Environ Syst Decis (2017) 37:34–41
DOI 10.1007/s10669-016-9617-2

Robust optimization evaluation of reliance on locally produced foods

Bayram Dundar[1] · Christine Costello[2] · Ronald G. McGarvey[3]

Published online: 11 November 2016

Abstract While local food production may be beneficial in terms of developing the local economy and reducing greenhouse gases from transportation, sustainability strategies focused on local food production may generate their own risks due to yield variability. We have developed a robust optimization (RO) model to determine the minimum amount of land (cropland and pasture) required to grow food items that would satisfy a local population's (accounting for gender and age) calorie and nutrient needs. This model has been applied to Boone County, Missouri, which has a population of approximately 170,000. Boone County is 1790 km^2, with 16% of the land defined as cropland and 30% defined as pasture. The model includes 27 nutrients from 17 potential foods that could be produced: six fruits and vegetables, five grains and six animal-sourced foods. Yield estimates are based on the predominate methods of agriculture in the USA. We first run our model assuming no variability, using the midpoint yield estimates. Then, to quantify uncertainty in yield for different food types, we use historical yield data over 10 years to estimate this variability and run our RO model under these variability estimates. We compare the two model results to illustrate the impact of data uncertainty on meeting sustainable local food for communities. Solutions suggest that nutrition needs can be met for the Boone County population within the land area defined.

Keywords Land use · Local food production · Robust optimization

✉ Ronald G. McGarvey
mcgarveyr@missouri.edu

Bayram Dundar
bd5zc@mail.missouri.edu

Christine Costello
costelloc@missouri.edu

1 Department of Industrial and Manufacturing Systems Engineering, University of Missouri, Columbia, MO, USA

2 Department of Bioengineering, University of Missouri, Columbia, MO, USA

3 Department of Industrial and Manufacturing Systems Engineering, and Truman School of Public Affairs, University of Missouri, Columbia, MO, USA

1 Introduction

The local food movement has flourished over the past two decades. The USDA reported that in 2012, 163,675 farms or 7.8% of U.S. farms were marketing foods locally (Low et al. 2015). Local and regional food markets are praised for improving rural economies and are often thought to be more environmentally sustainable. Evidence that local food production benefits rural economies is mounting (Johnson et al. 2014), though conclusive findings are hindered by data collection (Low et al. 2015). The environmental benefits, particularly those associated with reduced transportation, are not as clear (Weber and Matthews 2008; Mariola 2008). And there are relatively few investigations of the feasibility of meeting a population's food needs through local food systems, particularly in dense, urban environments (Peters et al. 2009; Zumkehr and Campbell 2015). There is even less inquiry as to how localized food systems might introduce risk to supported populations with studies using either average or theoretical maximum yields only (Zumkehr and Campbell 2015). There is very little work done to assess how populations would meet their nutritional needs throughout the year and what staple crops would be required.

There is no universally accepted definition of a "local food." Generally speaking, the term refers to minimizing the geographical distance between food producers and consumers; however, it can also refer to social, environmental and supply chain characteristics (Martinez et al. 2010). Perhaps one of the most widely applied distances cited in relation to a food being considered local is 100 miles (160 km). Others consider a food to be local if it is grown within their home county or bordering counties (Martinez et al. 2010). There are also alternative interpretations across countries: Some European nations, having less land area than some US counties, have a more limiting definition of what constitutes a local food. When expanding the definition from local to regional, the 2008 Farm Bill required that the distance from producer to consumer be less than 400 miles (Martinez et al. 2010). Other qualities that consumers attribute to "local" foods are foods that are produced using fewer pesticides and manufactured fertilizers and that apply fair farm labor practices and animal welfare (Martinez et al. 2010) and give consumers a sense of a direct connection to their food (Mount 2012).

A few other studies have analyzed the potential for cities in the United States (U.S.) to meet the nutritional needs of population centers from local foodsheds. Generally, these studies suggest that the majority of populations can meet their nutritional needs (defined in terms of required calories) with foods grown within a 100-mile radius (some assuming maximum achievable yields, others assuming average yields) with the exception of the most populous cities in the country, e.g., New York City, NY or Los Angeles, CA (Peters et al. 2009; Zumkehr and Campbell 2015). Another recent study (Monaco et al. 2016) examined the ability of local food production to satisfy the demand for staple foods in the metropolitan area of Milan, Italy.

In addition to considering whether or not local food systems are optimal for local economies or the environment, it is important to understand the limitations of localizing the food supply. In this work, we ask the question: "Is it feasible given land resources to meet the nutritional requirements of a given population within a local context?" To explore this question, we employed a linear programming (LP) optimization framework inspired by (Briend et al. 2003). To consider how a solution may be influenced by fluctuations in yield, this method was extended to include robust optimization (RO) techniques that select optimal solutions given the observed range in yields for the selected crops. This study's inclusion of high and low yields observed over a multi-year period as well as explicit inclusion of many nutrients provides additional details into this important question. We utilize both models, allowing us to obtain a "non-robust" and a "robust" solution for any location of interest.

To demonstrate the method, we selected Boone County, MO, and the most restrictive definition of local food. Note that we have conservatively assumed that there would be no trade and the potential for novel crops and cultivation are not included in this initial study. Moreover, this analysis assumes a continuation of the industrial monoculture model that currently dominates production in Boone County. To the extent that alternative production models can achieve increased yields on equal land, this analysis presents an a fortiori argument on the potential for local food production to satisfy a local community's needs, having accounted for yield variability.

2 Materials and methods

In this analysis, optimization models are employed to identify the minimum amount of cropland and pastureland required to produce a set of food items that satisfy the local population's nutritional requirements. This approach is consistent with an assumption that any existing cropland and pastureland that is not utilized to satisfy the local population's nutritional requirements can be put to some better use than agricultural production. We first present the formulation of the LP model, which utilizes the following sets and indices. I is the set of all food types, indexed by i. K is the set of all land types, indexed by k. J is the set of all nutrients having a minimum requirement level (e.g., calcium), indexed by j. R is the set of nutrients having a maximum allowable level (e.g., cholesterol), indexed by r.

Parameters used in the mathematical model are defined as follows: w_i is land in m^2 required to produce one kg of food type i per year, under the assumption that any unit of land area can be utilized for at most one food type ; n_{ij} is the amount of nutrient j obtained from one kg of food type i; t_{ir} is the amount of nutrient r obtained from one kg of food type i; β_j is the aggregate minimum required amount of nutrient j; θ_r is aggregate maximum allowed amount of nutrient r; α_{ik} is the area in m^2 of land type k required to produce 1 kg annually of food type i; L_k represents the total area available of land type k; γ is a scalar between 0 and 1 that defines the maximum amount of any single food type that can be grown (measured on a weight basis).

Decision variable x_i is the kg of food i to be produced to meet the population's annual nutrient needs. Note that it is assumed that any food that is produced can be stored in such a manner as to allow for its consumption throughout the year, although we executed the optimization models, assuming that 30% of all food production was lost to waste, consistent with estimates of U.S. food waste (Buzby et al. 2014).

2.1 Nominal objective function and constraints

The objective of the optimization model is to minimize the total amount of cropland and pasture needed to produce food satisfying the community's nutritional requirements.

$$\min \sum_{i \in I} w_i x_i \tag{1}$$

The population's nutritional requirements are represented by the following two inequality constraints. Constraint (2) is of type greater than minimum required amount, and constraint (3) is of type less than maximum allowable amount.

$$\sum_{i \in I} n_{ij} x_i \geq \beta_j \quad \forall j \in J \tag{2}$$

$$\sum_{i \in I} t_{\mathrm{ir}} x_i \leq \theta_r \quad \forall r \in R \tag{3}$$

In a similar fashion, the available amount of each land type is limited by a maximum allowable constraint, as defined by constraint (4).

$$\sum_{i \in I} \alpha_{ik} x_i \leq L_k \quad \forall k \in K \tag{4}$$

Additionally, we impose a diversification constraint (5) to prevent any single food product from comprising too large a percentage of the overall production amount, in an effort to avoid the risks of overreliance on any single food. Observe that, in order to maintain a linear function, we have rewritten the constraint such that the summation term does not appear in a denominator.

$$\gamma * \left(\sum_{i \in i} x_i \right) \geq x_i \quad \forall i \in I \tag{5}$$

2.2 Robust counterpart of LP model

The LP uses data parameter a_{ik} in constraint (4) to translate food production into land requirements. An LP model such as this can only utilize a point estimate for this yield parameter. Because yields are known to vary greatly across years, we developed an RO model that accounts for this uncertainty, using the ellipsoidal uncertainty set approach to yield parameters developed by Ben-tal, et al. (2009).

$$a_{ik} = a_{ik} + \rho_{ik} \hat{a}_{ik} \tag{6}$$

Here, a_{ik} is the nominal value of the yield parameter for food type i and land type k, which is calculated as the midpoint of the maximum and minimum yield data observed over multi-year period for which data were available, and $\hat{a}_{ik}$ is the maximum allowable deviation, equal to one half of the difference between these maximum and minimum yield values. ρ_{ik} is our measure of uncertainty, which is constrained for each land type k such that $\sqrt{\sum_{i \in I} (\rho_{ik})^2} \leq 1$. Accordingly, each a_{ik} takes values between $[a_{ik} - \hat{a}_{ik}, a_{ik} + \hat{a}_{ik}]$.

In our RO model, we replace constraint (4) with the following constraint.

$$\sqrt{\sum_{i \in I} (\hat{a}_{ik} x_i)^2} \leq L_k - \sum_{i \in I} a_{ik} x_i, \quad \forall k \in K \tag{7}$$

2.3 Data

2.3.1 *Land requirements for food types included*

Crop yields were obtained from the USDA (USDA NASS 2016); these data are presented in Table 1. Foods were selected based on their current prevalence in the American agricultural system and for their nutritional qualities. Where possible, Boone County yield data were used, if that was not available MO state or national data were used. The amount of land used for each animal-based product was estimated using the approach of Costello et al. (2015), with the exception of land area required for beef, which followed USDA recommendations (USDA NRCS 2009). For animal-based products, we assumed a yield variability range equal to 50% of the midpoint value. We incorporated a total food loss of 30% in yield parameters (Buzby et al. 2014), assuming that this waste occurs between production and consumption. The use of conventionally grown crops and conventionally raised livestock may be a conservative estimate of the potential nutritional output from land as alternative management options may be capable of yielding higher nutritional content. For example, intercropping of grains and legumes can lead to higher grain yields and protein content (Bedoussac et al. 2015).

2.3.2 *Land availability*

A raster layer of National Land Cover Database (NLCD) for 2011 was used to represent Boone County land area (NLCD 2011). These raster data were input to a GIS to calculate the amount of cropland and pasture at a 30 × 30 m resolution, as depicted in Fig. 1. Total pasture and cropland are approximately 534 and 282 km^2, respectively.

2.3.3 *Nutrition contained in food types*

Nutrition data for each food type included were retrieved from the USDA national nutrient database (USDA ARS 2016). Initial attempts to solve the LP model determined the problem to be infeasible because there was no set of

Table 1 Data type, years, nominal and range values for each food

Foods	Data type	Years	Midpoint (m^2/kg)		Range (m^2/kg)	
			Cropland	Pasture	Cropland	Pasture
Corn	State	1995–2015	1.49	–	0.63	–
Soy	State	1995–2015	4.30	–	1.10	–
Wheat	State	1995–2015	3.13	–	0.69	–
Oats	State	1995–2015	5.18	–	1.02	–
Barley	County	1997–2012*	3.65	–	0.40	–
Apples	State	2007–2014	0.69	–	0.38	–
Grapes	State	2007–2014	1.88	–	0.61	–
Peaches	State	2007–2014	1.55	–	0.31	–
Tomatoes	National	1998–2014	0.28	–	0.02	–
Potatoes	State	1995–2014	0.35	–	0.12	–
Broccoli	National	2006–2014	0.52	–	0.03	–
Beef	National	2002	8.89	30.00	4.44	15.00
Chicken			9.30	–	4.65	–
Eggs			7.30	–	3.65	–
Milk			1.29	3.70	0.64	1.85
Turkey			10.40	–	5.20	–
Pork			21.89	–	10.94	–

* Data for barley were only available in 5-year intervals (i.e., 1997, 2002, 2007 and 2012)

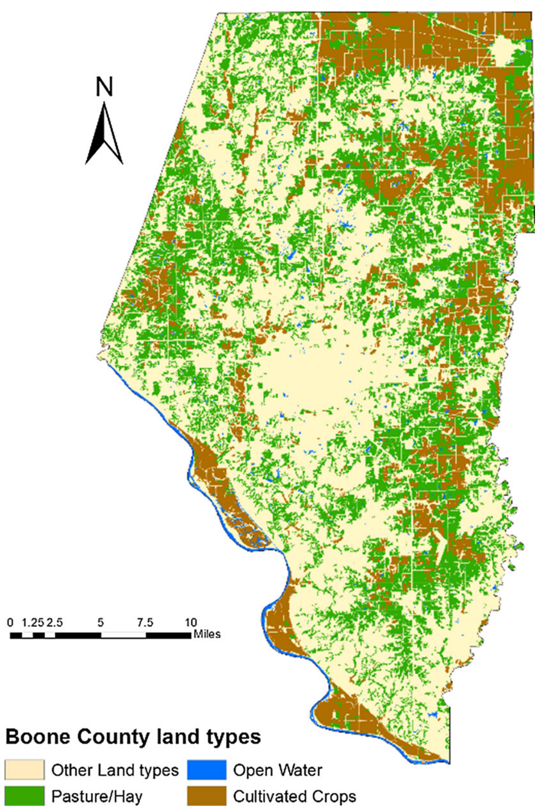


Fig. 1 Map showing the distribution of cropland and pasture in Boone County, MO

food items that could be produced on the available land that was capable of satisfying total vitamin D requirements. In order to allow the population's vitamin D requirements to be satisfied, we assumed that corn, wheat and oats were enriched with vitamin D at a level of 10 mcg per kg, well below the levels contained in typical breakfast cereals (USDA ARS 2016).

2.3.4 Population's nutritional requirements

USDA Center for Nutrition Policy and Promotion has proposed daily food nutrient requirements by age and gender groups (USDA CNPP 2016). Boone County population, by age and sex, is estimated from the 2014 U.S. Census Bureau database (US Census, 2016). The product of these values was used to estimate the county's total annual requirements for each nutrient.

3 Results

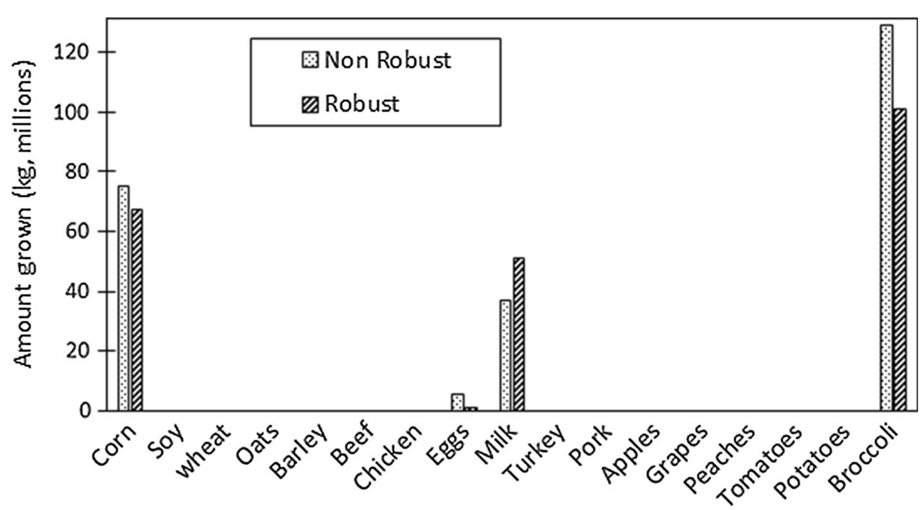
When using the LP model with average yield (kg food per square meter of land utilized) values, our model was able to identify a solution that satisfied the nutritional requirements of Boone County's population using only the land assumed to be available in Boone County, without accounting for diet diversification considerations (i.e., with parameter α set equal to 1). This solution utilized only four food types (corn, eggs, milk and broccoli) and consumed a total of 269 and 138 km^2 of cropland and pasture, respectively. The RO model (accounting for variations in yields) found a solution that utilizes the same four food types (in different amounts), consuming a total of 228 and 190 km^2 of cropland and pasture, respectively, based on average yields; in the worst-case scenario considered the RO solution would consume all 282 km^2 of available cropland (pastureland would not be constrained in the worst case). The amounts of each food type produced for each solution appear in Fig. 2.

In an effort to increase the diversity of the solution diet, we imposed diversification constraint (5) for smaller values of α and were able to obtain a solution that had no single food type accounting for more than 17% of the total food grown, by weight. This solution utilized nine food types (corn, soy, wheat, eggs, milk, apples, tomatoes, potatoes and broccoli) and consumed a total of 282 km^2 of cropland (the total amount available) and 191 km^2 of pasture. However, when considering the variability observed in historical yield values, the RO model was unable to find a feasible solution that satisfied all requirements at a diversification level of 17%. This is because the production of wheat and apples cannot be achieved, in light of yield uncertainty, at the levels necessary to allow for reduced production of corn, tomatoes, potatoes and broccoli at a level commensurate with a 17% diversification level. The amounts of each food type produced for this non-robust solution appear in Fig. 3.

The RO model could not identify a feasible solution with a diversification level less than 21%. At this 21% diversification level, the non-robust (i.e., average yields) solution utilized seven food types (corn, wheat, eggs, milk, tomatoes, potatoes and broccoli) and consumed a total of 259 and 181 km^2 of cropland and pasture, respectively. The RO solution utilized five food types (wheat and eggs were excluded from the list of seven food types for the corresponding non-robust solution) and consumed a total of 228 and 219 km^2 of cropland and pasture, respectively, based on average yields; in the worst-case scenario considered this RO solution would consume all 282 km^2 of available cropland (pastureland would not be constrained in the worst case). The amounts of each food type produced for each of these solutions appear in Fig. 4.

It is apparent on observation that the diet associated with these optimal production levels differs greatly from the typical American's diet. The optimization models presented in this study can be utilized to identify the impact of requiring production of certain food items. Consider, for example, a scenario in which the production of beef is maximized, subject to all of the other model constraints. In the RO model, the maximum amount of beef that can thusly be generated for Boone County is 3.8 million kg, corresponding to 22.2 kg/person/year, considerably less than the US per capita production of 34.6 kg/person in 2014 (USDA ERS 2016). This result can be achieved as part of a solution that had no single food type accounting for more than 40% of the total food grown, by weight. At this 40% diversification level, the non-robust (i.e., average yields) solution utilized five food types (corn, soy, beef, milk and broccoli) and consumed a total of 282 and 260 km^2 of cropland and pasture, respectively, with all available cropland being utilized. The RO solution utilized four food types (soy was excluded from the list of five food types for the corresponding non-robust solution) and consumed a total of 225 and 387 km^2 of cropland and pasture, respectively, based on average yields; in the worst-case scenario considered this RO solution would consume all 282 km^2 of available cropland, along with all 534 km^2 of pasture. The amounts of each food type produced for each of these solutions appear in Fig. 5.

Fig. 2 Optimal production amounts, with no diversification level



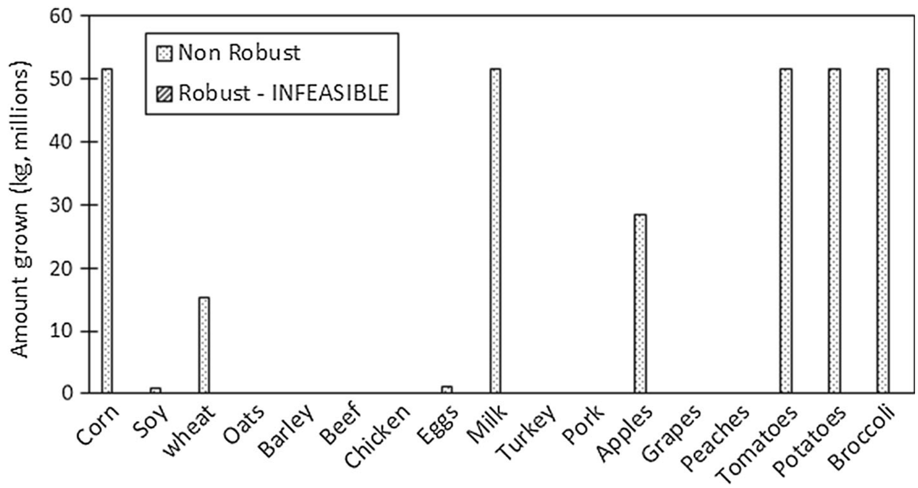


Fig. 3 Optimal production amounts at 17% diversification level

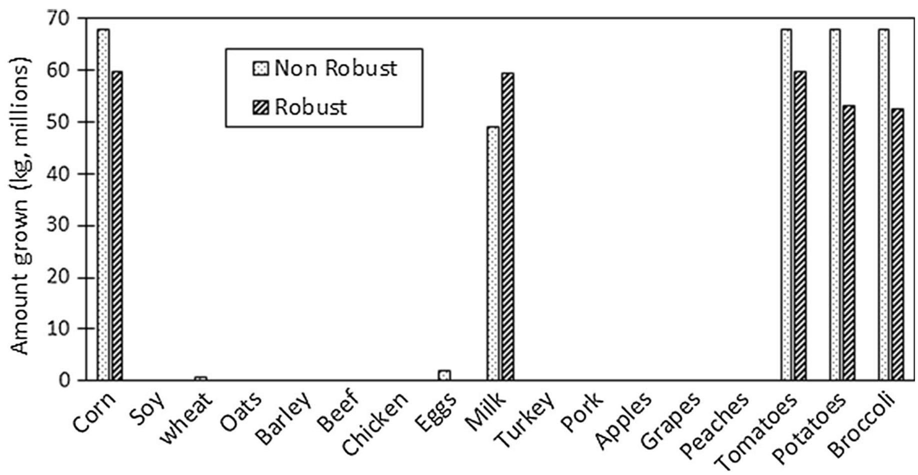


Fig. 4 Optimal production amounts at 21% diversification level

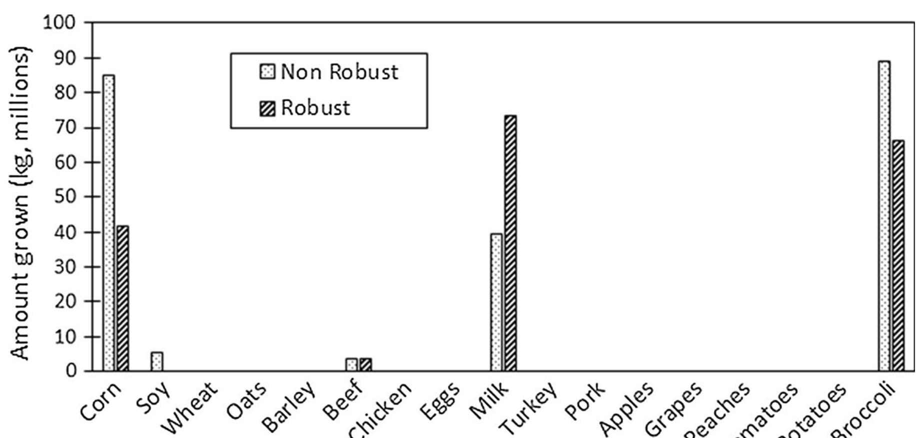


Fig. 5 Optimal production amounts, maximizing beef production, at 40% diversification level

4 Discussion

Our results suggest that Boone County, MO, an area with substantial existing agricultural production and relatively small population, could potentially grow sufficient food within its current cropland and pasture to satisfy the nutritional needs of its inhabitants. When considering average yield levels, these requirements can be satisfied while incorporating a diet that is composed of no more than 17% one food type (by weight). The primary contribution of this research is an extension to include consideration of the 10-year variability in observed yields. Our robust optimization model is able to identify a solution that would satisfy the population's diet requirements, even in the event that yields vary by amounts consistent with those observed over the multi-year period considered, although in this case the optimal diet is more heavily reliant on a few foods, with up to 21% of the total production (by weight) in each of corn, milk and tomatoes. This diet concentrates production in food types that are less subject to yield variability and requires 12% less cropland and 21% more pastureland, based on average yields, than does the non-robust solution. Moreover, the robust solution would be feasible even if yields varied from their average estimates (in this solution, the worst-case allowable variations would be 70, 70, 11, 3 and 2% of the observed maximum variation for corn, milk, tomatoes, potatoes and broccoli, respectively), whereas the non-robust solution would not be feasible in this scenario, as it would require 312 km^2 of cropland. At a diversification level of 21%, the non-robust LP solution was at its constraint limits for three nutrients (calcium, vitamin D and vitamin B12), whereas the robust solution was at its constraint limits for only two nutrients (calcium and vitamin D). A more diverse set of foods would be likely in more realistic representations of local food systems and diets; however, it was not our intention to identify a specific mix of foods, rather to ensure that nutritional needs of a given population could be met. The model presented herein could be expanded to include additional foods.

These findings are generally consistent with previous studies of local food production (Peters et al. 2009, Zumkehr and Campbell 2015), which found considerable potential for local self-sufficiency. Our findings support this conclusion for a region with relatively small population and relatively large existing agricultural land, even when accounting for a more nuanced definition of nutritional requirements and for variability in yields. It is unclear whether these results would extend to a more urban environment (e.g., Chicago), which was deemed capable of self-sufficiency in (Zumkehr and Campbell 2015), given that the production levels for Boone County varied considerably when yield uncertainty was considered. Future work may also include the potential for urban agriculture, permaculture and agroforestry to contribute to food supplies (Clark and Nicholas 2013; Smith et al. 2012). To the extent that alternative agricultural production paradigms can achieve multiplicative effects through coproduction of animals and crops, models similar to those presented here could be extended, through the inclusion of binary-restricted decision variables, to account for these multiplicative effects.

References

Bedoussac L, Journet E-P, Hauggaard-Nielsen H, Naudin C, Corre-Hellou G, Jensen ES, Prieur L, Justes E (2015) Ecological principles underlying the increase of productivity achieved by cereal-grain legume intercrops in organic farming. A review. Agron Sust Dev 35(3):911–935

Ben-Tal A, Ghaoui LE, Nemirovski A (2009) Robust optimization. Princeton University Press, Princeton

Briend A, Darmon N, Ferguson E, Erhardt JG (2003) Linear programming: a mathematical tool for analyzing and optimizing children's diets during the complementary feeding period. J Pediatr Gastroenterol Nutr 36:12–22

Buzby JC, Wells HF, Hyman J (2014) The estimated amount, value, and calories of postharvest food losses at the retail and consumer levels in the United States. United States Department of Agriculture Economic Research Service, Bull Number 121

Clark KH, Nicholas KA (2013) Introducing urban food forestry: a multifunctional approach to increase food security and provide ecosystem services. Landsc Ecol 28:1649–1669

Costello C, Xue X, Howarth RW (2015) Comparison of production-phase environmental impact metrics derived at the farm and national-scale for United States agricultural commodities. Environ Res Lett 10(11):114004

Johnson T, Rossi J, Hendrickson M, Cantrell R, Scott JR (2014) Economic impacts of local food systems in the rural midwest: evidence from Missouri and Nebraska

Low SA, Adalja A, Beaulieu E, Key N, Martinez S, Melton A, Perez A, Ralston K, Stewart H, Suttles S, Vogel S, Jablonski BBR (2015)Trends in U.S. local and regional food systems. AP-068, U.S. Department of Agriculture, Economic Research Service

Mariola MJ (2008) The local industrial complex? Questioning the link between local foods and energy use. Agric Hum Values 25(2):193–196

Martinez S, Hand M, Da Pra M, Pollack S, Ralston K, Smith T, Vogel S, Clark S, Lohr L, Low S, Newman C (2010) Local food systems: concepts, impacts, and issues. ERR 97, U.S. Department of Agriculture, Economic Research Service

Monaco F, Sali G, Mazzocchi C, Corsi S (2016) Optimizing agricultural land use options for complying with food demand: evidences from linear programming in a metropolitan area. Aestimum 68:45–59

Mount PA (2012) Local food, scale and conventionalization: mid-scale farms and the governance of "Local Beef" Chains. Ph.D. thesis, University of Guelph, Guelph, Ontario, Canada. https://atrium2.lib.uoguelph.ca/xmlui/bitstream/handle/10214/3971/UG_Thesis_Mount_2012.pdf. Accessed 13 Oct 2016

National Land Cover Database (NLCD) (2011) http://www.mrlc.gov/nlcd11_data.php. Accessed 13 Oct 2016

Peters CJ, Bills NL, Lembo AJ, Wilkins JL, Fick GW (2009) Mapping potential foodsheds in New York State: a spatial model for evaluating the capacity to localize food production. Renew Agric Food Syst 24(1):72–84

Smith J, Pearce BD, Wolfe MS (2012) A European perspective for developing modern multifunctional agroforestry systems for sustainable intensification. Renew Agric Food Syst 27(4):323–332

United States Census Bureau (2016) http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_14_1YR_S0101&prodType=table. Accessed 13 Oct 2016

United States Department of Agriculture (USDA), Agricultural Research Service (ARS), national nutrient database for standard reference release 28. https://ndb.nal.usda.gov/ndb/search. Accessed 13 Oct 2016

USDA Center for Nutrition Policy and Promotion (CNPP). http://www.cnpp.usda.gov/sites/default/files/myplate_miplato/table3.pdf. Accessed 13 Oct 2016

USDA, Economic Research Service (ERS) (2016) Food Availability (Per Capita) Data System. http://www.ers.usda.gov/data-products/food-availability-(per-capita)-data-system/.aspx. Accessed 13 Oct 2016

USDA, National Agricultural Statistics Service (NASS) http://quickstats.nass.usda.gov/. Accessed 13 Oct 2016

USDA Natural Resources Conservation Service (NRCS) (2009) Balancing your Animals with your Forage: Small Scale Solutions for your Farm. http://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/stelprdb1167344.pdf. Accessed 13 Oct 2016

Weber C, Matthews HS (2008) Food-miles and the relative climate impacts of food choices in the United States. Environ Sci Technol 42(10):3508–3513

Zumkehr A, Campbell JE (2015) The potential for local croplands to meet U.S. food demand. Frontiers. Ecology 13(5):244–248