinations were also higher than the yields of all the remaining combinations (Figure 3D). MKWC was reduced by 1.44% in MKWC_TOP30 compared with Xianyu335 (Figure 3). Among these TOP 30 hybrids, the performance of 11 hybrids was evaluated in the field (Table S3, Figure S5), and the other 19 novel hybrids were predicted and selected that have not yet

been tested. Therefore, these hybrids are promising candidates for the next evaluation stage in field trials.

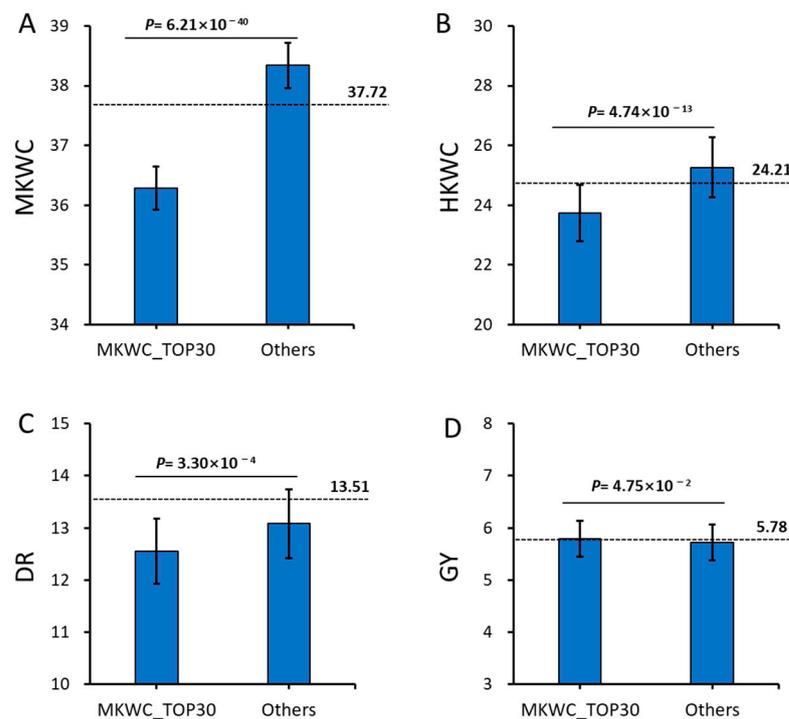


Figure 3. Comparison of predicted breeding values of 561 hybrid combinations. Bar chart showing the predicted 30 hybrids with the lowest MKWC compared to the remaining hybrids (A) and the predicted breeding values of HKWC, DR, and GY obtained for these 30 hybrids were also compared to the rest of the hybrids (B–D). P values were calculated from a two-tailed Student’s *t*-test. The Figure’s dashed line indicates the breeding values for the control Xianyu335.

4. Discussion

4.1. Accurate Identification of KWC in Northern China

The development and application of maize varieties with low HKWC is a prerequisite to ensure that mechanical harvesting can be applied. The use of whole genomic prediction technology holds practical importance for the rapid selection of low KWC maize varieties. Because KWC is controlled by multiple small-effect QTLs and is highly influenced by environmental conditions, one of the biggest challenges lies in obtaining accurate KWC values, particularly at the kernel’s physiological maturity stage. A number of methods have been developed to measure the KWC. For example, moisture determination indicators by detecting changes in capacitance [62], the SK-300 moisture-determination meter [63], the digital timber-moisture meter [64], the wooden moisture meter Voltcraft FM-200 Humidity [65], the grain moisture meter with microwave attenuation at 10.5 GHz [66], and the hand-held moisture meter [67]. Although the classical drying method is regarded as the most reliable for determining MKWC values, it is labor-intensive and time-consuming, making it unsuitable for large-scale phenotyping in genomic prediction studies. To balance precision and practicality, we chose the Protimeter Moisture Meter (BLD5605) for the phenotypic acquisition of MKWC. This method provided reliable readings after calibration with corresponding temperature and humidity data, ensuring accuracy without the excessive resource demands of classical methods.

Considering that MKWC was mainly influenced by varietal factors (genetics), we recorded MKWC during the maturity period of kernel development of each hybrid to minimize variability caused by differing maturation stages. While environmental factors significantly affect MKWC, we standardized the harvesting period to ensure a consistent dehydration process and reduce environmental errors. Our current results indicated that

there were significant correlations among the four growing locations for KWC-related traits and also a high heritability of 0.75–0.85. These findings suggest that our methodological choices, including the Protimeter Moisture Meter and standardizing the harvest period, effectively addressed the trade-offs between precision, efficiency, and practicality. This allowed us to obtain accurate KWC values, which were instrumental in achieving the desired genomic prediction accuracy.

4.2. Influence of Different Statistical Models on the Prediction Accuracy of Traits

Prediction accuracy varied widely among agronomic traits in this study, with HWKC having the highest prediction accuracy of 0.876 and GY having the lowest prediction accuracy of 0.386. It was evident from the heritability that the prediction accuracy reaching large variations was significantly affected by different genetic structures, as shown in previous studies [68,69]. We used different model predictions for comparative analysis, and the results showed that for traits with higher heritability, statistical models had less effect on prediction accuracy, such as MKWC, HWKC, and DR, which reached similar prediction accuracy with all models (Table 3). For the lower heritability trait GY, the prediction accuracy of different statistical models varied considerably. For example, the lowest prediction accuracy of ANN was 0.386 for GY, and the highest prediction accuracy of the G:AD(AA) model was 0.520, which was an improvement of 0.134. Therefore, our results were consistent with previous studies that the prediction accuracy was mainly determined by the genetic structure [53]. For traits with low heritability, larger sample sizes or multi-environment experimental data can be used to improve the generalization ability of the model. However, by using several different models for prediction, it is possible to select the optimal model for different traits. For example, the highest prediction accuracy of Bayes A was 0.816 for MKWC.

According to earlier studies [70–73], trait-specific markers have demonstrated advantages for genomic prediction by capturing key genetic signals. In this study, we tested various subsets of specific SNPs using the G:A (RKHS) model, incorporating specific SNPs as fixed effects and the remaining SNPs as random effects. Consistent with previous findings [73,74], certain combinations significantly increased prediction accuracy. These results reinforced the importance of leveraging trait-specific markers to enhance accuracy, particularly in breeding programs targeting complex traits. By aligning with earlier studies while applying these insights to the unique genetic and environmental context of maize breeding, our findings provided further evidence for the efficacy of this approach and offered a practical framework for improving molecular breeding efficiency.

4.3. Prediction Accuracy and Practical Application of Genomic Prediction for Hybrid Performance

In this study, we evaluated the prediction accuracy of hybrid performance using two types of genetic models: one assuming a single genetic effect for the hybrid (G model) and another defining genetic effect based on allele origin (GCA model). Additive, dominance, and epistatic effects were considered. The results showed that the G model outperformed all GCA models across the four traits, with the G:AD(AA) model achieving the highest prediction accuracy for GY when non-additive effects were included. While previous studies have reported similar predictive power between G and GCA models in hybrid populations [75,76], our findings suggest a clear advantage of the G model for trait prediction. The lower performance of the GCA model in predicting hybrids may stem from its over-reliance on the general combining ability of the parents while neglecting the unique non-additive effects in hybrids. In this study, the G model more comprehensively captured the genetic potential of hybrid combinations by integrating additive, dominance, and epistatic effects, thereby demonstrating higher prediction accuracy. Nonetheless, the GCA model remains valuable for estimating parental line values and supporting the selection of next-cycle parental lines.

Results from many previous studies have demonstrated the potential of genomic prediction in crop breeding, particularly in maize hybrid breeding [53]. For example, GP

has been successfully applied to traits such as GY and KWC in maize in America and Europe, highlighting its versatility in different environmental contexts and breeding programs [52,77,78]. Additionally, in wheat and rice, GP has been employed to improve traits such as grain quality, disease resistance, and abiotic stress tolerance in regions ranging from South Asia to South America [79–82]. These studies underline the global applicability and adaptability of GP models across species and regions. In this study, utilizing GP, we identified MKWC_TOP30 hybrids with the lowest MKWC, with an average MKWC of 36.28%. These hybrids also exhibited high GY, demonstrating the efficiency of this approach in selecting hybrids suited for mechanized maize harvesting. The GP model, developed using hybrid populations specifically from northern China, has significant practical implications. It not only reduces costs and time in the breeding process but also directly supports the adoption of mechanized harvesting by enabling the rapid selection of hybrids with optimal moisture content. Compared with previous studies conducted in regions like America, where kernel dry-down rates have been improved through both genomic and phenotypic selection, our findings emphasize the necessity of tailoring prediction models to specific regional climates. In northern China, where rapid temperature drops hinder natural kernel drying, the ability to predict and select hybrids with low MKWC was particularly impactful. These results advanced the regional breeding strategies for northern China and contributed to global efforts in optimizing crop performance.

5. Conclusions

Our study aimed to identify superior low MKWC hybrids and evaluate the potential of genomic prediction using maize hybrids from northern China. The results showed that prediction models had minimal impact on trait prediction accuracy, which largely depended on the genetic basis of the trait. Combining genetic effects and trait-specific markers significantly improved yield prediction accuracy. These findings highlighted the potential of GP to enhance breeding efficiency by enabling precise selection of hybrids with low MKWC and high yield, reducing the time and costs of traditional phenotypic selection. This approach was particularly valuable for supporting mechanized harvesting in northern China, boosting agricultural efficiency and farmer profitability. Moreover, the study contributed to understanding the genetic basis of key traits and offered a foundation for applying GP to other crops and regions with similar challenges.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/agronomy14122795/s1>, Figure S1: Normal distribution diagram of HKWC (A), DR (B), and GY (C). Figure S2: Manhattan plot of marker effect size distribution of HKWC (A), DR (B), and GY (C) obtained through GWAS. Figure S3: The comparison of $-\log^{10}(P)$ value of SNPs from the *ts_SNP*s dataset and *random_SNP*s dataset for the remaining three traits, HKWC(A), DR(B), and GY(C), respectively. Figure S4: Prediction accuracy of two SNP datasets (Number of SNP = 100) for four traits in a single environment. Figure S5: Promising hybrid combinations that have not been generated in the breeding program so far. Table S1: Basic description and correlation analysis of four traits. Table S2: Performance of the best 30 predicted hybrid combinations (Predicted values by GBLUP). Table S3: The kernel water content at maturity (%) of the best 11 tested hybrid combinations.

Author Contributions: Methodology, P.L., R.Y., M.L., J.W., D.Z., Z.Z., J.H., X.Z., J.F., X.L., Z.H. and Z.N.; software, P.L. and J.F.; investigation, P.L., R.Y., L.Z., J.Y., H.W., H.Y., R.Z., W.L., F.W., W.G., X.X. and K.Y.; writing—original draft preparation, P.L., R.Y., X.Z. and J.F.; writing—review and editing, P.L., H.Y., M.L., J.W., D.Z., Z.Z., J.H., X.Z., J.F., X.L., Z.H. and Z.N.; resources, L.Z., J.Y. and W.G.; data curation, L.Z., H.W., R.Z., W.L., F.W., X.X. and K.Y.; formal analysis, J.Y., visualization, W.L. and F.W.; conceptualization, X.Z., X.L., Z.H. and Z.N.; project administration, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Sustainable Development International Cooperation Program from the Bill & Melinda Gates Foundation (2022YFAG1002), the Sci-Tech Innovation 2030 Program of China (2023ZD0402804), the National Key Research and Development Program of China (2020YFE0202300), Agricultural Science and Technology Innovation Program of CAAS, and the China Scholarship Council. The funding agencies were not involved in the design of the study; the collection, analysis, and interpretation of the data; or the preparation of the manuscript.

Data Availability Statement: All other relevant data and R codes are available from the corresponding authors upon request.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Xie, R.Z.; Ming, B.; Gao, S.; Wang, K.; Hou, P.; Li, S. Current state and suggestions for mechanical harvesting of corn in China. *J. Integr. Agric.* **2022**, *21*, 892–897. [\[CrossRef\]](#)
- Brooking, I.R. Maize ear moisture during grain-filling, and its relation to physiological maturity and grain-drying. *Field Crop. Res.* **1990**, *23*, 55–68. [\[CrossRef\]](#)
- Schmidt, J.L.; Hallauer, A.R. Estimating harvest date of corn in the field. *Crop. Sci.* **1966**, *6*, 227–231. [\[CrossRef\]](#)
- Johnson, D.R.; Tanner, J.W. Calculation of the rate and duration of grain filling in corn (*Zea mays* L.). *Crop. Sci.* **1972**, *12*, 485–486. [\[CrossRef\]](#)
- Sala, R.G.; Andrade, F.H.; Camadro, E.L.; Ceron, J.C. Quantitative trait loci for grain moisture at harvest and field grain drying rate in maize (*Zea mays* L.). *Theor. Appl. Genet.* **2006**, *112*, 462–471. [\[CrossRef\]](#)
- Nass, H.G.; Crane, P.L. Effect of endosperm mutants on drying rate in corn (*Zea mays* L.). *Crop. Sci.* **1970**, *10*, 141–144. [\[CrossRef\]](#)
- Mišević, D.; Alexander, D.E.; Dumanović, J.; Kerečki, B.; Ratković, S. Grain moisture loss rate of high-oil and standard-oil maize hybrids. *Agron. J.* **1988**, *80*, 841–845. [\[CrossRef\]](#)
- Crane, P.L.; Miles, S.R.; Newman, J.E. Factors associated with varietal differences in rate of field drying in corn. *Agron. J.* **1959**, *51*, 318–320. [\[CrossRef\]](#)
- Purdy, J.L.; Crane, P.L. Influence of pericarp on differential drying rate in “mature” corn (*Zea mays* L.). *Crop. Sci.* **1967**, *7*, 379–381. [\[CrossRef\]](#)
- Cavaliere, A.J.; Smith, O.S. Grain filling and field drying of a set of maize hybrids released from 1982. *Crop. Sci.* **1985**, *25*, 856–860. [\[CrossRef\]](#)
- Hicks, D.R.; Geadelmann, G.L.; Peterson, R.H. Drying rates of frosted maturing maize. *Agron. J.* **1976**, *68*, 452–455. [\[CrossRef\]](#)
- Sweeney, P.M.; Martin, S.K.S.; Clucas, C.P. Indirect inbred selection to reduce grain moisture in maize hybrids. *Crop. Sci.* **1994**, *34*, 391–396. [\[CrossRef\]](#)
- Troyer, A.F.; Ambrose, W.B. Plant characteristics affecting field drying rate of ear corn. *Crop. Sci.* **1971**, *11*, 529–531. [\[CrossRef\]](#)
- Beavis, W.D.; Smith, O.S.; Grant, D.; Fincher, R. Identification of quantitative trait loci using a small sample of topcrossed and F₄ progeny from maize. *Crop. Sci.* **1994**, *34*, 882–896. [\[CrossRef\]](#)
- Melchinger, A.E.; Utz, H.F.; Schön, C.C. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics* **1998**, *149*, 383–403. [\[CrossRef\]](#)
- Austin, D.F.; Lee, M.; Veldboom, L.R.; Hallauer, A.R. Genetic mapping in maize with hybrid progeny across testers and generations: Grain yield and grain moisture. *Crop. Sci.* **2000**, *40*, 30–39. [\[CrossRef\]](#)
- Ho, J.; McCouch, S.; Smith, M. Improvement of hybrid yield by advanced backcross QTL analysis in elite maize. *Theor. Appl. Genet.* **2002**, *105*, 440–448. [\[CrossRef\]](#)
- Kebede, A.Z.; Woldemariam, T.; Reid, L.M.; Harris, L.J. Quantitative trait loci mapping for Gibberella ear rot resistance and associated agronomic traits using genotyping-by-sequencing in maize. *Theor. Appl. Genet.* **2016**, *129*, 17–29. [\[CrossRef\]](#)
- Song, W.; Shi, Z.I.; Xing, J.; Duan, M.; Su, A.; Li, C.; Zhang, R.; Zhao, Y.; Luo, M.; Wang, J.; et al. Molecular mapping of quantitative trait loci for grain moisture at harvest in maize. *Plant Breed.* **2017**, *136*, 28–32. [\[CrossRef\]](#)
- Zhou, G.; Hao, D.; Xue, L.; Chen, G.; Lu, H.; Zhang, Z.; Shi, M.; Huang, X.L.; Mao, Y. Genome-wide association study of kernel moisture content at harvest stage in maize. *Breed. Sci.* **2018**, *68*, 622–628. [\[CrossRef\]](#)
- Liu, J.; Yu, H.; Liu, Y.; Deng, S.; Liu, Q.; Liu, B.; Xu, M. Genetic dissection of grain water content and dehydration rate related to mechanical harvest in maize. *BMC Plant Biol.* **2020**, *20*, 118. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yin, S.; Liu, J.; Yang, T.; Li, P.; Xu, Y.; Fang, H.; Xu, S.; Wei, J.; Xue, L.; Hao, D.; et al. Genetic analysis of the seed dehydration process in maize based on a logistic model. *Crop. J.* **2020**, *8*, 182–193. [\[CrossRef\]](#)
- Zhang, J.; Zhang, F.Q.; Tang, B.J.; Ding, Y.; Xia, L.K.; Qi, J.S.; Mu, X.Y.; Gu, L.M.; Lu, D.W.; Chen, Y.H. Molecular mapping of quantitative trait loci for grain moisture at harvest and field grain drying rate in maize (*Zea mays* L.). *Physiol. Plant.* **2020**, *169*, 64–72. [\[CrossRef\]](#)
- Li, W.; Yu, Y.; Wang, L.; Luo, Y.; Peng, Y.; Xu, Y.; Liu, X.; Wu, S.; Jian, L.; Xu, J.; et al. The genetic architecture of the dynamic changes in grain moisture in maize. *Plant Biotechnol. J.* **2021**, *19*, 1195–1205. [\[CrossRef\]](#) [\[PubMed\]](#)

25. Xu, Y.; Zhang, X.; Li, H.; Zheng, H.; Zhang, J.; Olsen, M.; Varshney, R.; Prasanna, B.M.; Qian, Q. Smart breeding driven by big data, artificial intelligence, and integrated genomic-environmental prediction. *Mol. Plant* **2022**, *15*, 1664–1695. [[CrossRef](#)]
26. Crossa, J.; Perez-Rodriguez, P.; Cuevas, J.; Montesinos-Lopez, O.; Jarquin, D.; De Los Campos, G.; Burgueno, J.; Gonzalez-Camacho, J.M.; Perez-Elizalde, S.; Beyene, Y.; et al. Genomic selection in plant breeding: Methods, models, and perspectives. *Trends Plant Sci.* **2017**, *22*, 961–975. [[CrossRef](#)] [[PubMed](#)]
27. Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **2001**, *157*, 1819–1829. [[CrossRef](#)]
28. VanRaden, P.M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **2008**, *91*, 4414–4423. [[CrossRef](#)] [[PubMed](#)]
29. Karkkainen, H.P.; Sillanpaa, M.J. Back to basics for Bayesian model building in genomic selection. *Genetics* **2012**, *191*, 969–987. [[CrossRef](#)]
30. Perez, P.; De los Campos, G. Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **2014**, *198*, 483–495. [[CrossRef](#)]
31. Piepho, H.P. Ridge regression and extensions for genome wide selection in maize. *Crop. Sci.* **2009**, *49*, 1165–1176. [[CrossRef](#)]
32. Endelman, J.B. Ridge regression and other kernels for genomic selection with R package rrBLUP. *Plant Genome* **2011**, *4*, 255–258. [[CrossRef](#)]
33. Friedman, J.; Hastie, T.; Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **2010**, *33*, 1–22. [[CrossRef](#)]
34. Li, Z.T.; Sillanpaa, M.J. Overview of LASSO-related penalized regression methods for quantitative trait mapping and genomic selection. *Theor. Appl. Genet.* **2012**, *125*, 419–435. [[CrossRef](#)]
35. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
36. Chen, X.; Ishwaran, H. Random forests for genomic data analysis. *Genomics* **2012**, *99*, 323–329. [[CrossRef](#)]
37. Holliday, J.A.; Wang, T.L.; Aitken, S. Predicting adaptive phenotypes from multilocus genotypes in sitka spruce (*Picea sitchensis*) Using random forest. *G3 Genes Genomes Genet.* **2012**, *2*, 1085–1093. [[CrossRef](#)]
38. Maenhout, S.; De Baets, B.; Haesaert, G.; Bockstaele, E.V. Support vector machine regression for the prediction of maize hybrid performance. *Theor. Appl. Genet.* **2007**, *115*, 1003–1013. [[CrossRef](#)] [[PubMed](#)]
39. Montesinos-Lopez, O.A.; Martin-Vallejo, J.; Crossa, J.; Gianola, D.; Hernandez-Suarez, C.M.; Montesinos-Lopez, A.; Juliana, P.; Singh, R. A Benchmarking between deep learning, support vector machine and Bayesian threshold best linear unbiased prediction for predicting ordinal traits in plant breeding. *G3 Genes Genomes Genet.* **2019**, *9*, 601–618. [[CrossRef](#)]
40. Gonzalez-Camacho, J.M.; De los Campos, G.; Perez, P.; Gianola, D.; Cairns, J.E.; Mahuku, G.; Babu, R.; Crossa, J. Genome-enabled prediction of genetic values using radial basis function neural networks. *Theor. Appl. Genet.* **2012**, *125*, 759–771. [[CrossRef](#)]
41. Yan, J.; Xu, Y.T.; Cheng, Q.; Jiang, S.Q.; Wang, Q.; Xiao, Y.J.; Ma, C.; Yan, J.B.; Wang, X.F. LightGBM: Accelerated genomically designed crop breeding through ensemble learning. *Genome Biol.* **2021**, *22*, 271. [[CrossRef](#)] [[PubMed](#)]
42. Cui, Z.; Dong, H.; Zhang, A.; Ruan, Y.; He, Y.; Zhang, Z. Assessment of the potential for genomic selection to improve husk traits in maize. *G3 Genes Genomes Genet.* **2020**, *10*, 3741–3749. [[CrossRef](#)]
43. Labroo, M.R.; Ali, J.; Aslam, M.U.; de Asis, E.J.; Dela Paz, M.A.; Sevilla, M.A.; Lipka, A.E.; Studer, A.J.; Rutkoski, J.E. Genomic prediction of yield traits in single-cross hybrid rice (*Oryza sativa* L.). *Front. Genet.* **2021**, *12*, 692870. [[CrossRef](#)]
44. Michel, S.; Löschenberger, F.; Ametz, C.; Pachler, B.; Sparr, E.; Bürstmayr, H. Simultaneous selection for grain yield and protein content in genomics-assisted wheat breeding. *Theor. Appl. Genet.* **2019**, *132*, 1745–1760. [[CrossRef](#)] [[PubMed](#)]
45. Tong, H.; Nikoloski, Z. Machine learning approaches for crop improvement: Leveraging phenotypic and genotypic big data. *J. Plant Physiol.* **2020**, *257*, 153354. [[CrossRef](#)]
46. Wang, X.; Xu, Y.; Hu, Z.; Xu, C. Genomic selection methods for crop improvement: Current status and prospects. *Crop. J.* **2018**, *6*, 330–340. [[CrossRef](#)]
47. Zhang, A.; Pérez-Rodríguez, P.; San Vicente, F.; Palacios-Rojas, N.; Dhliwayo, T.; Liu, Y.; Cui, Z.; Guan, Y.; Wang, H.; Zheng, H.; et al. Genomic prediction of the performance of hybrids and the combining abilities for line by tester trials in maize. *Crop. J.* **2022**, *10*, 109–116. [[CrossRef](#)]
48. Zhu, X.; Chi, R.; Du, Y.; Qin, J.; Xiong, Z.; Zhang, W.; Li, X. Experimental study on the key factors of low-loss threshing of high-moisture maize. *Int. J. Agric. Biol. Eng.* **2020**, *13*, 23–31. [[CrossRef](#)]
49. Campos, H.; Cooper, M.; Edmeades, G.O.; Löffler, C.; Schussler, J.R.; Ibañez, M. Changes in drought tolerance in maize associated with fifty years of breeding for yield in the US corn belt. *Maydica* **2006**, *51*, 369.
50. Li, Y.; Tao, C.; Zhe, Q.; Li, K.; Yin, X.; Han, D.; Yan, B.; Zhao, D.; Zhang, D. Development and application of mechanized maize harvesters. *Int. J. Agric. Biol. Eng.* **2016**, *9*, 15–28. [[CrossRef](#)]
51. Chu, Z.; Ming, B.; Li, L.; Xue, J.; Zhang, W.; Hou, L.; Xie, R.; Hou, P.; Wang, K.; Li, S. Dynamics of maize grain drying in the high latitude region of Northeast China. *J. Integr. Agric.* **2022**, *21*, 365–374. [[CrossRef](#)]
52. Ni, P.; Anche, M.T.; Ruan, Y.; Dang, D.; Morales, N.; Li, L.; Liu, M.; Wang, S.; Robbins, K.R. Genomic prediction strategies for dry-down-related traits in maize. *Front. Plant Sci.* **2022**, *13*, 930429. [[CrossRef](#)] [[PubMed](#)]
53. Luo, P.; Wang, H.W.; Ni, Z.Y.; Yang, R.S.; Wang, F.; Yong, H.J.; Zhang, L.; Zhou, Z.Q.; Song, W.; Li, M.S.; et al. Genomic prediction of yield performance among single-cross maize hybrids using a partial diallel cross design. *Crop. J.* **2023**, *10*, 109–116. [[CrossRef](#)]
54. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **2015**, *67*, 1–48. [[CrossRef](#)]

55. Hallauer, A.R.; Carena, M.J.; Miranda Filho, J.D. *Quantitative Genetics in Maize Breeding*; Springer: New York, NY, USA, 2010. [[CrossRef](#)]
56. Guo, Z.; Wang, H.; Tao, J.; Ren, Y.; Xu, C.; Wu, K.; Zou, C.; Zhang, J.; Xu, Y. Development of multiple SNP marker panels affordable to breeders through genotyping by target sequencing (GBTS) in maize. *Mol. Breed.* **2019**, *39*, 37. [[CrossRef](#)]
57. González-Diéguez, D.; Legarra, A.; Charcosset, A.; Moreau, L.; Lehermeier, C.; Teyssèdre, S.Z.; Vitezica, G. Genomic prediction of hybrid crops allows disentangling dominance and epistasis. *Genetics* **2021**, *218*, iyab026. [[CrossRef](#)]
58. Vitezica, Z.G.; Legarra, A.; Toro, M.A.; Varona, L. Orthogonal estimates of variances for additive, dominance, and epistatic effects in populations. *Genetics* **2017**, *206*, 1297–1307. [[CrossRef](#)]
59. Covarrubias-Pazarán, G. Genome-assisted prediction of quantitative traits using the R package sommer. *PLoS ONE* **2016**, *11*, e0156744. [[CrossRef](#)]
60. López, O.A.M.; González, B.A.M.; López, A.M.; Crossa, J. Statistical Machine-Learning Methods for Genomic Prediction Using the SKM Library. *Genes* **2023**, *14*, 1003. [[CrossRef](#)]
61. Bernardo, R. Genome wide selection when major genes are known. *Crop. Sci.* **2014**, *54*, 68–75. [[CrossRef](#)]
62. Reid, L.M.; Zhu, X.; Morrison, M.J.; Woldemariam, T.; Voloaca, C.; Wu, J.H.; Xiang, K. A non-destructive method for measuring maize kernel moisture in a breeding program. *Maydica* **2010**, *55*, 163–171.
63. Qian, Y.L.; Zhang, X.Q.; Wang, L.F.; Chen, J.; Chen, B.R.; Lv, G.H.; Wu, Z.C.; Guo, J.; Wang, J.; Qi, Y.C.; et al. Detection of QTLs controlling fast kernel dehydration in maize (*Zea mays* L.). *Genet. Mol. Res.* **2016**, *15*, 3. [[CrossRef](#)]
64. Yang, J.; Carena, M.J.; Uphaus, J. Area Under the Dry Down Curve (AUDDC): A method to evaluate rate of dry down in maize. *Crop. Sci.* **2010**, *50*, 2347–2354. [[CrossRef](#)]
65. Filipenco, A.; Mandache, V.; Valsan, G.; Ivan, F.; Ciocazanu, I. Inheritance of grain dry-down in corn (*Zea mays* L.). *Bull. Univ. Agric. Sci. Veter. Med. Cluj-Napoca. Agric.* **2013**, *70*, 223–226. [[CrossRef](#)]
66. Kim, K.B.; Noh, S.H.; Kim, J.H. Development of grain moisture meter using microwave attenuation at 10.5 GHz and moisture density. *IEEE Trans. Instrum. Meas.* **2000**, *51*, 72–77. [[CrossRef](#)]
67. Freppon, J.T.; Martin, S.K.S.; Pratt, R.C.; Henderlong, P.R. Selection for low ear moisture in corn, using a hand-held meter. *Crop. Sci.* **1992**, *32*, 1062–1064. [[CrossRef](#)]
68. Casa, A.M.; Mitchell, S.E.; Smith, O.S.; Register, J.C.; Wessler, S.R.; Kresovich, S. Evaluation of Hbr (MITE) markers for assessment of genetic relationships among maize (*Zea mays* L.) inbred lines. *Theor. Appl. Genet.* **2002**, *104*, 104–110. [[CrossRef](#)] [[PubMed](#)]
69. Guo, T.; Li, H.; Yan, J.; Tang, J.; Li, J.; Zhang, Z.; Zhang, L.; Wang, J. Performance prediction of F1 hybrids between recombinant inbred lines derived from two elite maize inbred lines. *Theor. Appl. Genet.* **2013**, *126*, 189–201. [[CrossRef](#)]
70. Zhang, Z.; Liu, J.F.; Ding, X.D.; Bijma, P.; de Koning, D.J.; Zhang, Q. Best linear unbiased prediction of genomic breeding values using a trait-specific marker-derived relationship matrix (genomic selection using TABLUP). *PLoS ONE* **2010**, *5*, e12648. [[CrossRef](#)]
71. Bhandari, A.; Bartholome, J.; Hamadoun, T.V.; Kumari, N.; Frouin, J.; Kumar, A.; Ahmadi, N. Selection of trait-specific markers and multi environment models improve genomic predictive ability in rice. *PLoS ONE* **2019**, *14*, e0208871. [[CrossRef](#)]
72. Zhao, Y.S.; Mette, M.F.; Gowda, M.; Longin, C.F.H.; Reif, J.C. Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. *Heredity* **2014**, *112*, 638–645. [[CrossRef](#)] [[PubMed](#)]
73. Zhang, Z.; Erbe, M.; He, J.L.; Ober, U.; Gao, N.; Zhang, H.; Simianer, H.; Li, J.Q. Accuracy of whole-genome prediction using a genetic architecture-enhanced variance-covariance matrix. *G3 Genes | Genomes | Genet.* **2015**, *5*, 615–627. [[CrossRef](#)]
74. Li, G.; Dong, Y.; Zhao, Y.; Tian, X.; Würschum, T.; Xue, J.; Chen, S.; Reif, J.; Xu, S.; Liu, W. Genome-wide prediction in a hybrid maize population adapted to Northwest China. *Crop. J.* **2020**, *8*, 830–842. [[CrossRef](#)]
75. Technow, F.; Schrag, T.A.; Schipprack, W.; Bauer, E.; Simianer, H.; Melchinger, A.E. Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. *Genetics* **2014**, *197*, 1343–1355. [[CrossRef](#)] [[PubMed](#)]
76. Alves, F.C.; Granato, Í.S.C.; Galli, G.; Lyra, D.H.; Fritsche-Neto, R.; De Los Campos, G. Bayesian analysis and prediction of hybrid performance. *Plant Methods* **2019**, *7*, 1–18. [[CrossRef](#)] [[PubMed](#)]
77. Millet, E.J.; Kruijer, W.; Coupel-Ledru, A.; Prado, S.A.; Cabrera-Bosquet, L.; Lacube, S.; Charcosset, A.; Welcker, C.; Eeuwijk, F.; Tardieu, F. Genomic prediction of maize yield across European environmental conditions. *Nat. Genet.* **2019**, *51*, 952–956. [[CrossRef](#)]
78. Gunundu, R.; Shimelis, H.; Mashilo, J. Genomic selection and enablers for agronomic traits in maize (*Zea mays*): A review. *Plant Breed.* **2023**, *142*, 573–593. [[CrossRef](#)]
79. Bhavani, S.; Singh, P.K.; Qureshi, N.; He, X.; Biswal, A.K.; Juliana, P.; Dababat, A.; Mourad, A.M.I. Globally Important wheat diseases: Status, challenges, breeding and genomic tools to enhance resistance durability. In *Genomic Designing for Biotic Stress Resistant Cereal Crops*; Kole, C., Ed.; Springer: Cham, Switzerland, 2021; pp. 59–128. [[CrossRef](#)]
80. Cui, Y.; Li, R.; Li, G.; Zhang, F.; Zhu, T.; Zhang, Q.; Ali, J.; Li, Z.; Xu, S. Hybrid breeding of rice via genomic selection. *Plant Biotechnol. J.* **2020**, *18*, 57–67. [[CrossRef](#)]

81. Dar, M.H.; Bano, D.A.; Waza, S.A.; Zaidi, N.W.; Majid, A.; Shikari, A.B.; Ahangar, M.A.; Hossain, M.; Kumar, A.; Singh, U.S. Abiotic stress tolerance-progress and pathways of sustainable rice production. *Sustainability* **2021**, *13*, 2078. [[CrossRef](#)]
82. Yadav, S.; Sandhu, N.; Dixit, S.; Singh, V.; Catolos, M.; Mazumder, R.R.; Rahman, M.A.; Kumar, A. Genomics-assisted breeding for successful development of multiple-stress-tolerant, climate-smart rice for southern and southeastern Asia. *Plant Genome* **2021**, *14*, e20074. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.