Effects of projected climate change on quantity and quality of soybean yield under different emission scenarios

Hamidreza Ahmadzadeh Araji[1,*], Aimrun Wayayok[1,2], Jahanfar Daneshian[3], Majid Mirzaei[4], Ali Reza Massah Bavani[5], C. B. S. Teh[6], Ahmad Fikri Abdullah[1] and Parisa Ahmadi[7]

[1]Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400, Serdang, Selangor Darul Ehsan, Malaysia
[2]Smart Farming Technology Research Centre (SFTRC), Faculty of Engineering, Universiti Putra Malaysia, 43400, Serdang, Selangor Darul Ehsan, Malaysia
[3]Seed and Plant Improvement Institute, Agricultural Research, Education and Extension Organization (AREEO) Karaj, Iran
[4]Department of Civil Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur, Malaysia
[5]Department of Irrigation and Drainage Engineering, Aburaihan Campus, University of Tehran, Iran
[6]Department of Land Management, Faculty of Agriculture, Universiti Putra Malaysia, 43400, Serdang, Selangor Darul Ehsan, Malaysia
[7]Institute of Graduate Studies Building, Institute of Ocean and Earth Science (IOES), University of Malaya, 50603 Kuala Lumpur

Soybean is one of the most important oilseed crops in the world. Its economic value is based on the concentration of protein and oil produced in the seeds. However, in climate change studies, a crop simulation model like AquaCrop is unable to predict the qualitative yield of crops. Therefore, this study aims to simulate qualitative soybean yield based on regression models between observed dry grain yield (Yd) from 12 treatments as independent variables with their corresponding observed values for oil and protein contents as dependent variables. The *P*-value (<0.05) and R^2 value of the linear regression model showed that oil content was positively regressed with yield, whereas protein content was negatively regressed with yield. On the other hand, predicted values of Yd from the calibrated AquaCrop model over general circulation models based on weighted multi-model ensemble means of five emission scenarios have been used for simulation of soybean oil and protein contents in the future. The results obtained by comparing historical period (1985–2010) to the future period (2020–2039) centred on the 2030s, showed that soybean oil content increased similarly as yield increased in the future period while protein content decreased inversely with yield. Overall, statistical indicators showed that the linear regression model performed well to predict the soybean oil and protein content when AquaCrop model not able to simulate the qualitative yield.

*For correspondence. (e-mail: hamid_araji@yahoo.com)

Keywords: Dry grain yield, linear regression, oil contents, protein contents, soybean.

SOYBEAN is one of the most important oilseed crops in the world. In general, dry soybeans by weight have approximately 20% fat (oil) and 36% protein. The role of oil and protein from soybean in human diet and industrial application cannot be neglected. The economic value of soybean is based on the concentration of protein and oil produced in seeds.

Some studies show a correlation between mean daily temperature and soybean oil and protein content[1,2].

Thomas *et al.*[3] studied the effects of climate change on the composition of soybean seeds by focusing on increased CO_2 and elevated temperature. The findings showed that oil concentration was highest at 32/22°C (day/night) and decreased with further increase in temperature. Taub *et al.*[4] did the meta-analysis for 228 studies to observe the effect of elevated atmospheric CO_2 rate on the protein concentration of major food crops. The results showed that increased CO_2 led to a 1.4% reduction in protein concentration of soybean. On the other hand, AquaCrop is a crop simulation model which was designed by FAO as a water-driven model to simulate and predict quantitative yield response in relation to water supply, but the model is incapable to simulate and predict qualitative yield. Some linear regression models can be suggested as an additional function to simulate and predict oil and protein contents of soybean in the future. Therefore, this study aims to evaluate proposed linear regression models to simulate oil and protein contents and consequently predict their values under projected climate change scenarios.

The study area with a cold semi-arid climate was located in experimental fields within the latitude of 35°47′30″N and 35°48′0″N, and longitude of 50°54′30″E and 50°55′30″E, at research department of oilseed crops, Karaj seed and plant improvement institute, Iran. In this study, the soybean dry grain yield (Yd) was collected from a field experiment in 2010, which were implemented using a completely randomized block design with three replications in 2010. Four soybean varieties including L17 (V1), Williams*Hobbit (W*H) (V2), M9 (V3), and M7 (V4) under three irrigation treatments including without water stress (I1), mild water stress (I2), and severe water stress (I3) were sown on 27 June. The oil and protein contents were measured for each treatment by using inframatic grain analyser.

Both protein and oil were a function of the rate of dry matter accumulation in soybean seed[5]. Accordingly, environmental influences on seed size can lead to changes in overall protein and oil concentration. Therefore, two linear regression models were carried out between observed Yd from twelve treatments as independent variables with their corresponding observed values of oil and protein contents separately as dependent variables.

The results of simulated Yd from AquaCrop model that had already been calibrated by Araji *et al.*[6] were applied to equations derived from regression models to simulate oil and protein contents in percentage. Next, the observed and simulated oil and protein content were respectively multiplied to the observed and simulated Yd to determine the weight of oil and protein for each treatment. Some statistical tests were implemented between the observed and simulated values to investigate the validity of the linear regression model.

In order to predict the percentage of oil and protein content for the future period (2020–2039) centered on 2030s, the mean of simulated Yd in the historical (1985–2010) and the future period (2030s) under five emission scenarios (B1, A1B, A2, RCP2.6, and RCP8.5) was applied to the equations of the linear regression models.

Based on observed and simulated values, statistical indicators were applied to evaluate the accuracy of linear regression models. Equations (1) to (6) for RMSE, RMSE*n*, the index of agreement (*d*), mean bias error (MBE), model efficiency (ME), and Nash Sutcliff efficiency (NSE) respectively were utilized in judging the model performance.

$$\text{RMSE} = \left[\frac{\sum_{i=1}^{n}(Pi - Oi)^2}{n}\right]^{1/2}, \tag{1}$$

$$\text{RMSE}_n = 100\frac{\text{RMSE}}{\bar{O}}, \tag{2}$$

$$d = 1 - \frac{\sum_{i=1}^{n}(Pi - Oi)^2}{\sum_{i=1}^{n}(|Pi - \bar{O}| + |Oi - \bar{O}|)^2}, \tag{3}$$

$$\text{MBE} = \frac{1}{n}\left[\sum_{i=1}^{n} Pi - Oi\right], \tag{4}$$

$$\text{ME} = \frac{\sum_{i=1}^{n}(Oi - \bar{O})^2 - \sum_{i=1}^{n}(Pi - Oi)^2}{\sum_{i=1}^{n}(Oi - \bar{O})^2}, \tag{5}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}(Oi - Pi)^2}{\sum_{i=1}^{n}(Oi - \bar{O})^2}, \tag{6}$$

where *Pi* and *Oi* are the predicted (simulated) and observed values, $\bar{O}$ is the mean observed values, and *n* is the number of measurements.

RMSE values close to zero show that the residual estimation error decreases. On the other hand RMSE*n* values are excellent if smaller than 10%, good if between 10% and 20%, fair if between 20% and 30% and poor if larger than 30% (ref. 7). The index of agreement (*d*) ranges from 0 to 1, and the model simulates the appointed parameter better for the value close to 1 (ref. 8). Moreover, a positive value in MBE represents overestimation in estimated values, whereas the negative values indicate underestimation in estimated values[9]. ME ranges from negative infinity to positive 1 and the closer values to 1 show more model robustness[10]. The model also shows better simulation efficiency when the value of NSE approaches to 1 (ref. 11).

The results of the regression models between observed Yd from twelve treatments with their corresponding observed values of oil and protein contents are respectively shown in Figures 1 and 2. The *P*-value (<0.05) and R^2 value of the linear regression model showed that oil content was positively regressed with Yd, whereas protein content was negatively regressed with Yd. The results were consistent with results obtained by Chung *et al.*[12], which found that oil content was positively correlated with soybean yield whereas the protein was negatively correlated with yield.

The observed and simulated values of Yd, oil and protein contents for different treatments in 2010 (Table 1) show that simulated oil content decreased in treatments with water stress which was in line with decreasing trend of simulated yield obtained by Araji *et al.*[6], while protein content inversely increased with water stress. The results

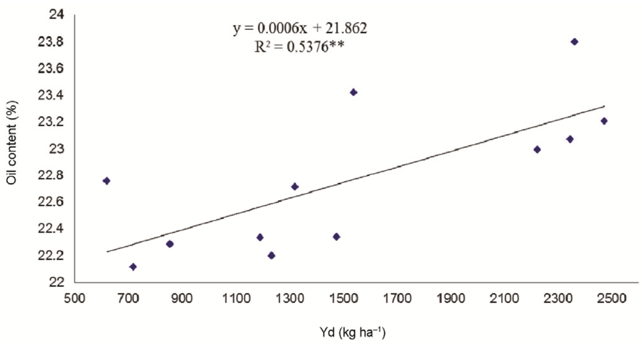

Figure 1. Linear regression model between soybean Yd and oil content.

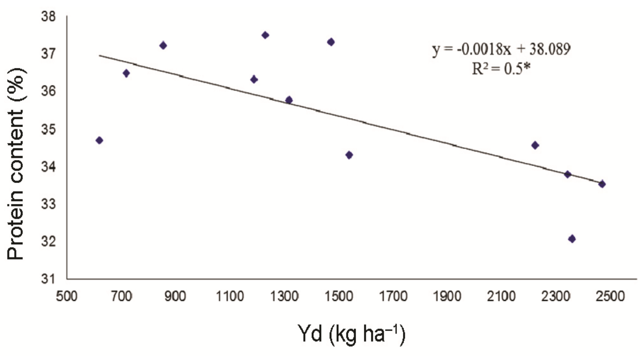

Figure 2. Linear regression model between soybean Yd and protein content.

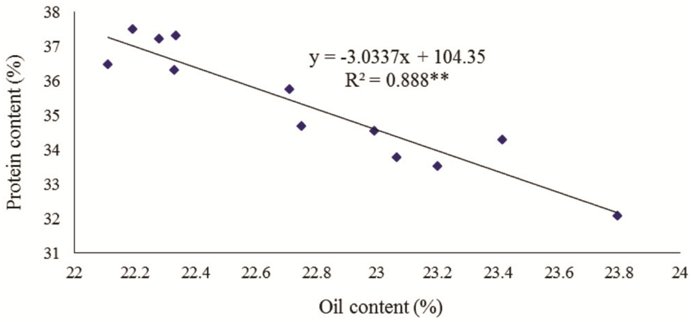

Figure 3. Linear regression model between soybean oil and protein contents.

Table 1. Observed and simulated values of Yd, oil and protein contents for different treatments in 2010

	Yd (kg ha^{-1})		Oil content (%)		Protein content (%)	
Varieties	Observed	Simulated	Observed	Simulated	Observed	Simulated
V1I1	2473	2570	23.20	23.40	33.53	33.46
V1I2	1540	1421	23.41	22.71	34.29	35.53
V1I3	720	750	22.11	22.31	36.47	36.74
V2I1	2346	2392	23.07	23.30	33.79	33.78
V2I2	1320	1355	22.71	22.68	35.75	35.65
V2I3	621	710	22.75	22.29	34.68	36.81
V3I1	2224	2245	22.99	23.21	34.55	34.05
V3I2	1233	1285	22.19	22.63	37.50	35.78
V3I3	856	721	22.28	22.29	37.22	36.79
V4I1	2363	2390	23.80	23.30	32.07	33.79
V4I2	1475	1416	22.34	22.71	37.31	35.54
V4I3	1191	1010	22.33	22.47	36.31	36.27

Table 2. Statistical tests between the weight of oil and protein from the observed data versus simulated data by the linear regression model

	Oil (kg ha^{-1})		Protein (kg ha^{-1})	
	Observed	Simulated	Observed	Simulated
Mean	351	349	532	530
SD	152	160	204	216
Minimum	574	601	215	261
Maximum	141	158	792	808
N	12		12	
RMSE	22		37	
RMSEn (%)	6		7	
d	0.995		0.992	
MBE	–1		–2	
ME	0.979		0.968	
NSE	0.979		0.968	

agree with Dornbos Jr and Mullen[13], who found that severe drought, increased protein content by 4.4%, while oil content decreased by 2.9%. As drought stress increased, protein content increased linearly and oil content decreased linearly at each air temperature.

The linear regression model between soybean oil and protein contents (Figure 3) showed that protein contents of soybean were negatively regressed with seed oil contents, which were in agreement with similar results reported by Chung *et al.*[12]. The results were also in agreement with Piper and Boote[14], who declared that oil and protein concentration were inversely related in response to heat stress during seed filling period.

The model prediction accuracy tested by statistical indicators is shown in Table 2. The RMSE*n* and *d* values indicated that the model had excellent performance for simulation of oil and protein. These indices also showed that the model could simulate oil better than protein. The results of MBE showed that the model underestimated both oil and protein. The ME values showed the robustness of the linear regression model, and NSE showed that model performance is excellent in the simulation of oil and protein contents. Therefore, the linear regression model had enough validity to predict oil and protein contents in the future period.

The mean annual yields over the historical (1985–2010) and the future period (2030s) under different emission scenarios simulated by AquaCrop model are shown in Table 3. The results of simulated yield in Table 3 were applied to the equations from the linear regression model in Figures 1 and 2 to estimate oil and protein content of each treatment in the historical and the future period respectively. Future changes in soybean oil and protein content under different emission scenarios are shown in Figure 4. Under projected climate change for all

Table 3. Mean annual yield over the historical (1985–2010) and the future period (2030s) under different emission scenarios simulated by AquaCrop model

Varieties	Historical (1985–2010)	2.6 (2030s)	8.5 (2030s)	A2 (2030s)	A1B (2030s)	B1 (2030s)
V1I1	2359	2672	2765	2772	2824	2695
V1I2	1304	1477	1553	1553	1610	1511
V1I3	641	763	809	738	932	720
V2I1	2211	2507	2596	2602	2649	2531
V2I2	1241	1405	1476	1476	1528	1435
V2I3	619	733	784	731	890	705
V3I1	2083	2360	2445	2452	2500	2383
V3I2	1196	1354	1422	1419	1470	1377
V3I3	639	757	798	742	902	724
V4I1	2225	2520	2610	2622	2660	2560
V4I2	1319	1484	1557	1567	1612	1527
V4I3	943	1061	1116	1113	1173	1086

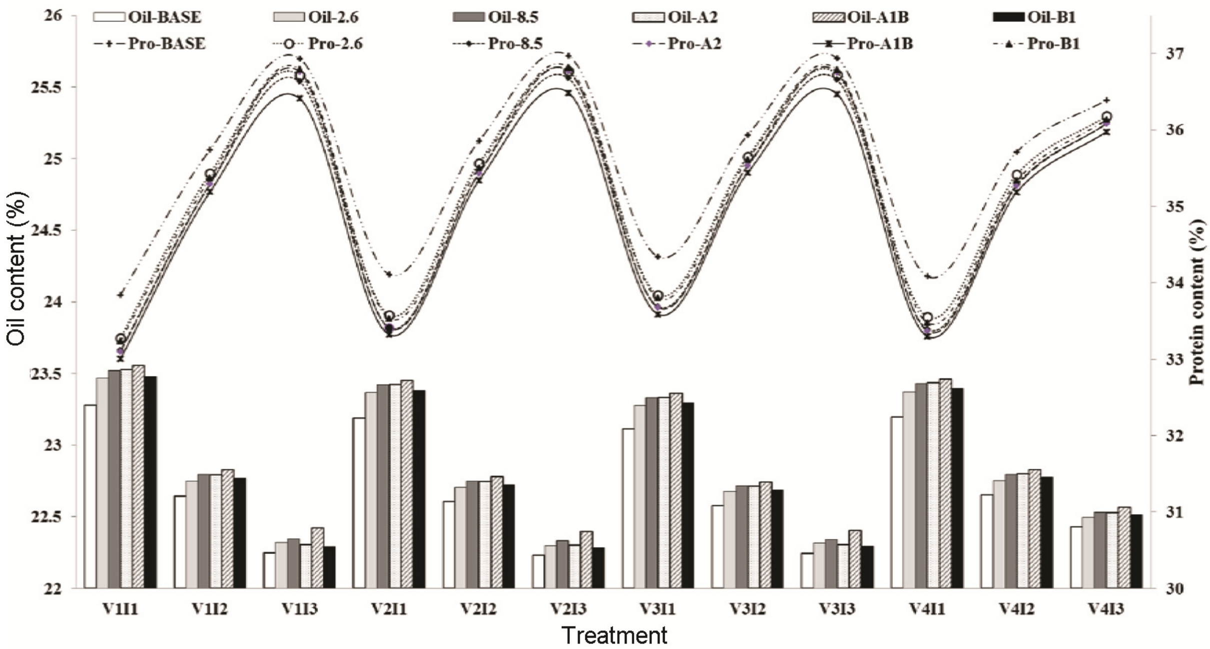

Figure 4. Future changes in soybean oil and protein content under different emission scenarios estimated by the linear regression model.

treatments and emission scenarios, the oil content increased while protein content decreased.

The mean of oil content within all emission scenarios under the control treatment showed that L17 had the highest increase, followed by M7, W*H and M9, in that order. The estimated oil content under mild water stress showed that M7 had the highest increase, followed by L17, W*H and M9, in that order. Furthermore, the estimated oil content under severe water stress showed that M7 had the highest increase, followed by L17, M9 and W*H, in that order.

The mean of protein content within all emission scenarios under the control treatment showed that L17 had the highest decrease, followed by M7, W*H and M9, in that order. The estimated protein content under mild water stress showed that M7 had the highest decrease, followed by L17, W*H and M9, in that order. Moreover, the estimated protein content under severe water stress showed that M7 had the highest decrease, followed by L17, M9 and W*H, in that order.

Based on the results derived from Table 4 (ref. 7), the increased mean temperatures were in agreement with those obtained by Howell and Cartter[15] who reported that high temperatures during soybean seed filling period could elevate seed oil. The results also were consistent with findings by Divsalar[16], who reported that the minimum oil content and maximum protein content were found in treatment with water stress during seed filling period. The increased CO_2 rate due to the future climate change decreased the protein concentration of soybean[17], which were in agreement with the present findings.

Oil concentration increased by an increment in temperature with optimum at 25°C to 28°C, above which the oil concentration declined[13,14,18]. At temperatures greater than 28°C, the protein content increased linearly with temperature[13,18,19]. Following Table 4 and Figure 4, it can be concluded that under projected climate change, mean temperature during seed feeling stage exceeds the optimum threshold (above 25°C) for seed oil storage which were in agreement with Wolf *et al.*[19] who found oil content was positively correlated with increasing temperature from 25°C to 36°C.

The linear regression model had validity to predict the qualitative yield of soybean. The results can be linked to the results of Yd simulated by AquaCrop model to estimate oil and protein content of soybean. The results of the oil content were precisely in line with those obtained for yield under projected climate change scenarios. However, protein content decreased due to its reversed relation with soybean yield and oil content. Moreover, increased temperature in cold semi-arid climate of the study area may prepare an optimal temperature for seed oil storage during the flowering and seed filling periods.

It has been suggested that under projected climate change, not only estimation of qualitative yield in the future is important but also other crop production parameters such as oil content, protein content and water productivity should be estimated simultaneously among different varieties to enhance the productivity of soybean yield in the future. The present results are practical for agricultural decision-makers to cultivate the best varieties in terms of yield (quantitative and qualitative) and irrigation levels and consequently to adapt themselves with the future climate change. Under future climate changes, this simulation allows irrigation engineers to manage the supply of water matching with a suitable variety of soybean to increase the production of oil and protein for the best pricing values.

1. Kumar, V., Rani, A., Solanki, S. and Hussain, S., Influence of growing environment on the biochemical composition and physical characteristics of soybean seed. *J. Food Compos. Anal.*, 2006, **19**, 188–195.
2. Song, W. *et al.*, Analyzing the effects of climate factors on soybean protein, oil contents, and composition by extensive and high-density sampling in china. *J. Agric. Food Chem.*, 2016, **64**, 4121–4130.
3. Thomas, J., Boote, K., Allen, L., Gallo-Meagher, M. and Davis, J., Elevated temperature and carbon dioxide effects on soybean seed composition and transcript abundance. *Crop Sci*., 2003, **43**, 1548–1557.
4. Taub, D. R., Miller, B. and Allen, H., Effects of elevated CO_2 on the protein concentration of food crops: A meta-analysis. *Glob. Change Biol.*, 2008, **14**, 565–575.
5. Pipolo, A. E., Sinclair, T. R. and Camara, G. M., Effects of temperature on oil and protein concentration in soybean seeds cultured in vitro. *Ann. Appl. Biol*., 2004, **144**, 71–76.
6. Araji, H. A., Wayayok, A., Bavani, A. M., Amiri, E., Abdullah, A. F., Daneshian, J. and Teh, C., Impacts of climate change on soybean production under different treatments of field experiments considering the uncertainty of general circulation models. *Agric. Water Manage.*, 2018, **205**, 63–71.
7. Raes, D., Steduto, P., Hsiao, T. C. and Fereres, E., Fao, land and water division Rome, Italy, 2012.
8. Stricevic, R., Cosic, M., Djurovic, N., Pejic, B. and Maksimovic, L., Assessment of the fao aquacrop model in the simulation of rainfed and supplementally irrigated maize, sugar beet and sunflower. *Agric. Water Manage.*, 2011, **98**, 1615–1621.
9. Iqbal, M. A. *et al.*, Evaluation of the fao aquacrop model for winter wheat on the north china plain under deficit irrigation from field experiment to regional yield simulation. *Agric. Water Manage.*, 2014, **135**, 61–72.
10. Araya, A., Habtu, S., Hadgu, K. M., Kebede, A. and Dejene, T., Test of aquacrop model in simulating biomass and yield of water deficient and irrigated barley (*Hordeum vulgare*). *Agric. Water Manage.*, 2010, **97**, 1838–1846.
11. Mustafa, S. M. T., Vanuytrecht, E. and Huysmans, M., Combined deficit irrigation and soil fertility management on different soil textures to improve wheat yield in drought-prone Bangladesh. *Agric. Water Manage.*, 2017, **191**, 124–137.
12. Chung, J. *et al.*, The seed protein, oil, and yield qtl on soybean linkage group *i*. *Crop Sci.*, 2003, **43**, 1053–1067.
13. Dornbos Jr, D. and Mullen, R., Soybean seed protein and oil contents and fatty acid composition adjustments by drought and temperature. *J. Am. Oil Chem. Soc.*, 1992, **69**, 228–231.
14. Piper, E. L. and Boote, K. I., Temperature and cultivar effects on soybean seed oil and protein concentrations. *J. Am. Oil Chem. Soc.*, 1999, **76**, 1233–1241.
15. Howell, R. W. and Cartter, J. L., Physiological factors affecting composition of soybeans: Ii. Response of oil and other constituents of soybeans to temperature under controlled conditions 1. *Agron. J.*, 1958, **50**, 664–667.
16. Divsalar, M., Study the effect of drought stress on oil percent, protein percent and fatty acids composition of soybean grain. *J. Plant Ecophysiol*., 2017, **8**, 44–55.
17. DaMatta, F. M., Grandis, A., Arenque, B. C. and Buckeridge, M. S., Impacts of climate changes on crop physiology and food quality. *Food Res. Int.*, 2010, **43**, 1814–1823.
18. Gibson, L. and Mullen, R., Soybean seed composition under high day and night growth temperatures. *J. Am. Oil Chem. Soc*., 1996, **73**, 733–737.
19. Wolf, R., Cavins, J., Kleiman, R. and Black, L., Effect of temperature on soybean seed constituents: oil, protein, moisture, fatty acids, amino acids and sugars. *J. Am. Oil Chem. Soc.*, 1982, **59**, 230–232.

ACKNOWLEDGEMENTS. We acknowledge the Karaj Seed and Plant Improvement Institute, Department of Oil Seed Crops for the data of field experiments and measurements of soybean oil and protein contents. This project was funded by Universiti Putra Malaysia (grant vot no. 9573700). The support from all the UPM staff is greatly appreciated.

Received 7 August 2019; accepted 22 August 2019

doi: 10.18520/cs/v118/i1/103-107