Originally published as:

Romanovska, P., Schauberger, B., Gornott, C. (2023): Wheat yields in Kazakhstan can successfully be forecasted using a statistical crop model. - European Journal of Agronomy, 147, 126843.

DOI: https://doi.org/10.1016/j.eja.2023.126843

Wheat yields in Kazakhstan can successfully be forecasted using a statistical crop model

Paula Romanovska*[1], Bernhard Schauberger[1,2] and Christoph Gornott[1,3]

[1]Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, P.O. Box 60 12 03, 14412 Potsdam, Germany;

[2]Department of Sustainable Agriculture and Energy Systems, University of Applied Sciences Weihenstephan-Triesdorf, Am Staudengarten 1, 85354, Freising, Germany

[3]Agroecosystem Analysis and Modelling, Faculty of Organic Agricultural Sciences, University of Kassel, 37213 Witzenhausen, Germany;

paula.romanovska@pik-potsdam.de (P.R.; 0000-0002-1454-4960); schauber@pik-potsdam.de (B.S.; 0000-0001-7917-0392); gornott@pik-potsdam.de (C.G.; 0000-0003-3933-3358)

*Correspondence: paula.romanovska@pik-potsdam.de (P.R.)

Abstract

Wheat production in Kazakhstan is fundamentally contributing to food security in Central Asia and beyond. It gained even more importance after recent spikes in global food prices in 2022. Therefore, timely and reliable estimates of Kazakh wheat production are important for food security planning and management. In this study, we developed a statistical weather-driven crop model that can successfully hindcast wheat yields at the oblast level up to two months before the harvest. The hindcast of wheat yields for 1993 to 2021 produces a median R^2 of 0.69 for the full model run and R^2 values of 0.60 and 0.37 for two levels of out-of-sample validations, respectively. Based on these yield estimates we provide a robust hindcast of the total wheat production for Kazakhstan with R^2 values between 0.86 and 0.73. We forecast total wheat production in Kazakhstan for 2022 to be 12.4 million tonnes and the average yield to be 0.96 tonnes per hectare, which is 5 % above the production and yield of 2021 (assuming equal areas). The statistical model is run with publicly available weather and yield data and requires low computational power, making it easily replicable. The forecast model can be used as a replenishment to currently applied forecasting methods supporting countries in Central Asia to meet their food demand.

Keywords: Yield forecast; Statistical crop model; Wheat; Kazakhstan; Out-of-sample validation

1. Introduction

Global wheat production is fundamental for food security providing over 20 % of global calorie intake (Lephuthing et al., 2021). At the same time prices for grains reached their record high in spring 2022 after the invasion of Russian forces in Ukraine which has aggravated existing tensions in the food market (Glauben et al., 2022) and accelerated increasing energy prices. The war comes on top of existing challenges that have already disrupted the agricultural commodities market: the COVID-19 pandemic, shipping constraints, land degradation and recent extreme events that are rising under human-made climate change (Barbier & Hochard, 2018; Behnassi & El Haiba, 2022; Nicas, 2022). Subsequent high prices and grain scarcity severely affect food security in some vulnerable food-importing countries (Behnassi & El Haiba, 2022). This makes reliable wheat production in high-producing regions even more important. One of those regions is Kazakhstan. In 2020, the country was the top 14th producer of wheat with a national production of above 14 million tonnes. This is around 70 % of the production of Ukraine and 17 % of the production of the Russian Federation (Bureau of National Statistics Kazakhstan, 2022; FAO, 2022a). Together the three countries were often referred to become the world's "bread basket" (Swinnen et al., 2017). The Kazakh wheat production is important for food security in the whole region as the main share of wheat exports goes to other Central Asian countries and the wheat quality is high (Schierhorn et al., 2020; USDA, 2010). However, Kazakhstan has shown highly variable annual wheat production ranging between 4.7 and 22.7 million tonnes since 1992 (FAO, 2022a). The export share of wheat also showed high variability with values between 20 % and 65 % of the annual production (USDA, 2010). Kazakhstan exported around 9 million tonnes of cereals in 2021 (FAO, 2022b) and was the 12th largest wheat exporter (WITS, 2022). Key importers of Kazakh wheat are Uzbekistan, Tajikistan and Afghanistan (USDA, 2022b), which underlines the high importance of Kazakh wheat production for food security in Central Asia. The wheat trading in the region was and is disrupted by the Russian grain export restrictions, political uncertainties after the Taliban takeover in Afghanistan and pandemic-related border restrictions, especially in China. This led to increased exports to several countries outside of Central Asia, namely to Italy, Azerbaijan and Turkey, while still Central Asia and Afghanistan remain the main destination for Kazakh grain exports (USDA, 2022b).

Wheat yields are generally low in Kazakhstan compared to the other high-producing countries due to adverse climate conditions (a short growing season with high temperatures in summer and harsh winter temperatures as well as limited water availability) and low input use (fertilizers and pesticides) (Schierhorn et al., 2020). Kazakhstan is estimated to have the highest wheat yield gap of the high-producing countries reaching 70 % (Senapati et al., 2022). Kazakh rainfed wheat production is largely dependent on the weather conditions. The influence of precipitation could be seen in the dry year of 2021 where the production was 21 % lower than the year before (Bureau of National Statistics Kazakhstan, 2022). Agricultural production is at risk due to a tendency towards intensive land degradation and desertification, salinisation, and decreasing soil fertility (Saparov, 2014). Furthermore, the Asian Arid Zone, within which Kazakhstan is located, is particularly sensitive to climate change and showed a past temperature increase above the global average over land (Salnikov et al., 2015; Schierhorn et al., 2020).

Given the low average yields in Kazakhstan, their high variability in percentage terms as well as further pressure on global and regional yields due to climate change, it is essential to have good estimates of production before harvest to timely plan the distribution of food and ensure food security in Central Asia and beyond. Numerous studies have developed yield forecasts based on satellite data, climatic variables and soil properties on various spatial scales (Cai et al., 2019; Luo et al., 2022). Nevertheless, according to our knowledge, little research has been done on national yield forecasting in Central Asia. This is also shown by a systematic review of yield forecasting approaches done by Schauberger et al.

(2020). They analysed 362 scientific studies conducting yield forecasts published between 2004 and 2019 and found none in a Central Asian country. To our best knowledge, national forecasts of wheat production from the government of Kazakhstan are mainly based on work-intensive field measurements. Operational yield forecasts on the national level are provided furthermore by the European Commission with the MARS Crop Yield Forecasting System, the United States Department of Agriculture (USDA) and the Food and Agricultural Organisation (FAO). The forecasting tool by the European Commission is based on satellite observations, meteorological data, meteorological forecasts, agro-meteorological and biophysical modelling, and statistical analyses (European Commission, 2023). USDA's crop production forecasts in Kazakhstan use information from agricultural experts, official governmental statistics, the analysis of economic data and satellite imagery (USDA, 2022c). The FAO bases its forecasts on ground-based information and satellite data (FAO, 2020).

Therefore, in this study, we provide a statistical wheat yield hindcast for Kazakhstan at the oblast level from 1993 to 2021 based on weather indices and use it to forecast total national wheat production in a fully blind validation for 2022. To assess the skill of the model we included two levels of out-of-sample validations. The forecast for 2022 is compared to national forecasts by FAO, USDA and the European Commission (European Commission, 2022; FAO, 2022b; USDA, 2022c).

2. Data and Methods

2.1 The study Region

Kazakhstan is located in the centre of Asia and covers 2.7 million square km making it the ninth biggest country in the world (Dana et al., 2009). The climate is dominated by its high continentality leading to hot summers and cold winters as well as low annual rainfall (Salnikov et al., 2015). Wheat production is mainly rainfed and thus sensitive to climate variability and climate change (Karatayev et al., 2022). Frequent drought events have been observed since 1770 (Zhang R. et al., 2017) disrupting stable agricultural production.

Average wheat production showed high variability in the past due to fluctuations in yield and total production area. After peaking at 19.6 million hectares in 1969 the wheat growing area steadily declined in the 1970s as low-yielding fields were taken out of grain production. The production area declined again in the 1990s after the breakup of the Soviet Union. The wheat area started to rebound only in the 2000s (USDA, 2010). Yields were and are generally low compared to other world regions and only reach more than two tonnes per hectare in a few years and for some oblasts. Higher yields can be achieved in years of high rainfall, but the drought-prone area suffers from low yields due to little rainfall in around two out of five cropping seasons (Shmelev et al., 2021; USDA, 2010).

More than 90 % of the wheat is sown as spring wheat (Abugalieva et al., 2010), as the climate in most of the country is more favourable for spring wheat than winter wheat. Winter wheat is only grown in parts of the south of Kazakhstan. This justifies a focus on the spring wheat growing season within this study. Spring wheat is sown from March to May and harvest begins in August (Alchemyka, 2022; Wang et al., 2022) with the earliest planting in the south and the latest harvesting in the north. While Kazakhstan consists of 14 administrative territories (oblasts), 75 % of the production is in the plains and lowlands of three oblasts located in the north: Aqmola, North Kazakhstan, and Qostanay (USDA, 2010; Wang et al., 2022).

2.2 Input data

The inputs of the regression model are observed yield data and agroclimatic indices based on weather data. As a source for past weather data, we chose ERA5 (fifth generation of European Centre for Medium-Range Weather Forecast (ECMWF) atmospheric reanalyses), the latest high-resolution reanalysis data set produced by ECMWF. It combines vast amounts of historical observations into global estimates using advanced modelling and data assimilation systems (Hersbach et al., 2019, 2020). ERA5 is available around two months behind real-time. Additionally, we used the initial release data ERA5T for June and July 2022 as it is available about 5 days behind real-time and thus needed for a timely forecast.

We obtained data on yield, production and production area at the oblast level from the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan - Bureau of National Statistics (Bureau of National Statistics Kazakhstan, 2022). The complete time series of yield, production and production area of wheat from 1993 to 2021 is visualized in the Supplementary Information Figure S1 and the wheat production in 2021 is shown in Figure 1. The data displayed shows the combined winter and spring wheat, as to our best knowledge separate data for the two cultivars is not available for the complete time period since 1993 from a public source. The data time series is complete for all oblasts but Atyrau, Mangghystau and South Kazakhstan, three oblasts with low or no production. For Atyrau and South Kazakhstan, the last years contain missing entries and for Mangghystau no data was reported for any year. The arid climate in the southwest of Kazakhstan hampers wheat production and could explain the missing entries in Atyrau and Mangghystau. The production is highest in the North of Kazakhstan (Aqmola, North Kazakhstan, and Qostanay) and decreases towards the South, where the reported production declines to a few thousand tonnes per oblast. Personal discussions with agricultural experts in Kazakhstan testify to the reported annual variability of wheat production. Namely, the high-yielding years of 2009 and 2011 as well as the low-yielding years of 2010 and 2012 were frequently mentioned. As an exception, it was pointed out that the total wheat production in 2008 was likely up to 3 million tonnes below the reported one.

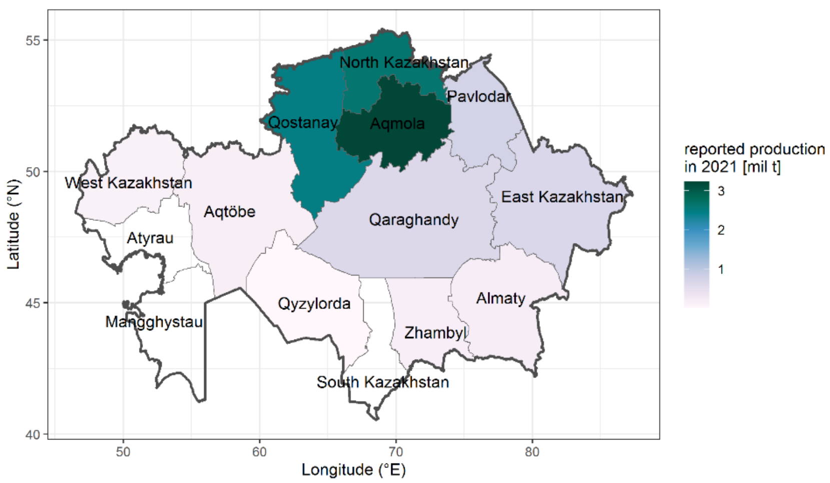

Figure 1: Reported wheat production in 2021 [mil t] in each oblast. The data source is the Bureau of National statistics Kazakhstan (2022).

Additionally, crop planting and harvesting dates of spring wheat were used in this study to define the growing season. These were taken from conversations with agricultural experts from the Analytical Centre of the Economic Policy in Agricultural Sector (ACEPAS) - a local private crop forecast agency - and verified with crop calendars of the Center for Sustainability and the Global Environment (Sacks et al., 2010) and the United States Department of Agriculture - Foreign Agricultural Service (USDA, 2022a). The earliest planting date was thereby set to the 15^{th} of April and the latest harvesting date to the 30^{th} of September. This covers the planting period of spring wheat in the whole of Kazakhstan with the earliest planting in the south and the latest harvesting in the north. Due to a lack of region-specific data, we used the same crop calendar for the whole of Kazakhstan.

2.3 Model setup

We developed a statistical regression model following the approach of Gornott & Wechsung (2016), Laudien et al. (2022) and Schauberger et al. (2017) to hindcast wheat yields for 1993-2021 and forecast wheat yields for 2022. We used the R (R Core Team, 2022) packages *caret, glmnet, tidyr, dplyr* and *ggplot.* To account for the diverse climate conditions in Kazakhstan, we set up a different model for each oblast following the flow chart in Figure 2 and steps A to E described in detail below. A more fine-grained analysis at a smaller administrative level was not possible due to the unavailability of reliable yield data.

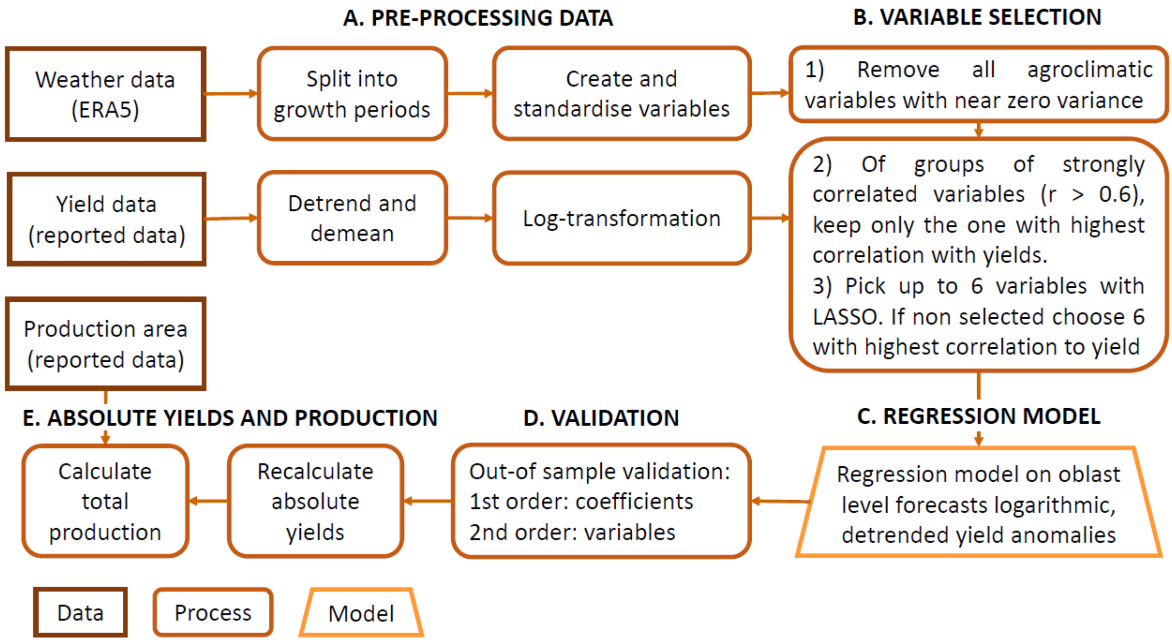

Figure 2: Flow chart of yield model.

A. Pre-processing of agroclimatic variables and yield data

Based on daily maximum and minimum temperature, precipitation and near-surface relative humidity, we created 20 agroclimatic variables that describe either potential stress or growth factors (listed in Table 1). In the literature, these variables were found to be relevant for wheat production or were successfully used in statistical crop models.

Temperature and water availability have a main influence on potential crop growth. We aim at covering these aspects in the model by including the agroclimatic input variables listed in Table 1 and described in the following.

Because of the connection between temperature and plant development, growing degree days (GDD) are often used as a predictive indicator for crop development (Dar et al., 2018). Next to the median minimum and maximum temperature in the growing season, we considered variables related to low and high-temperature extremes as both can decrease the rate of dry matter production or even lead to crop failure (Grace, 1988; Porter & Gawith, 1999). Plant development of wheat is slower below an optimum temperature range, considered to be inactive below 0°C and plants might be irreversibly lost beyond a lethal limit that depends on the growth stage and is well below 0°C (Nagai & Makino, 2009; Porter & Gawith, 1999). To cover these aspects we included the region-specific long-term 5th and 1st percentile of minimum daily temperature, the sum of frost degree days (below 0°C) and the sum of chilling degree days (below 10°C). High temperatures can adversely affect the flowering and reproductive growth of wheat and are associated with decreasing relative humidity affecting yield potential (Meng et al., 2017). To represent different levels of high-temperature extremes, we included heat degree days above 35°C as well as the 95^{th} and 99^{th} percentile of maximum daily temperature as variables. Additionally, crop yields respond to the diurnal temperature range (DTR, difference between maximum and minimum daily temperature) since some plant processes (e.g. photosynthesis) happen only during the day, while others (e.g. crop development) are nonlinearly related to temperature during the day and night. A higher DTR is furthermore related to higher solar radiation, which can benefit crop growth, especially in case of sufficient water availability. A higher DTR can also lead to decreases in yields in case of higher frost occurrence (Lobell, 2007). To cover the named processes related to the DTR we included the mean DTR as well as the standard deviation of the DTR as variables.

As wheat production in Kazakhstan is mainly rainfed, it highly depends on water provided by rainfall during the growing season. Therefore, we included the precipitation sum as an agroclimatic variable. On the one hand, heavy rainfall events can adversely affect crop growth by damaging the crop canopies or limiting root and plant functions through anoxic soil conditions (van der Velde et al., 2012) and on the other hand, it can positively impact crop production in semi-arid areas by wetting deeper soil layers. Taking those effects into consideration, we included the number of high rainfall events above 15 mm and above 30 mm. Long dry spells during the sowing season lead to a high risk of crop failure (Liakatas, 1997). To account for the damaging impacts of dry spells, we included consecutive dry spells of at least 7 and 20 days. Drought conditions were additionally covered by the standardized precipitation–evapotranspiration index (SPEI) (Vicente-Serrano et al., 2010). The SPEI is based on the calculation of a water balance and incorporates the effects of temperature variability on evapotranspiration. Precipitation before the growing season was shown to also have a positive effect on yields through conserved soil moisture (Meng et al., 2017). Thus, we considered the 12-month SPEI taken on the 31^{st} of July which covers the average drought conditions of the whole year next to the 4-month SPEI which covers the drought conditions of April-July.

Humidity can affect crop development in two ways. First, it can directly affect the plant by altering the water content of the plant. Second, humidity can affect growth indirectly by influencing photosynthesis, pollination, leaf growth and pests and diseases (Zhang P. et al., 2015). Accounting for possible negative and positive effects of humidity, we included mean near-surface relative humidity as a variable.

Variable	Definition
gdd	growing degree days with tbase = 0°C and values above 30°C are counted as 30°C
tasmin	median of daily minimum temperature
tasmax	median of daily maximum temperature
ahdd	accumulated heat degree days (summing degrees over 35°C up)
afdd	accumulated frost degree days (summing degrees below 0°C up)
acdd	accumulated chilling degree days (summing degrees below 10°C up)
tmax95	95^{th} percentile of maximum daily temperature
tmax99	99^{th} percentile of maximum daily temperature
tmin05	5^{th} percentile of minimum daily temperature
tmin01	1^{st} percentile of minimum daily temperature
DTR	mean diurnal temperature range (difference between minimum and maximum daily temperature)
DTR_sd	standard deviation of diurnal temperature range
pr	mean precipitation
15mm	number of days with rainfall above 15mm/day

30mm	number of days with rainfall above 30mm/day
dry7	number of dry spells with a length of at least 7 days (day with < 0.5 mm is defined as dry)
dry20	number of dry spells with a length of at least 20 days (day with < 0.5 mm is defined as dry)
spei4	Standardised Precipitation-Evapotranspiration Index (SPEI) at a 4-month time scale taken at 31st of July
spei12	SPEI at a 12-month time scale taken on the 31st of July
hurs	mean near-surface relative humidity

Table 1: Definition of agroclimatic variables used as input for the regression model.

The variables were separately calculated for three phases of the growing season, the first one from planting until the beginning of stem elongation, the second one until the end of flowering and the third stage until the forecast date 31st of July. The phases were defined for each year separately according to GDD thresholds as defined by Miller et al. (2001) (Table 2). These thresholds of plant growth stages according to GDDs were taken from experiments with hard red wheat in semiarid Saskatchewan (southern Canada) owing to the scarcity of equivalent data in Kazakhstan. The values are meaningful approximations for wheat in northern Kazakhstan due to the similar climate and the classification of most wheat in northern Kazakhstan as a hard red spring type (Morgounov et al., 2007). Still, the GDD thresholds might not match all cultivars, especially those in the south of Kazkahstan. To hindcast and forecast yields before harvest, we included only the weather data up to two months before the last day of reported harvest (30th of September), thus in this case until 31st of July.

Phase of growing season	Growth stage	Number of GDDs
Phase 1	until the beginning of stem elongation	0 – 592 GDD
Phase 2	until the end of flowering	593 – 901 GDD
Phase 3	until forecast date	902 GDD – end

Table 2: Definition of growing season phases.

As the last step, we standardized the weather variables by subtracting the mean and dividing this by the standard deviation of each variable. This allows for a consistent interpretation of the variable coefficients.

To eliminate long-term trends due to management changes, we detrended the yield time series with the method (mean, linear, quadratic, and cubic) that resulted in the lowest Akaike Information Criterion (AIC). Then we took the logarithmic values of yield to adjust the yield input data to a Gaussian distribution.

B. Variable selection

At first, we removed all variables that showed near zero variance. Specifically, we excluded predictors that are subject to both of the following two characteristics: (1) the ratio of the frequency of the most common value to the second most common value is more than 10 and (2) the most common value is present for more than 50 % of the total sample. The thresholds were defined visually and empirically. The chosen values have been proven to be useful here to exclude not only predictors that have no variance but also predictors which have only differing values for one or two years, as was the case e.g. for heavy rainfall events in some oblasts. Next, we checked which variables are collinear (Pearson correlation coefficient > 0.6). From each group of collinear variables, we only kept the one variable with the highest correlation to yield anomalies. The final variable selection was done with LASSO and was set to allow for a maximum of 6 variables. In case LASSO selected more than 6 variables the ones with the lowest correlation to yield were removed. We set this maximum to avoid overfitting and tested that allowing for more variables does not improve the out-of-sample validations further.

C. Regression model

Finally, we applied an individual linear regression model for each oblast following Eq 1.

$$\log(y_{ti}) = \beta_0 + \sum_{j=1}^{J} \beta_{ji} x_{jti} + \varepsilon_{ti} \quad (1)$$

y – detrended and demeaned response variable (i.e. yield)
β – beta coefficient
β_0 – intercept
x – standardized input variables (i.e. agroclimatic indices)
ε – error term
t – T years (t = 1,…T)
i – N spatial units (i.e. oblast) (i = 1,…N)
j – J variables (j = 1,…J); $J \leq 6$

D. Model testing and validation

We performed the Breusch-Pagan test against the heteroscedasticity of model residuals and the Shapiro-Wilk test for a normal distribution of residuals.

The robustness of the yield model is assessed by performing two out-of-sample validations based on Laudien et al. (2020).

First level out-of-sample: we selected the agroclimatic variables with LASSO based on data from all years. Then we subsequently removed one year from the data set and fitted the coefficients to simulate yields for the removed year.

Second level out-of-sample: we subsequently removed one year from the data set and selected the variables with LASSO based on the reduced data set. We then fitted the coefficients and simulated yields for the removed year. The second level out-of-sample validation represents the operational case as it does not use any information from the year to be hindcasted.

We used the standard model evaluation index R^2 (coefficient of determination) to explain the degree that our input variables (agroclimatic indices) can explain our output variable (yield). R^2 ranges from 0 to 1, whereby 1 indicates a perfect model fit.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (2)$$

SS_{res} – sum of squares of residuals (modelled minus observed value)

SS_{tot} – total sum of squares (sum over all squared differences between the observations and their overall mean)

E. Recalculating yields and total production

The regression model results in estimates of detrended, logarithmic yield anomalies. To obtain the predicted actual yields, we transformed the detrended, logarithmic yield anomalies back to their actual values.

Finally, we calculated the predicted total production in Kazakhstan by multiplying the predicted yield with the production area for each oblast and summing over all oblasts. To forecast the production for 2022, we assumed the same production area for wheat in 2022 as in 2021. This approximation is plausible according to data by USDA (USDA, 2022b) estimating a similar total production area and the JRC MARS Bulletin (European Commission, 2022) estimating a 4 % higher area for 2022 as was reported for the previous year. We included only oblasts for which an official yield time series since 1993 is available without missing years to ensure the comparability between all years. This excludes Atyrau and South Kazakhstan. Due to their low wheat production in the reported years (around 2 % of the national production) and the unavailability of data in recent years (Figure S1c), we assume that the model covers approximately 98 % of total production.

3. Results

3.1 Validating the hindcast model

The model can hindcast actual yields reasonably well in all oblasts with R^2 values ranging between 0.45 and 0.93 and a median R^2 of 0.69 (R^2 values for each oblast are shown in Figure 3). The median R^2 decreases to 0.60 and 0.37 for the two out-of-sample validations, respectively. The R^2 values are slightly lower for the logarithmic, detrended yield anomalies (shown in Figure S2). LASSO selected between 1 and 13 variables for the different oblasts (on average 4.3 variables) whereby only in two oblasts more than the 6 variables, that we restricted the selection to later manually, were selected. Residuals were normally distributed for ten out of thirteen oblasts (Shapiro-Wilk test). Heteroscedasticity did not appear in any of the models (Breusch-Pagan test).

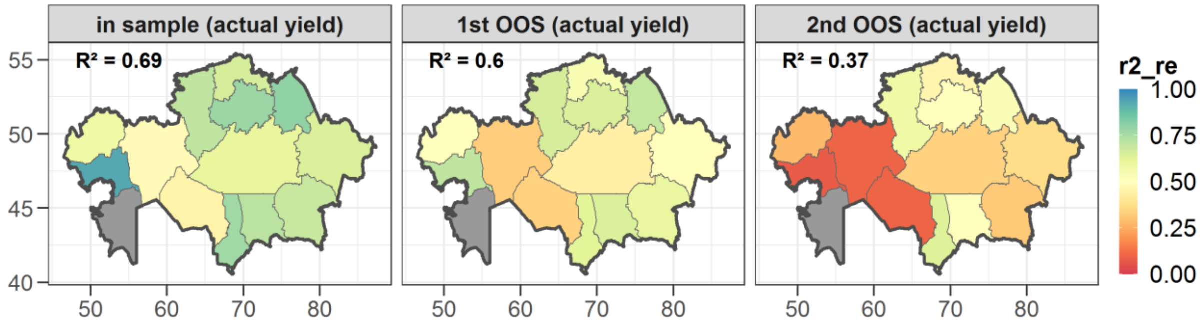

Figure 3: Model performance measured as R^2 for the actual (recalculated) yields from 1993 to 2021 at the oblast level. The left map (in sample) shows the model performance for the complete time series, the map in the middle (1st OOS) shows the 1st out-of-sample validation and the right map (2nd OOS) the 2nd out-of-sample validation. The median R^2 over all oblasts is shown in the upper left corner of each map.

Nevertheless, the model cannot capture all extreme years and shows some outliers for individual years and oblasts (Figure 4). The 2nd order out-of-sample simulations show unreasonable high yields in single years in Atyrau and Aqtöbe. The out-of-sample model performance is lower for the low-producing and arid oblasts in the west of the country. For the high-producing oblasts, Aqmola, North Kazakhstan, and Qostanay, the out-of-sample validations show more robust results with no outliers and R^2 values for the 2nd order out-of-sample of 0.51, 0.46 and 0.58 respectively (Figure 4 & Table S1). These values are well above the country median of 0.37.

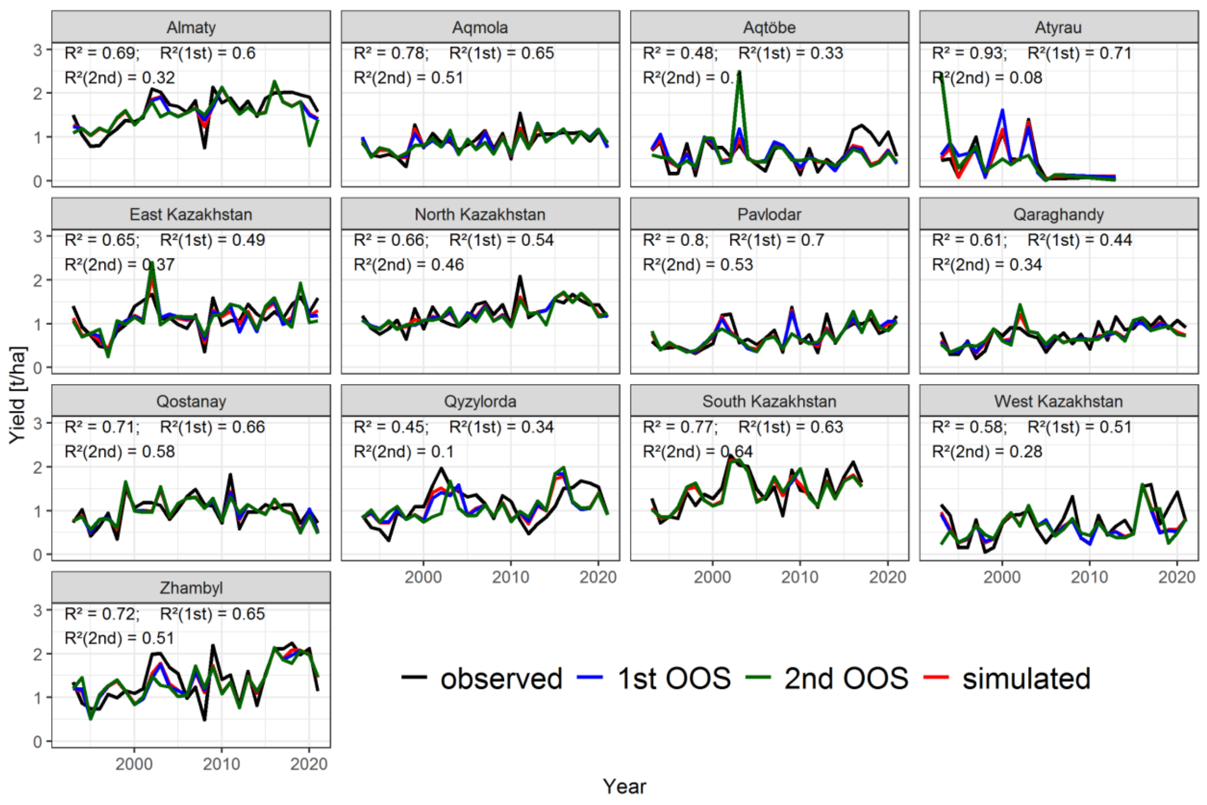

Figure 4: Time series of observed and simulated yield (in t per hectare) for each oblast including the 1st and 2nd order out-of-sample. The R^2 values for the full model and the 1st and 2nd out-of-sample are written in the upper part of each graph.

The variables that were selected by Lasso in the model are related to precipitation sums in the whole growing season (pr, sprei4, spei12) and in the second growth period (pr_2), single high rainfall events in all three growth periods (mm15_1, mm30_1, mm15_2, mm30_2, mm15_3), dry spells in all three growth periods (dry20_1, dry7_1, dry20_2, dry7_2, dry20_3), cold temperatures (acdd, acdd_2, acdd_3, tasmin), cold temperature extremes (tmin05_3), high temperature extremes (ahdd_1, ahdd_3, tmax99_2), the temperature range (DTR_sd, DTR_sd_2, DTR_sd_3, DTR_1) and mean near-surface relative humidity (hurs, hurs_1, hurs_3). Precipitation sums, a positive SPEI value, single precipitation events and near-surface relative humidity show consistently positive coefficients while dry spells show negative ones. Higher mean minimum temperatures and less chilling degree days show, with one exception, a negative influence on yields. Coefficients of high and low-temperature extremes are negative in the second and third growth stages (ahdd_3, tmax99_2, tmin05_3) and positive in the first growth stage (ahdd_1). The DTR (DTR_1) was once selected by Lasso with a negative coefficient. The standard deviation of the DTR (DTR_sd, DTR_sd_2, DTR_sd_3) shows predominantly positive influences on wheat production.

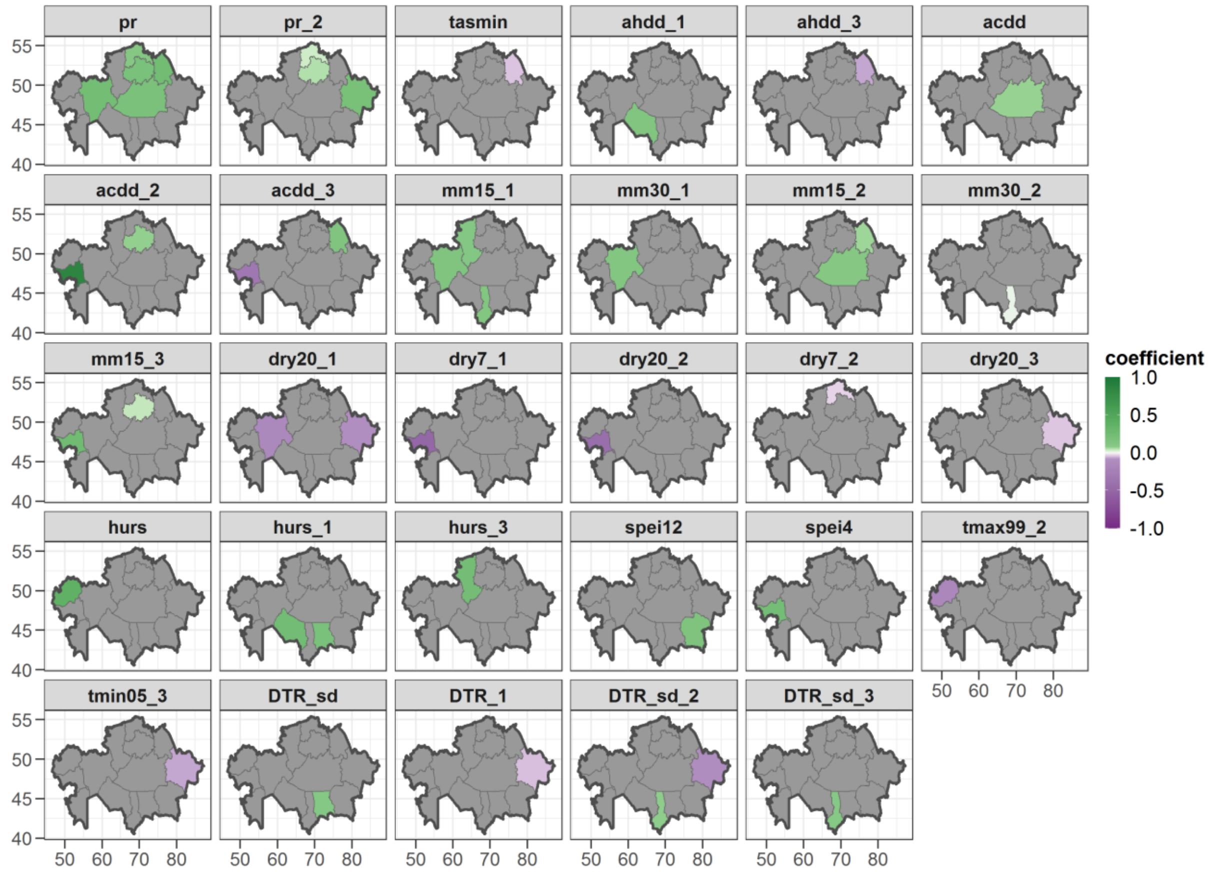

Figure 5: Regression coefficients for all selected variables. The number (_1, _2 and _3) indicate the growth stage. E.g. "pr_2" thus stands for precipitation sum in the second growth stage. No ending indicates the average over the complete growing period up to the forecast date. Oblasts in which the variable was not selected, are shown in grey.

3.2 Hindcast of total Kazakh wheat production

We further used the yield hindcast model at the oblast level to estimate total wheat production in Kazakhstan by multiplying the predicted yield by the wheat production area of each oblast and summing the production over all oblast (Figure 6a). The time series of observed and simulated total production are shown in Figure 6b. The R^2 values for the total production hindcast range between 0.86 and 0.73.

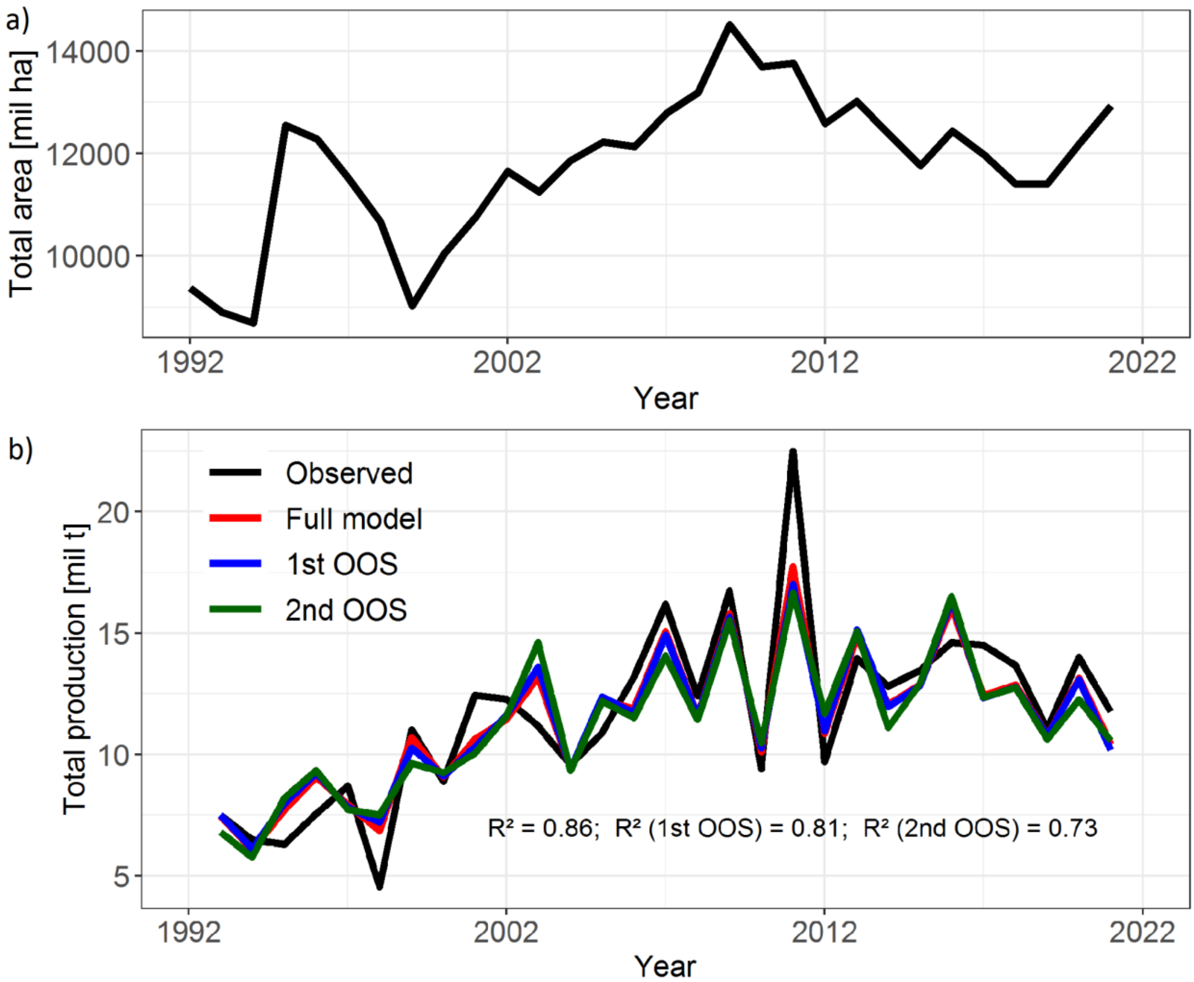

Figure 6: Time series of a) total observed wheat production area and b) observed and simulated total wheat production in Kazakhstan including the 1st and 2nd order out-of-sample. The R^2 values for the full model and the 1^{st} and 2^{nd} out-of-sample are written within the graph.

We additionally run the total production model considering only the three oblasts with the highest production (Aqmola, North Kazakhstan and Qostanay). The graphs for the total production are shown in the Supplementary Information (Figure S3) and largely agree with Figure 6b) since these three oblasts are dominating the total Kazakh wheat production. Still, the R^2 values are slightly higher when considering only the three oblasts with 0.88, 0.83 and 0.78 compared to 0.86, 0.81 and 0.73.

3.3 Forecasting wheat yield and production for 2022

We applied the model to the year 2022 to get a completely independent forecast of yields for each oblast and the total Kazakh production. This is tantamount to an operational forecasting mode, where no information on yields from the target year is available at the time of forecasting. The forecasted yields at the oblast level are displayed in Figure 7 and listed in the Supplementary Information, Table S1. The forecasted total production in the 11 oblasts, excluding Atyrau, Mangghystau and South Kazakhstan, is around 12.4 million tonnes for 2022. The forecasted wheat production in the three oblasts with the highest production (Aqmola, North Kazakhstan and Qostanay) sums to 9.85 million tonnes.

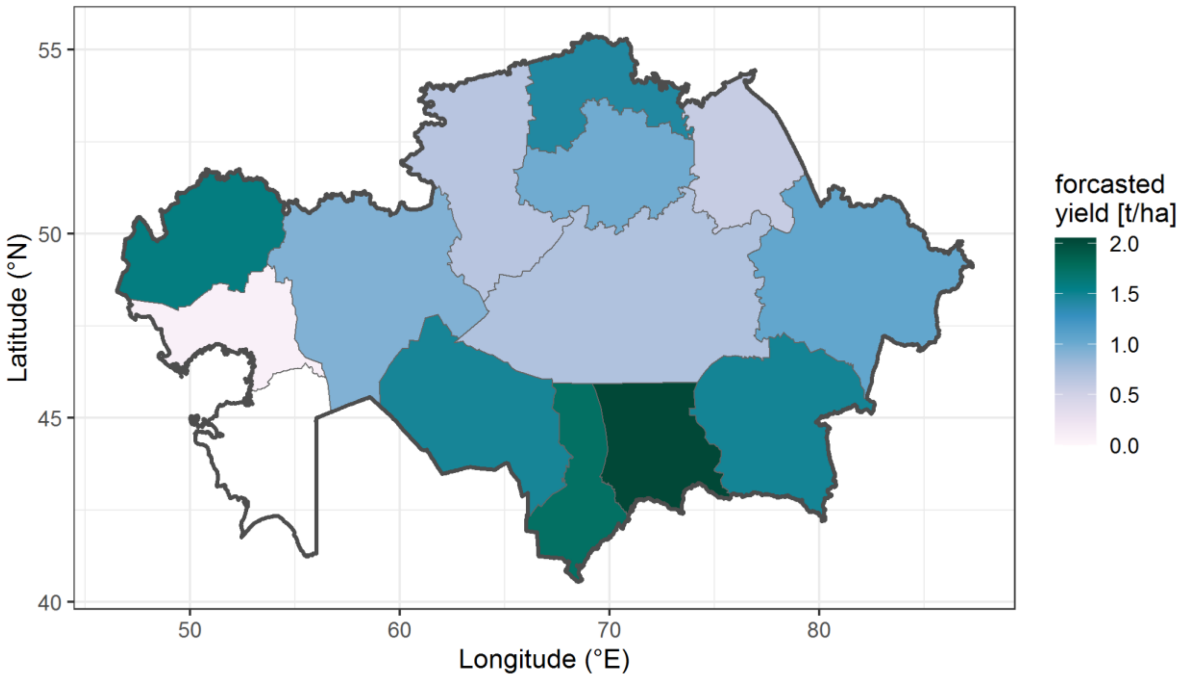

Figure 7: Yield forecast for 2022 in tonnes per hectare at the oblast level.

4. Discussion

We provide and validate a seasonal hindcast of wheat yields at the oblast level and total wheat production at the national level in Kazakhstan from 1993 to 2021. Furthermore, we run a completely independent national production forecast for 2022. The model has the advantage of requiring low computational power and little input as it is solely based on yield and weather data.

The model shows high accuracy in hindcasting wheat yields during the growing season up to two months before harvest (R^2 of 0.69) and attains satisfying out-of-sample validations (median R^2 of 0.60 and 0.37). Thus, the model can explain at least 37 % of the yield variability (when applying the rigorous second-level validation) only by accounting for weather anomalies up to the 31st of July. This underlines the high influence of weather on wheat production in Kazakhstan.

The model clearly shows the importance of precipitation sums, high precipitation events, high humidity, and the absence of longer dry spells and droughts for wheat production in the whole of Kazakhstan and throughout all growth stages. Sufficient precipitation was found to be most relevant for yield in the second growth stage between stem elongation and the end of flowering. Water scarcity during the growing season is also stated and analysed as the major stress factor for wheat production in Kazakhstan by various other studies (Babkenov et al., 2020; Fehér et al., 2016; Pavlova et al., 2014; Shmelev et al., 2021). The variable selection shows that minimum temperatures below 10°C have a predominantly positive effect on wheat yields and high minimum temperatures were negatively associated with yields in all growth stages. Night temperatures might influence crop production through the following two processes. First, high night temperatures (20°C and above) over a longer period were shown to lead to decreased photosynthesis and lower grain yields (Prasad et al., 2008). Second, lower night temperatures are associated with higher solar radiation as cloud-free nights are cooler. The DTR and its standard deviation can also give some information on the variability of cloud cover within a growth stage. This, in return, influences solar radiation as well as precipitation. Our results regarding the DTR and its standard deviation point to the correlation that a higher mean DTR in the first growth stage can lead to lower yields while a higher variability of DTR in any growth stage can lead to higher yields. The DTR in the second growth stage in East Kazakhstan does not fit this as the model coefficient for this variable points in the opposite direction. Variables related to mean maximum temperature and growing degree days were not selected pointing to the already sufficient heat units available during the growing season in all years. According to the variable selection, the correlation between high-temperature extremes and yields is negative except for the first growth stage, where high temperatures show a positive effect on crop growth.

Most of the variables were selected throughout different growth stages and the coefficients show the same sign. Only for heat degree days, the sign of coefficients differs between growth stages, pointing to a positive influence of hot temperatures on yield in the first growth stage and a negative one in the third growth stage. Thus, the different influence of weather on wheat depending on the growth stage is not clearly shown by the model results. This might also be due to the uniform definition of the growth stages according to GDDs over the whole of Kazakhstan despite the use of different cultivars.

Compared to indices related to precipitation, variables related to high-temperature extremes were not as often selected by the model. Wang et al. (2022) and Schierhorn et al. (2020) analysed the influence of climate on crop production and also found that extremely high temperatures had little impact on wheat production in Kazakhstan. We can conclude that under current climate conditions drought stress has a higher negative impact on wheat production than high-temperature stress. As the effect of temperature extremes on crop growth is discontinuous, the negative impact of high-temperature extremes might still come into play in the future after a specific warming level has reached under climate change. According to our modelling results, high minimum temperatures and the absence of enough chilling degree days are already now negatively correlated to yields. This will become more relevant under further increasing temperatures. Precipitation amount and distribution showed a high influence on crop yields. Precipitation in the future would need to rise at least by the amount of additional evaporation due to global warming to achieve similar yields in the future. While the latest climate models project a robust increase in mean annual precipitation in Central Asia, the changes also include a seasonal shift with a wetter spring and a dryer summer season (Jiang et al., 2020). Only a few studies exist that analyse subsequent influences on future wheat yields in Kazakhstan and those do not agree on a future trend (compare IPCC WG2 (2022)). Furthermore, the influences of climate change on wheat production are spatially diverging over Kazakhstan according to Schierhorn et al. (2020) and Wang et al. (2022).

The R^2 of the total production hindcast in Kazakhstan ranges between 0.86 (fitted) and 0.73 (level 2 out-of-sample) underlining the operability of the production forecast. The hindcast for total production also covers extreme years. The high-producing years 2009 and 2011 as well as the low-producing years 2010 and 2012 are to a satisfying degree represented. When running the model only with the high-producing oblasts, the R^2 values for the total production hindcast can even further be increased with values between 0.88 (fitted) and 0.78 (level 2 out-of-sample). This might be due to first, a higher data uncertainty in areas with low production, second, a high proportion of crop failure which is only partially recognized by the statistical model, third, the uniformly set GDD thresholds that fit better to the wheat production in the north, and, forth, the uniformly chosen crop calendar for spring wheat that is likely apt for the highly productive areas in the north of Kazakhstan, but not for the west or to the winter wheat in the south. Differences between modelled and reported data might not only come from model errors but also to some extent from errors in the reported data.

The high R^2 values for the total production hindcast are only to some extent due to the high yield model performance. We see two additional determents for the goodness of fit: First, errors with opposite signs can cancel each other out by summing the production over all oblasts and, second, including the known production area increases the goodness of the model. The second point underlines the importance of having good estimates of the production area to ensure that the model can perform in a real case equally well.

With this model, we provided a completely independent yield and production forecast up to two months before harvest at the oblast and national level for 2022. Total wheat production in 2022 is forecasted to be 12.4 million tonnes for the whole country but Atyrau, Mangghystau and South Kazakhstan. The actual numbers at the national level might be slightly higher due to the exclusion of oblasts and the possibly higher production area in 2022 (European Commission, 2022). The average yield is forecasted to be 0.96 tonnes per hectare and thus 5 % above the low-yielding last year but still well below the high-yielding year 2020. Other wheat yield forecasts that compile and harmonize data from different sources project a higher increase of 11.7 % (FAO), 19.4 % (MARS Crop Yield Forecasting System) and 11.8 % (USDA) for 2022 compared to 2021 (European Commission, 2022; FAO, 2022b; USDA, 2022c). Forecasts for winter wheat in the south of Kazakhstan are especially high with a 57 % increase compared to 2021 (European Commission, 2022) and partially drive the high overall increase in wheat. As our statistical model is not fit to winter wheat, this potentially high increase in production in the south is likely not covered and underestimated. Official observational data from the Bureau of National Statistics have not yet been released.

Processes that cannot or only to a limited extent be covered by the model are pests and diseases, increasing soil degradation and varying agricultural management practices like fertilizer input. High costs of fertilizer in 2022 are for example not accounted for in the model and might have led to less fertilizer use and thus lower yields. Errors in observations (weather and yield) as well as the necessary aggregation of agroclimatic variables over a large area add further uncertainties to the model. Separate data on spring wheat yield instead of the combined winter and spring wheat data could lead to more accurate results in the south of Kazakhstan. Additionally, not all influences of weather on yield might be covered by the chosen agroclimatic variables. Namely, non-linear processes and interactive terms were partially neglected. Existing discontinuous threshold responses to temperature and precipitation, might not be caught by the model. The model could further be improved by introducing differing growing seasons and definitions of growth stages instead of uniform ones over the whole of Kazakhstan.

A large part of Kazakh wheat was exported in the last years. After a disrupted global agricultural commodity market following the invasion of Russian forces in Ukraine, the Kazakh government introduced grain export restrictions in April 2022 leading to a limit on wheat export to guarantee domestic grain stocks and control national wheat prices (USDA, 2022b). The restrictions were suspended in September 2022 (FAO, 2022b) contributing to the stabilization of the global wheat trading market. In line with this, FAO forecasts wheat exports from Kazakhstan to be almost 9 million tonnes in the season 2022/23 and thus only slightly below the average of the recent years (FAO, 2022b). Nevertheless, even constant total export rates can lead to considerably lower wheat availability in single countries as export flows from Kazakhstan have substantially changed their distribution between buyer countries in recent years. Decreases in wheat exports were already observed in 2021 especially to geographically close countries, namely Afghanistan (-11 %), Tajikistan (-17 %), Kyrgyzstan (-79 %), China (-92 %) and Russia (-46 %) (USDA, 2022b). Rising wheat prices might aggravate this tendency this year constraining food security in Central Asia despite stable wheat production in the main producing country Kazakhstan.

In light of multiple conflicts and economic crises, increasing climate change and degrading soil conditions in Kazakhstan, food security in the region is at risk. Efforts in improved agronomy and developing and applying drought-resistant varieties can contribute to closing the existing high yield gap (Morgounov et al., 2013; Senapati et al., 2022) and thus can help to minimize the risk of food insecurity in Kazakhstan and the whole of Central Asia.

5. Conclusions

In this paper, we developed a statistical wheat yield and production hindcast model at the oblast level in Kazakhstan based on yield and weather data from 1993 to 2021. The hindcast of yield at the oblast level produces a median R^2 of 0.69. We stringently evaluate our model in a double out-of-sample validation that produces a median R^2 of 0.60 and 0.37. The selection of climate variables in the model allows us to conclude that mean precipitation and precipitation distribution are most crucial for wheat production and that high minimum temperatures hamper wheat growth in Kazakhstan. This highlights the risk for yield declines due to increasing climate change, whereby spatial differences were observed. The hindcast of total production at the national level shows high performance with R^2 values between 0.86 and 0.73. Higher R^2 values are reached for the forecast in the three high producing oblasts in the north. This underpins the operability of this forecast, especially in northern Kazakhstan, which we advise using as a replenishment to field measurements that are currently the main source for national forecasts. The total wheat production in Kazakhstan for 2022 is forecasted to be 12.4 million tonnes and yields are forecasted to be 5 % above the year 2021. The forecasted increase compared to the previous year is in line with other forecasts, namely by FAO, USDA and the European Commission. Still, all three other sources forecast a higher increase between 11.7 and 19.4 %. Due to a disrupted trading market, the relatively high production might still lead to less wheat availability in neighbouring countries. The high performance and the interpretability of the model show that wheat production can be reasonably well estimated two months before the end of harvest using a simple statistical model based on publicly available weather and yield data. Such models can be used for operational forecasts to inform about looming food shortages and needed agronomic inventions to avert food shortages.

Acknowledgements

We highly appreciate the creation and provision of datasets by all institutions and spatial product creators. Namely, we would like to thank the Analytical Centre of the Economic Policy in Agricultural Sector (ACEPAS), the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan - Bureau of National Statistics and the Copernicus Climate Change Service (C3S) and the European Centre for Medium-Range Weather Forecast (ECMWF).

This work has been funded by the Green Central Asia project financed by the German Federal Foreign Office (Auswärtiges Amt).

For their valuable feedback and input we would like to thank Rahel Laudien, Abror Gafurov, Gulnaz Iskakova, Dauren Oshakbayev, Evgeniy Kim, Viktoriya Krylova and Iulii Didovets.

Statements & Declarations

Funding

This work has been funded by the Green Central Asia project financed by the German Federal Foreign Office (Auswärtiges Amt).

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Data Availability

The datasets analysed during the current study are available at the Copernicus repository (https://cds.climate.copernicus.eu/#!/home), the Center for Sustainability and the Global Environment repository (https://sage.nelson.wisc.edu/data-and-models/datasets/crop-calendar-dataset/) and at the Bureau of National Statistics of Kazakhstan (https://stat.gov.kz/official/industry/14/statistic/7).

Code availability

Code can be made available upon request.

Authors Contributions

***Paula Romanovska**: Conceptualisation, Methodology, Formal analysis, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualisation; **Bernhard Schauberger**: Conceptualisation, Methodology, Resources, Writing – Review & Editing, Supervision; **Christoph Gornott**: Conceptualisation, Methodology, Writing – Review & Editing, Supervision, Funding acquisition;*

References

Abugalieva, A., Roberto, •, & Peña, J. (2010). Grain Quality of Spring and Winter Wheat of Kazakhstan. *The Asian and Australasian Journal of Plant Science and Biotechnology*.

Alchemyka. (2022). *Озимая пшеница*. https://alchemyka.kz/kulturyi/ozimaya-pshenicza.html

Babkenov, A. T., Babkenova, S. A., Abdullayev, K. K., & Kairzhanov, Y. K. (2020). Breeding Spring Soft Wheat for Productivity, Grain Quality, and Resistance to Adverse External Factors in Nothern Kazakhstan. *Journal of Ecological Engineering*, *Vol. 21*(nr 6), 8–12. https://doi.org/10.12911/22998993/123160

Barbier, E. B., & Hochard, J. P. (2018). Land degradation and poverty. *Nature Sustainability 2018 1:11*, *1*(11), 623–631. https://doi.org/10.1038/s41893-018-0155-4

Behnassi, M., & El Haiba, M. (2022). Implications of the Russia–Ukraine war for global food security. *Nature Human Behaviour 2022 6:6*, *6*(6), 754–755. https://doi.org/10.1038/s41562-022-01391-x

Bureau of National Statistics Kazakhstan. (2022). *Statistics of agriculture, forestry, hunting and fisheries*. https://stat.gov.kz/official/industry/14/statistic/7

Cai, Y., Guan, K., Lobell, D., Potgieter, A. B., Wang, S., Peng, J., Xu, T., Asseng, S., Zhang, Y., You, L., & Peng, B. (2019). Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agricultural and Forest Meteorology*, *274*, 144–159. https://doi.org/10.1016/J.AGRFORMET.2019.03.010

Dana, L. P., Han, M., Ratten, V., & Welpe, I. M. (2009). Handbook of research on Asian entrepreneurship. *Handbook of Research on Asian Entrepreneurship*. https://doi.org/10.4337/9781781952658

Dar, E. A., Brar, A. S., & Yousuf, A. (2018). Growing degree days and heat use efficiency of wheat as influenced by thermal and moisture regimes. *Journal of Agrometeorology*.

European Commission. (2022). *JRC MARS Bulletin - Global outlook Crop monitoring European neighbourhood Kazakhstan - September 2022*. https://doi.org/10.2760/933261

European Commission. (2023). *Crop yield forecasting*. EU Science Hub. https://joint-research-centre.ec.europa.eu/scientific-activities-z/crop-yield-forecasting_en

FAO. (2020). *Global Information and Early Warning System*. Markets and Trade. https://www.fao.org/3/ca7518en/ca7518en.pdf

FAO. (2022a). *FAOSTAT Statistical Database*. https://www.fao.org/faostat

FAO. (2022b). *GIEWS Country Brief on Kazakhstan*. GIEWS - Global Information and Early Warning System on Food and Agriculture. https://www.fao.org/giews/countrybrief/country.jsp?code=KAZ

Fehér, I., Lehota, J., Lakner, Z., Kende, Z., Bálint, C., Vinogradov, S., & Fieldsend, A. (2016). Kazakhstan's wheat production potential. *The Eurasian Wheat Belt and Food Security: Global and Regional Aspects*, 177–194. https://doi.org/10.1007/978-3-319-33239-0_11/COVER

Glauben, T., Svanidze, M., Götz, L., Prehn, S., Jamali Jaghdani, T., Đurić, I., & Kuhn, L. (2022).

The War in Ukraine, Agricultural Trade and Risks to Global Food Security. *Intereconomics 2022 57:3*, *57*(3), 157–163. https://doi.org/10.1007/S10272-022-1052-7

Gornott, C., & Wechsung, F. (2016). Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany. *Agricultural and Forest Meteorology*, *217*, 89–100. https://doi.org/10.1016/J.AGRFORMET.2015.10.005

Grace, J. (1988). Temperature as a determinant of plant productivity. *Symposia of the Society for Experimental Biology*, *42*, 91–107. https://europepmc.org/article/med/3270210

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., … Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/https://doi.org/10.1002/qj.3803

Hersbach, H., Bell, W., Berrisford, P., Horányi, A., J., M.-S., Nicolas, J., Radu, R., Schepers, D., Simmons, A., Soci, C., & Dee, D. (2019). *Global reanalysis: goodbye ERA-Interim, hello ERA5*. *159*, 17–24. https://doi.org/10.21957/vf291hehd7

IPCC. (2022). Climate Change 2022: Impacts, Adaptation and Vulnerability Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. In *[P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, J. Malley, (eds.)] Cambridge University Press. Cambridge University Press, Cambridge, UK an*. https://doi.org/10.1017/9781009325844.

Jiang, J., Zhou, T., Chen, X., & Zhang, L. (2020). Future changes in precipitation over Central Asia based on CMIP6 projections. *Environmental Research Letters*, *15*(5), 054009. https://doi.org/10.1088/1748-9326/AB7D03

Karatayev, M., Clarke, M., Salnikov, V., Bekseitova, R., & Nizamova, M. (2022). Monitoring climate change, drought conditions and wheat production in Eurasia: the case study of Kazakhstan. *Heliyon*, *8*(1), e08660. https://doi.org/10.1016/J.HELIYON.2021.E08660

Laudien, R., Schauberger, B., Makowski, D., & Gornott, C. (2020). Robustly forecasting maize yields in Tanzania based on climatic predictors. *Scientific Reports*. https://doi.org/10.1038/s41598-020-76315-8

Laudien, R., Schauberger, B., Waid, J., & Gornott, C. (2022). A forecast of staple crop production in Burkina Faso to enable early warnings of shortages in domestic food availability. *Scientific Reports 2022 12:1*, *12*(1), 1–10. https://doi.org/10.1038/s41598-022-05561-9

Lephuthing, M. C., Tolmay, V. L., Baloyi, T. A., Hlongoane, T., Oliphant, T. A., & Tsilo, T. J. (2021). Relationship of grain micronutrient concentrations and grain yield components in a doubled haploid bread wheat (Triticum aestivum) population. *Crop and Pasture Science*, *73*(2), 116–126. https://doi.org/10.1071/CP21206

Liakatas, A. (1997). Growth and yield of rainfed wheat on the seasonally dry Aegean Islands. *Undefined*, *58*(1–2), 43–56. https://doi.org/10.1007/BF00867431

Lobell, D. B. (2007). Changes in diurnal temperature range and national cereal yields. *Agricultural and Forest Meteorology*, *145*(3–4), 229–238. https://doi.org/10.1016/J.AGRFORMET.2007.05.002

Luo, Y., Zhang, Z., Cao, J., Zhang, L., Zhang, J., Han, J., Zhuang, H., Cheng, F., & Tao, F. (2022). Accurately mapping global wheat production system using deep learning algorithms. *International Journal of Applied Earth Observation and Geoinformation*, *110*, 102823. https://doi.org/10.1016/J.JAG.2022.102823

Meng, T., Carewa, R., Florkowski, W. J., & Klepacka, A. M. (2017). Analyzing Temperature and Precipitation Influences on Yield Distributions of Canola and Spring Wheat in Saskatchewan. *Journal of Applied Meteorology and Climatology*, *56*(4), 897–913. https://doi.org/10.1175/JAMC-D-16-0258.1

Miller, P., Lanier, W., & Brandt, S. (2001). Using Growing Degree Days to Predict Plant Stages E-5. *Montana State University.*, *MT200103 A*. http://store.msuextension.org/publications/AgandNaturalResources/MT200103AG.pdf

Morgounov, A., Abugalieva, A., & Martynov, S. (2013). Effect of Climate Change and Variety on Long-term Variation of Grain Yield and Quality in Winter Wheat in Kazakhstan. *Cereal Research Communications*, *42*(1), 163–172. https://doi.org/10.1556/CRC.2013.0047

Morgounov, A., Rosseeva, L., Koyshibayev, M., Morgounov, A., Rosseeva, L., & Koyshibayev, M. (2007). Leaf rust of spring wheat in Northern Kazakhstan and Siberia: incidence, virulence, and breeding for resistance*. *Australian Journal of Agricultural Research*, *58*(9), 847–853. https://doi.org/10.1071/AR07086

Nagai, T., & Makino, A. (2009). Differences Between Rice and Wheat in Temperature Responses of Photosynthesis and Plant Growth. *Plant and Cell Physiology*, *50*(4), 744–755. https://doi.org/10.1093/PCP/PCP029

Nicas, J. (2022). Ukraine war threatens to cause a global food crisis. *New York Times*. https://www.nytimes.com/2022/03/20/world/americas/ukraine-war-global-food-crisis.html

Pavlova, V. N., Varcheva, S. E., Bokusheva, R., & Calanca, P. (2014). Modelling the effects of climate variability on spring wheat productivity in the steppe zone of Russia and Kazakhstan. *Ecological Modelling*, *277*, 57–67. https://doi.org/10.1016/J.ECOLMODEL.2014.01.014

Porter, J. R., & Gawith, M. (1999). Temperatures and the growth and development of wheat: a review. *European Journal of Agronomy*, *10*(1), 23–36. https://doi.org/10.1016/S1161-0301(98)00047-1

Prasad, P. V. V., Pisipati, S. R., Ristic, Z., Bukovnik, U., & Fritz, A. K. (2008). Impact of Nighttime Temperature on Physiology and Growth of Spring Wheat. *Crop Science*, *48*(6), 2372–2380. https://doi.org/10.2135/CROPSCI2007.12.0717

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Sacks, W. J., Deryng, D., Foley, J. A., & Ramankutty, N. (2010). Crop planting dates: An analysis of global patterns. *Global Ecology and Biogeography*, *19*(5), 607–620. https://doi.org/10.1111/j.1466-8238.2010.00551.x

Salnikov, V., Turulina, G., Polyakova, S., Petrova, Y., & Skakova, A. (2015). Climate change in Kazakhstan during the past 70 years. *Quaternary International*, *358*, 77–82. https://doi.org/10.1016/J.QUAINT.2014.09.008

Saparov, A. (2014). Soil Resources of the Republic of Kazakhstan: Current Status, Problems and Solutions. *Environmental Science and Engineering*, 61–73. https://doi.org/10.1007/978-3-319-01017-5_2/COVER

Schauberger, B., Gornott, C., & Wechsung, F. (2017). Global evaluation of a semiempirical model for yield anomalies and application to within-season yield forecasting. *Global Change Biology*, *23*(11), 4750–4764. https://doi.org/10.1111/gcb.13738

Schauberger, B., Jägermeyr, J., & Gornott, C. (2020). A systematic review of local to regional yield forecasting approaches and frequently used data resources. *European Journal of Agronomy*, *120*, 126153. https://doi.org/10.1016/J.EJA.2020.126153

Schierhorn, F., Hofmann, M., Adrian, I., Bobojonov, I., & Müller, D. (2020). Spatially varying impacts of climate change on wheat and barley yields in Kazakhstan. *Journal of Arid Environments*, *178*, 104164. https://doi.org/10.1016/J.JARIDENV.2020.104164

Senapati, N., Semenov, M. A., Halford, N. G., Hawkesford, M. J., Asseng, S., Cooper, M., Ewert, F., van Ittersum, M. K., Martre, P., Olesen, J. E., Reynolds, M., Rötter, R. P., & Webber, H. (2022). Global wheat production could benefit from closing the genetic yield gap. *Nature Food 2022 3:7*, *3*(7), 532–541. https://doi.org/10.1038/s43016-022-00540-9

Shmelev, S. E., Salnikov, V., Turulina, G., Polyakova, S., Tazhibayeva, T., Schnitzler, T., & Shmeleva, I. A. (2021). Climate Change and Food Security: The Impact of Some Key Variables on Wheat Yield in Kazakhstan. *Sustainability 2021, Vol. 13, Page 8583*, *13*(15), 8583. https://doi.org/10.3390/SU13158583

Swinnen, J., Burkitbayeva, S., Schierhorn, F., Prishchepov, A. V., & Müller, D. (2017). Production potential in the "bread baskets" of Eastern Europe and Central Asia. *Global Food Security*, *14*, 38–53. https://doi.org/10.1016/J.GFS.2017.03.005

USDA. (2010). *Kazakhstan Agricultural Overview*. https://ipad.fas.usda.gov/highlights/2010/01/kaz_19jan2010/

USDA. (2022a). *Crop Calendars for Kazakhstan*. United States Department of Agriculture - Foreign Agricultural Service. https://ipad.fas.usda.gov/rssiws/al/crop_calendar/kz.aspx

USDA. (2022b). *Grain and Feed Annual- Republic of Kazakhstan*. https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Grain and Feed Annual_Nur-Sultan %28Astana%29_Kazakhstan - Republic of_KZ2022-0008.pdf

USDA. (2022c). *World Agricultural Production - Circular Series WAP 7-22; July 2022*. https://downloads.usda.library.cornell.edu/usda-esmis/files/5q47rn72z/3197zt68h/cz30r0703/production.pdf

van der Velde, M., Tubiello, F. N., Vrieling, A., & Bouraoui, F. (2012). Impacts of extreme weather on wheat and maize in France: Evaluating regional crop simulations against observed data. *Climatic Change*, *113*(3–4), 751–765. https://doi.org/10.1007/S10584-011-0368-2/FIGURES/8

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index. *Journal of Climate*, *23*(7), 1696–1718. https://doi.org/10.1175/2009JCLI2909.1

Wang, D., Li, R., Gao, G., Jiakula, N., Toktarbek, S., Li, S., Ma, P., & Feng, Y. (2022). Impact of Climate Change on Food Security in Kazakhstan. *Agriculture 2022, Vol. 12, Page 1087*, *12*(8), 1087. https://doi.org/10.3390/AGRICULTURE12081087

WITS. (2022). *Kazakhstan Cereals; meslin and wheat other than durum exports by country | 2018 | Data*. World Integrated Trade Solution. https://wits.worldbank.org/trade/comtrade/en/country/KAZ/year/2018/tradeflow/Exports/partner/ALL/product/100190

Zhang, P., Zhang, J., Chen, M., Costello, C., Dawson, D., Deschênes, O., Kolstad, C. D., Wang, C., Zhang, B., & Zheng, X. (2015). *Economic Impacts of Climate Change on Chinese Agriculture: The Importance of Humidity and Other Climatic Variables*.

Zhang, R., Shang, H., Yu, S., He, Q., Yuan, Y., Bolatov, K., & Mambetov, B. T. (2017). Tree-ring-based precipitation reconstruction in southern Kazakhstan, reveals drought variability since A.D. 1770. *International Journal of Climatology*, *37*(2), 741–750. https://doi.org/10.1002/JOC.4736