Spatial Distribution of Greenhouse Commercial Horticulture in Kenya and the Role of Demographic, Infrastructure and Topo-Edaphic Factors

Faith Justus * and Danlin Yu

Department of Earth and Environmental Studies, Montclair State University, 1 Normal Ave, Montclair, NJ 07043, USA; E-Mail: yud@mail.montclair.edu

* Author to whom correspondence should be addressed; E-Mail: justusf2@mail.montclair.edu; Tel.: +1-973-674-0480.

Received: 22 September 2013; in revised form: 13 February 2014 / Accepted: 17 February 2014 / Published: 4 March 2014

Abstract: Greenhouse commercial horticulture in Kenya started more than two decades ago and has evolved to be a significant sector to the national economy. So far no studies have explored the spatial patterns and dynamics of the area under greenhouse cultivation. Google Earth archives alongside data from various portals provided an opportunity to study those farms’ spatial distribution. The roles of selected topo-edaphic, infrastructure and demographies factors that might influence current location within sub-watersheds in central highlands of Kenya are also examined. Results reveal a non-uniform spread with two high clusters; one in the semi-arid sub-watersheds 3AB shared by Kajiado and Machakos districts and the other is in sub-humid sub-watersheds 3BA shared by Kiambu and Nairobi districts. Multivariate linear regression analysis reveals four statistically significant parameters; population density ($p < 0.01$), number of dams ($p < 0.01$), average rainfall ($p < 0.01$) and average slope ($p < 0.05$) in predicting the number of greenhouse farms. Soil attributes are not significantly related with greenhouse farming in this study. Findings indicate that greenhouse commercial horticulture is heterogeneous, and rapidly expanding beyond the central highlands towards marginal semi-arid zones in Kenya. These findings are applicable in policy and decision making processes that aid the horticulture sector’s progress in a sustainable manner.

Keywords: greenhouse; commercial horticulture; spatial patterns; sub-watersheds
1. Introduction

Proliferation of large-scale intensive commercial horticulture, i.e., the production of high value fresh cut flowers, vegetables and fruits has attracted several Sub Saharan African economies in quest for unconventional export commodities [1,2]. The decline of traditional cash crops (coffee, tea and tobacco) due to changes in market demands, low return prices [3,4] and decreasing land resources for plantation farming [5] affected their production in Kenya. The horticulture sub sector has since been the fastest growing industry within the agricultural sector, scoring a yearly average growth of 15% to 20% in the last decade [6]. Approximately 4.5 million people are directly employed in production, processing, and marketing activities, while 3.5 million people benefit indirectly through trade and numerous other activities. Several reasons have promoted horticulture growth. These include a high demand for fresh produce at the international European markets [7], better trade terms and liberalization of economies, need for diversification [8,9], growing health and dietary awareness linked to fresh produce [10] and the desire of having year round availability of fresh produce. Other countries such as South Africa [11] and Tanzania [12] have successfully exploited the niche and often realized increased foreign income, creation of jobs for skilled and semi-skilled workers, and improved livelihoods for the urban and rural workers.

Success of the Kenyan horticultural sector is well documented [1,12]. Many studies attribute the continued progress to the strategic location of the country on the equator, which endows Kenya with good climatic conditions [9,13]. Additionally, availability of cheap labor [14], widened market access, frequent and reliable freights from Jomo Kenyatta International Airport and minimal governmental interference have also contributed to the success [15].

Recent studies show a general trend towards fewer and larger horticultural greenhouse growers [16], as evident in Kenya. Greenhouse horticulture production is intensive, occurring under a controlled environment which enables a year-round all season production. Crops are planted directly in native sterilized soils, under greenhouse field conditions, because of several reasons: (1) greenhouses offer a controlled environment against unexpected weather occurrences, e.g., frost, strong winds, hail, torrential rains, which can severely affect crops; (2) they also enable a timely planting, monitoring and management of crops, by controlling for humidity and carbon dioxide levels, which enable flower formation by the projected harvest dates; and (3), they offer protection from pests and diseases which are common in tropical climates. The sterilization of soil kills pathogens and nematodes which are a nuisance to crops like roses. Roses are commonly grown in Kenya on raised beds (~60 cm) (Figure A1) because the rose bushes (from where the flower stem sprouts), can last ~7 years under good management. The raised beds enhance deep root penetration and stability. However, other flower crops like Carnations, Static, Alstroemeria, Lilies, Hypericum and Lisianthus (also commonly grown in Kenya) [17], do not require raised beds, but are directly planted in the ground under greenhouses (Figure A2). Large amounts of land resources, water, labor, machinery, agrochemicals and technological skills are utilized through complex strategic planning. Small scale farmers play a minimal role, often contracted by larger companies.

The success of the industry with such an intensive nature of production has enormous impacts on the environment [18,19]. Yet such impact is not equally distributed since production is concentrated in specific sub-watersheds. Watersheds provide resources that are shared between a growing human
population, expanding cities, wildlife, plant species, and microorganisms. A threshold of resource use can be quickly reached if unsustainable means are applied in exploiting resources for commercial production needs. Furthermore, particular agricultural management practices can affect soil structure, composition and microorganism communities [20]. Fungal communities are affected by excessive soil fertilization, tillage [21], and soil pretreatments with effects observed on spores and mycelia densities in temperate and tropical agro-ecosystems [22]. The magnitude of production horticultural activities is also linked to declining water quality [23], due to nitrogen and phosphorus runoff, pesticides, soil enhancements and untreated pit latrine waste from settlements.

Failure to account for ecosystems’ goods and services (the environmental externalities) used in production when calculating returns is a large loop-hole contributing to resource over-use [24]. Until recently, environmental costs incurred in cleaning polluted waters [25] and soils through remediation were minimally recognized as items that needed to be included in production cost estimation. It becomes clear that watersheds experiencing intense and increased production activities are likely to deteriorate over years, and bouncing back to their productive status is fairly unlikely.

Currently, there is no national horticultural policy to guide the growth and sustainability of the horticulture sector [6]. This failure has resulted in overcrowding of farms in specific regions of central Kenya, which in turn necessitates a need to explore current spatial patterns and production dynamics, in order to generate information useful in forming a sustainable export culture. The current mainstream description of main horticulture production areas in Kenya is presented in a very general manner. It is stated that, “The main production areas are around Lake Naivasha, Mt. Kenya, Nairobi, Thika, Kiambu, Athi River, Kitale, Nakuru, Kericho, Nyandarua, Trans Nzoia, Uasin Gichu and Eastern Kenya” [17]. Such a description lacks precise geolocation that can be rectified by a distribution map based on point locations of actual locality of greenhouses. Our preliminary cartographic endeavor suggests greenhouse production is concentrated in a few sub watersheds within these broader administrative districts.

This work is the first in a series of studies seeking to establish the spatial extent of greenhouse horticulture production in Kenya, understand important factors influencing choice of location, and evaluate the environmental impacts of increased production activities on sub-watersheds.

The study’s objectives were (i) to determine current spatial distribution of green-housed commercial horticulture production and derive their acreage under cultivation within sub-watersheds in the central highland of Kenya; and (ii) to evaluate significance of topo-edaphic factors (i.e., soil pH, cation exchange capacity (cec), average bedrock, exchangeable sodium (exNa), exchangeable potassium(exK), average slope, rainfall and river density); infrastructure (road density and dams); and demography (population density) as factors determining location of farms.

2. Materials and Methods

2.1. Study Area

The study area covered 88 sub-watersheds in central highland of Kenya, confined within: 34°55’51.88”E, 0°55’10.51”N (upper-left); 38°07’09.91”E, 2°23’03.73”S (lower-right) and an area of 81,607.26 km² (Figure 1a). The study region was found to have various greenhouse horticultural
activities during the study period (2000–2011). Four lakes are found in the area (i.e., Lake Naivasha, Nakuru, Baringo and Elementaita) (Figure 1a). The region has six diverse agro-ecological zones that show a varied physical environment of an almost equatorial type of climate in upper-lands and semi-arid and arid environments in the lowlands [26]. The upper-lands occur in the upper northwest region of the study area and are generally humid to sub humid, often presenting thick forest vegetation. The lowlands show strong gradients in terms of average rainfall, temperature, and vegetation characteristics. Due to the mixed status, a variety of conditions exist that facilitate diverse farming. These include subsistence to export oriented commercial farms, and extensive large scale farms to very intensive small scale farms and livestock keeping [27].

Rainfall is highly variable in the study area, occurring in two seasons, April–May (long rains) and October–November (short rains). Around the Aberdare Mountains the average annual rainfall is about 1350 mm/year. However, this reduces substantially to ~600 mm/year advancing near Lake Naivasha. Another interesting feature to the east of the area is Mt. Kenya which stands at 5199 m above sea level. The mountain causes a strong shadow effect that highly influences climate conditions of the adjoining plateau land to more of a semi-arid to arid region, with rainfall of ~450 to 750 mm/year [28]. Due to low average annual rainfall, vegetation to the eastern and southern (760 mm/year) regions is dominated by woody vegetation of Acacia and Themeda grass cover that mainly occur in savanna environments [29].

Figure 1. Study area. (a) Map of Kenya (right), and inset (left) is the study area showing sub-watersheds in the central highlands considered in the study. (b) Major towns, rivers, lakes and mountain features—Mt. Kenya and the Aberdare ranges; (c) A Google Earth image of a greenhouse and a dam.
Temperature varies dramatically with altitude, attaining mean maximum and minimum temperatures of ~25 °C and 15 °C, respectively. Soils are predominantly fertile, deep (>160 cm) rich in clay and well drained, derived from mixed volcanic rocks. However, severe land degradation is widespread due to topsoil erosion attributed to continuous cultivation [30]. The upper zones are densely populated while drier areas exhibit sparse population.
2.2. Data Acquisition and Processing

This study attempts to predict spatial extent of greenhouse farms in sub-watersheds using eleven potential explanatory variables that are grouped to include topo-edaphic variables of soil pH, cation exchange capacity (cec), average bedrock, exchangeable sodium (exNa), exchangeable potassium (exK), average slope, rainfall and river density; infrastructure (road density and dams); and demography (population density). The crops are grown directly in the soil, under the greenhouses, and an incorporation of soil parameters in the model was considered meaningful. We describe data acquisition and preparation for each variable.

The sub-watershed data layer was created by the World Research Institute (WRI) from the 1992 Kenya National Water Master Plan, and is publically available (http://www.wri.org [31]). Each sub watersheds is assigned a unique code composed of a number and two letters, e.g., 5AB. The number represents one of the five river basins in Kenya, i.e., Lake Victoria (1); Rift valley and inland lakes (2); Athi River and coast (3); Tana River (4) and Ewaso Ng’iro (5) in which the sub-watershed resides. The first letter (in this case A) represents the sub catchment, while the second letter (in this case B) represents the major river branch in the particular sub-catchment. Given that a sub-catchment may have more than one major river branch, the second letters can continue progressively to as many as there are major river branches in the sub basin, for example, 5AA, 5AB, 5AC, 5AD, 4BA, 4BB, etc.

To determine current spatial distribution of green-housed commercial horticulture production and derive their area under cultivation within sub-watersheds, the geographic coordinates of the centroid point locations with greenhouses were identified, alongside the corresponding greenhouse images from Google Earth®. Google Earth uses the WGS84 (World Geodetic System 1984) and a terrain data derived from the Shuttle Radar Topographical Mission (3-arcsecond, ~90 m resolution at the equator) with a nominal accuracy of between 6 m and 8 m [32,33]. The increase in the use of Google Earth imageries in deriving area measurement, e.g., mapping the area under urban agriculture [34]; deriving the area of stone walled enclosures [35], and development of microscale meteorology models [36] may be an indication of its acceptance in the geospatial community as a means of making areal measurements with an acceptable level of accuracy relative to the scale of the project.

Both historical (2000–2003) and current imageries (2010–2011) were used (Figure 1c). The background imagery for Google Earth is a Landsat mosaic of 30 m resolution satellite images. However, where available, images of higher spatial resolution (1 m to 4 m) from commercial satellite programs such as DigitalGlobe and GeoEye are added and can be viewed when zoomed in. The clarity of individual images (Figure 1c) allowed for an accurate determination of greenhouse structure outlines and derivation of polygon area, by zooming in to a satisfactorily extent before taking measurements. From the images, the total area around the individual greenhouses at each identified farm locations was calculated. Historical imagery enabled identification of previous area under greenhouse production (if any) existed in that point. It also helped to establish probable year when the farm started. Images were carefully labeled with the dates of acquisition.

ImageJ, an image processing and analysis software developed by the United States National Institute of Health (NIH) [37] was used to derive the area under production from the saved images. ImageJ is widely used across disciplines due to its flexibility, and the ability to automate tasks using simple macros and plug in extensions. The measurement tool was calibrated before area calculation
was performed. Calibrating the image indicates to the software what an image pixel represents in real world in terms of size and distance, to enable conversion of image pixels to distance measurements [38]. Features of known distance are used in calibration, and in our study, a known distance of 100 m from a greenhouse was used. The area selection tool was then used to determine polygon area around individual greenhouses in sq. km.

Derived data on the area under greenhouses were then entered as attributes to the farms in a sub-watershed GIS layer. Sub-watersheds were used as the basic study units since this is a good balance between data collection and production of meaningful results and feedback, which would best serve watershed management plans in the future.

To evaluate the significance of topo-edaphic, infrastructure and demographic factors determining location of farms, various data layers of exploratory variables that potentially influence farm locations for greenhouse production and hence location suitability were obtained for each sub-watershed. The variables were based on literature and expert knowledge of the dominant factors influencing the choice of location [30,32–42].

Data layers of administrative boundaries (Districts), river and stream network, roads network, population and agro-ecological zones were available from [43].

Generally, the soil physical and chemical characteristics are important in determining its productivity and ability to support crop growth and development.

The soil parameters were derived from an Arc Info coverage of soil physical and chemical properties of Kenyan soils carried out by the Kenya soil survey in 1982 [44], and revised in 1997. It is publically available as a data layer from International Livestock Research Institute [43]. These data are the latest available at such scale and detail, and were recently used to characterize subsistence farming by different households in central Kenya [27].

Soil pH is a measure of the soil acidity or alkalinity as measured in a soil “water” solution, and considered a proxy of the active pH that affects plant growth [27]. The average cation exchange capacity (cec) measured in me/100g of the soil was used to give an estimate of the soil’s ability to attract, retain, and exchange cationic elements. The exchangeable potassium (exK) me/100 g, was considered important since potassium controls the stomata movements facilitating a plant’s efficiency in water use. It assists in recycling of nutrients to feed roots, leaves and fruits thereby promoting plant life [45]. It is also a key element besides nitrogen and phosphorus. The amount of exchangeable sodium (exNa), measured in me/100g, is an important factor determining a soil’s suitability for supporting plants [45] as it strongly influences water infiltration and soil aeration. Furthermore, soils of arid and semi-arid areas contain large quantities of sodium (Na), and irrigation agriculture in such areas is often a source of soluble salts even if excellent quality water is used [40]. Excesses of exNa (>2.5%) becomes toxic to the plant, and affects plant growth since it adversely alters the physical and chemical conditions of the soil (i.e., soil permeability, dispersion [40] and water permeability). Besides, sodium competes with calcium, magnesium, and potassium for uptake by plant roots prompting deficiencies of other cations.

The depth of average bedrock influences drainage of the site, water flow pattern and other hydro-geological functions [46] that might influence the decision to establish a greenhouse. In addition, it can offer a solid foundation for greenhouse construction.
The average slope was derived from a 250 m Digital Elevation Model (DEM) grid [31]. Slope influence drainage of an area [47,48] and the applicability of heavy farm equipment mainly utilized in large scale farming. It is an important parameter considered in installation of drip irrigation systems in greenhouses. While the use of a coarse-resolution DEM may affect the obtained slope values, it was a readily available dataset and compatible with our model. Furthermore, by deriving the average slope, the impact of likely high frequency noise introduced by the coarse DEM is reduced.

Average annual rainfall distribution (in millimeters) for sub watersheds was derived from data compiled by the Japanese International Co-operation Agency (JICA) under the National Water Master Plan, Kenya [31]. This data was applied in the model because rainfall is a major source of river and groundwater recharge, which occurs at the higher altitudes of Mt. Kenya and the Aberdare regions. Reduced rainfall affects river flow, which may directly impact irrigation and the area under production. River density, commonly known as drainage (total length of all the streams and rivers in a drainage basin divided by the total drainage area) was derived using the streams and river network data layer. It is a measure of how well a watershed is drained by rivers and stream channels.

Road density and number of dams were considered as two infrastructure exploratory variables. Road density was estimated for each sub-watershed by dividing the sum of all roads (km) in sub-watersheds by the area of sub-watersheds (sq. km). The number of dams within 500 m to identified farms was counted from images and entered as an attribute to the farms in the same GIS layer.

The census data follows administrative boundaries (i.e., districts and sub-locations). Sub-locations are administrative units ranked lower in hierarchy than districts and locations. The census population data in Kenya were obtained at this level, and aggregated to the district level. The 2009 census data obtained from Kenya National Bureau of Statistics [49] were joined to the district layer and population numbers for each sub-watershed extracted using ArcGIS®. Population density, representing a demographic exploratory variable, was derived by dividing the total population per sub-watershed by area of the sub-watershed.

Following the processing, the 11 exploratory variables were then used in a multivariate linear regression model to predict spatial location of observed greenhouse farms. The counted number of greenhouses and number of dams in the study area were logarithmically transformed to base ten. Such transformation is recommended to enable data normality [50] since greenhouses and dams were not evenly observed across all sub-watersheds. This was to enable applicability of ANOVA analysis, before which the Shapiro–Wilk test was done to check for data normality after the transformation. The test was statistically significant ($p < 0.01$) attesting data normality. A stepwise linear regression model [51] was used to relate the occurrence of greenhouses in the different sub-watersheds to all 11 potential explanatory variables. This approach was chosen since it assesses the contribution of individual variables to overall model fitness, and assists in choosing to exclude or include variables accordingly.

3. Results

3.1. Spatial Distribution and Area under Greenhouse Commercial Horticulture

The distribution of farms varied greatly in the sub-watersheds (Figure 2a). Our findings indicate that 24 out of the 84 sub-watersheds had greenhouse horticulture farms during the study period.
Clustering, defined as geographic concentrations of interconnected companies and institutions in a particular field [52] was observed in the Athi river area. This was in sub-watershed 3AB shared between Kajiado and Machakos districts, which have semi arid type of climate. An increased number of farms were also observed in sub-watershed 3BA that is shared by Kiambu and Nairobi districts, which are considered sub-humid to semi-arid areas. Other notable but moderately clustered sub-watersheds were 2GD which encompasses Lake Naivasha and Nakuru region, 3BB which covers Kiambu district, and 5BE that is shared by Meru Central and Laikipia Districts.

Figure 2. (a) Spatial distribution of greenhouse commercial horticulture in the central highlands of Kenya within sub-watersheds in 2000–2011; (b) Within districts.
Further examination of the spatial spread across the districts indicates that Kiambu, Nakuru, Thika and Kajiado, had more than 15 large scale commercial greenhouse farms each, while Laikipia, Meru Central and Kericho districts had 5 farms or less, individually (Figure 2b). Our observation is supported by [53] who report that production for export is concentrated in about two dozen, large-scale farms, accounting for 75% of the industry and 10%–15% by contracted small scale farmers.

Table 1. Total land area under greenhouse cultivation, and number of greenhouses in 2000–2003 and 2010–2011 in the study area.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>30 **</td>
<td>4 *</td>
</tr>
<tr>
<td>0.01–0.02</td>
<td>10</td>
<td>16</td>
</tr>
<tr>
<td>0.02–0.04</td>
<td>13</td>
<td>21</td>
</tr>
<tr>
<td>0.04–0.06</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>0.06–0.08</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td>0.08–0.10</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>0.10–0.12</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>0.12–0.14</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>0.14–0.16</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>0.16–0.18</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>0.18–0.20</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>0.20–0.22</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>&gt;0.4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Total Farms</td>
<td>66</td>
<td>97</td>
</tr>
</tbody>
</table>

* In 2011, four farms that initially had greenhouses but were currently not greenhouses, or were converted to other uses and so are assigned zero sq. km. ** Farms not existing in 2000–2003 but established in 2010–2011 are assigned a zero area under greenhouses for the 2000–2003. This reduces the number of current observed farms to 97.

Total land area and number of farms under greenhouse cultivation increased in the period 2000–2011 (Table 1). In 2000–2003, 66 farms covering and estimated area of 3.76 sq. km were identified. This number increased in 2010–2011, to 97 greenhouse farms covering 6.83 sq. km, which was a 3.07 sq. km. increase from previous period (2000–2003). An additional 31 greenhouses were identified in 2010–2011, which greatly increased the area under production between the periods. This observation agrees with a noted increase in real Gross Domestic Product annual growth from ~0.6% in 2002 to 6% in 2006 [53].

The majority of the greenhouse farms in 2010–2011 were approximately 0.01–0.12 sq. km (83.5%). 15 greenhouses had 0.12–0.22 sq. km (15.4%), and 1 farm had >0.40 sq. km (0.01%) under production. A notable increase in number of greenhouses occurred in 2010–2011 (31), which corresponded to an increase in area under production. What, then, are the possible conditions that drive such spatial distribution? Our multivariate regression might shed light in this regard.
3.2. Significance of Topo-Edaphic, Demographic and Infrastructure Factors in Predicting Farm Locations

The multivariate linear stepwise regression model results (Table 2) show that not all 11 variables were significant predictors of greenhouse locations. In fact, the final model identifies only number of dams, population density (2009), average precipitation and average slope are significantly related with the occurrence of greenhouse farms. The final model has an $r^2 = 0.88$, indicating that 88% variation in the observed spatial location of greenhouse farms in sub-watersheds was explained by these four variables. The statistically significant F value suggests a linear relationship between the transformed occurrence of greenhouse farms and the four explanatory variables. The other seven variables are not found to be significantly related with observed number of greenhouse farms.

Table 2. Multivariate linear step-wise regression model results. Coefficient of determination $r^2 = 0.88$, ($p < 0.01$, $t = 154.69$) showing significant variables of dams ($p < 0.01$), population density ($p < 0.01$), average rainfall ($p < 0.01$) and average slope ($p < 0.05$), in predicting greenhouse farms.

<table>
<thead>
<tr>
<th>Model</th>
<th>Adjusted R-Square</th>
<th>df</th>
<th>F</th>
<th>sig.</th>
<th>Model</th>
<th>B</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.84</td>
<td>83</td>
<td>449.82</td>
<td>0.00</td>
<td>1</td>
<td>0.127</td>
<td>2.826</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dams</td>
<td>0.642</td>
<td>21.209</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>0.86</td>
<td>83</td>
<td>449.82</td>
<td>0.00</td>
<td>2</td>
<td>0.013</td>
<td>0.249</td>
<td>0.804</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dams</td>
<td>0.613</td>
<td>20.899</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>POPden09</td>
<td>0.001</td>
<td>3.664</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>0.88</td>
<td>83</td>
<td>194.9</td>
<td>0.00</td>
<td>3</td>
<td>0.37</td>
<td>2.711</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dams</td>
<td>0.598</td>
<td>20.862</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>POPden09</td>
<td>0.001</td>
<td>4.537</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AvgRainfall</td>
<td>0.00</td>
<td>−2.813</td>
<td>0.006</td>
</tr>
<tr>
<td>4</td>
<td>0.88</td>
<td>83</td>
<td>154.69</td>
<td>0.00</td>
<td>4</td>
<td>0.535</td>
<td>3.513</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dams</td>
<td>0.595</td>
<td>21.254</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>POPden09</td>
<td>0.001</td>
<td>4.112</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AvgRainfall</td>
<td>0.00</td>
<td>−2.713</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AvgSlope</td>
<td>−0.059</td>
<td>−2.232</td>
<td>0.028</td>
</tr>
</tbody>
</table>

Our finding indicates the many soil parameters were not important factors influencing occurrence of greenhouse farms, which at the first glance, might be quite unanticipated. Considering the primary purpose of greenhouse farms, which is to seek the highest possible profit, such finding, however, does not seem to be unreasonable. We deduce that the variation of soil condition in our study area is not great enough to influence farm establishers for their locational decision. This is especially true due to increased use of soil amendments and soilless technologies in horticulture greenhouse farming. Instead, other factors that are more related with economic gains, such as population density, availability of water and slope factors, stand out as the most significant and alarming to decision makers who need to understand the horticulture practice is not sustainable in the long run.
4. Discussion

4.1. Spatial Distribution of Greenhouse Commercial Horticulture

Based on our analysis, the spatial spread of greenhouse farms varied greatly from 2000 to 2011, occurring primarily in 24 sub-watersheds out of the 84 considered in central highlands. Production is heterogeneous across the study area, creating hot spots which have successfully sustained horticulture production in Kenya, at least in the last decade. Dense clusters were found in the Athi River (3AB), Kiambu and outskirts of Nairobi (3AA), alongside moderate clusters that surround Lake Naivasha and Nakuru region (3BB), Meru Central and Laikipia districts (5BE) (Figure 2a,b). A minimal number of farms were observed in other sub-watersheds. Our finding is important because it helps clarify the standing assumption that production of horticultural commodities in Kenya occurs in many locations and is favored by good climatic conditions. We found fewer districts (Figure 2b) than those provided by Kenya Flower Council as regions of commercial horticulture farming, particularly cut flowers. The highly heterogeneous spread of greenhouses (Figure 2a) in sub-watersheds is likely related to availability of capital environmental resources that are employed in production of fresh horticultural produce. Increased demand for fresh produce at the international and domestic markets requires an all year round production which is met provided continuous resource availability.

One key characteristic of horticulture farming is intense water usage [18]. On the contrary, since water resources are unevenly distributed across and within sub-watersheds, we observe a limited number of farms in relatively arid areas and increased clustering in more humid regions where water is readily available (i.e., around Lake Naivasha). Greenhouse clustering can be an efficient use of land resources particularly in areas with high urbanization pressure [54]. From an environmental resource management perspective, such skewed patterns may result in imbalance in resource sharing, over-exploitation and faster degradation of sub-watersheds. This necessitates spatial planning and coordination of policy levels across institutions and various stakeholders. Our analysis is highly suggestive of a strong role played by combined production factors in explaining observed greenhouse concentration in specific sub-watersheds. Sub-watersheds with suitable soil characteristics but lacking in significant primary factors are possibly unsuitable as greenhouse sites. Presence of a diverse climate and vegetation gradients across sub-watersheds [28] further strengthens our observation that different factors may be influencing location of choice. For example, observed increases in greenhouses in Athi River (3AB), which is a semi-arid, low rainfall area and previously sparsely populated is likely due to the proximity of Jomo Kenyatta International Airport that readily provides a means of transport.

4.2. Changes in Area under Greenhouse Cultivation

A remarkable increase in number of greenhouses, and a corresponding increase in area under production (Table 1) were identified. Study analysis shows growth in production area from ~3.76 sq km to ~6.83 sq. km between 2000–2003 and 2010–2011 respectively. This observation agrees with a noted increase in real Gross Domestic Product annual growth estimated at 0.6% in 2002 and to 6% in 2006 attributed to growth of the horticulture sector [53]. Our calculated area shows the majority of all counted greenhouses (83%) occupied an area between 0.01–0.12 sq. km, while 15.4% had between 0.12–0.22 sq. km and one farm had more than 0.4 sq. km under greenhouses. In recent years,
studies show that greenhouses are generally occupying more space [55]. For instance in Canada, Spain, Great Britain and the Netherlands, new greenhouses occupy between 0.03 sq. km and 0.3 sq. km [56]. Meanwhile, a range is observed from 0.001 to 0.1 sq. km in Flanders [57].

Greenhouses provide a controlled environment for crop production, and application of technological advances. A couple of reasons may explain the observed growth dynamics. In the years following the decline of coffee prices and revenues [58], the majority of farmers sought alternative high value export commodities. Fresh produce commodities such as cut flowers, vegetables, fruits, and nuts provided this avenue. Diversification to high value produce was further strengthened by better trade terms such as the African Growth and Opportunity Agreement, increased produce demand at international market, and widened market access. Moreover, minimal government intervention in regulating the horticulture sector, unlike traditional cash crops, encouraged farmers to consider broader intensified production. The use of greenhouses to control for climate, pests invasion and weather effects on produce further enabled this type of production. From our analysis, greenhouses occupied as few as 0.01 sq. km, to as many as 0.22 to 0.40 sq. km., and a majority had between 0.01 and 0.12 sq. km. The magnitude of area may seem small but considering the intensity of production per unit area, the value is interesting. Our finding can relate to a study by [55] indicating a growth in greenhouse farms occupying 0.03 sq. km. It is rational to think that the ability to expand units of production highly depends on the investor's potential, more so because agricultural intensification requires labor and capital to enable the increased inputs necessary to raise the value of output per hectare [59]. Generally, the observed changes in area under production seem necessary and have supported horticulture sector growth atop traditional export commodities, and the economy [60,61]. However, rapid expansion of the sector has implications for resource utilization and sustainable management of watersheds resources, necessitating further evaluation of environmental impacts of increased production.

4.3. Role of Different Factors in Determining Spatial Location of Greenhouse Farms

4.3.1. Significant Predictors

Our multivariate regression analysis shows four significant factors related to location of greenhouses. These are average slope \((p < 0.05, t = -2.23)\), average rainfall \((p < 0.01, t = -2.713)\), number of dams \((p < 0.01, t = 21.25)\) and population density \((p < 0.01, t = 4.11)\) (Table 2). Our results show an average slope of 2.9% across the study area, with steepest slope at 6.8%, and lowest point at 0.77%.

The average slope considered in current study as a topography variable has significant effect on crop yield [48], and can affect mobility [62] and utilization of farm machinery in large scale farm operations. An increase in slope is shown to negatively affect corn yield reducing it by 0.79 bu./acre for each 1% increase in slope gradient [48]. Their work supports our results, which show a significant negative relationship between the average slope and greenhouse farms.

Similar claims are expounded by [63] who indicate that farmers recognize importance of slope in choice of farmland. The negative directionality of effect indicates decreasing number of greenhouses with an increase in average slope, implying that less step gradients are more favorable since they allow ease in use of farm machinery, while sustaining crop yield. A study carried out in the Gikuuri catchment...
in central Kenya [64] indicated flat to gentle slope (2%–15%) in the region, described by [65] as having slight hindrance to the use of heavy farm machinery. Large scale commercial horticulture chiefly use heavy machinery whose ability to work can be limited on steep slopes. This consequently reduces efficiency of mechanization while increasing cost of farm operations. In addition, the level of difficulty and cost in erecting greenhouses and fitting drip line systems is likely to increase at steeper grounds [48].

Results also show a significant negative relationship between numbers of greenhouse farms with average annual rainfall (Table 2). The majority of greenhouses are observed within zones receiving ~950 mm/year. Rainfall as a primary source of irrigation water [66] characterizes the long term quantity of water available for hydrological and agricultural purposes [67]. Study findings show an average increase in annual rainfall towards humid agro-ecological zones as the number of greenhouses declined. While rainfall availability may influence choice of a location for greenhouse farming, other climate factors, e.g., temperature and vegetation cover that change along rainfall gradient may hinder preference of a location for greenhouse farms. This relates to observed increase in area and number of farms in sub-humid to semi-arid and arid regions of the study area. More clustering is observed in semi-arid to sub-humid regions, where rainfall is moderate. This has implications for dry land horticultural farming particularly since rainfall variability in Kenya is shown to relate to ongoing drought and famine. Variability in rainfall is perpetuated by a changing climate, and shifting rain seasons as a result of El Nino North Atlantic Oscillation activity [68]. This may limit future spread and area under greenhouse production to specific areas creating an increased demand for water. Increasing rainfall variability may have an effect on the degree of water abstraction for farming purposes.

Based on our analysis, the number of dams is shown to be significantly (p < 0.01) related with greenhouse farming. Studies show increases in water abstraction via dams as farmers try to reduce shocks due to rainfall decline. Increased water abstraction for irrigation by large-scale commercial horticulture in Laikipia (5BE) was reported in [69]. This could elucidate findings of increases in area under production as facilitated by availability of dam water. Presence of dams can increase area under irrigation and sites of farming while enhancing productivity by increasing multi-cropping and the cultivation of water-intensive cash crops [70]. A report by the International Commission on Irrigation and Drainage [71] finds that dam construction significantly increases agricultural production and yield besides enabling farmers to substitute towards water-intensive crops. However, irrigated marginal lands in Spain are largely degrading due to increased extensive agriculture and population expansion is noted in [72]. We identify an urgent need to seek alternative sustainable sources of water, and investigations on water harvesting technologies that can benefit the sector to avoid resource exhaustion.

Population density was statistically significant (p < 0.01) (Table 2.) in predicting greenhouse farms in our study area. Our findings agree with [73] highlighting that commercial horticulture is labor intensive, often employing masses of workers. Currently, the sector employs close to 4 million people directly [61] in the processes of production, produce processing e.g., cleaning, sorting, grading, labeling, packing and post-harvest tasks, which require substantial skilled and semi-skilled labor. This is met through seasonal hire, contract or fulltime jobs afforded by the farms. It explains why population density as a production factor is highly significant in horticulture, and particularly this region of the world where advances in technological applications are still developing. Furthermore, scarcity of labor in commercial horticulture can be a limiting factor to production [73]. Substantial employment afforded by horticulture attracts job-seekers, prompting growth of unplanned settlements. Neighboring
towns expand [74] and businesses grow to meet public demand for basic household items and services. However, rampant increase in population has implications for resource use, often resulting in conflict between the stakeholders [8,18]. Land fragmentation to build rental shelter structures is noted as a current common scenario that results in declining land productivity and increased discharge of human waste into river system affecting water quality [22]. We identify a strong need for long term sustainable solutions to control overcrowding in sub-watersheds experiencing increased production.

4.3.2. Non Significant Factors: Soil Characteristics, River Density, Road Density and Depth of Bedrock

We find that neither soil parameters, river density, road density nor the average depth of the bedrock was statistically significant in predicting spatial extent of greenhouse farms. However, from our knowledge and reviewed literature, these variables can somewhat influence site selection for agricultural activities. Here below, we highlight the importance of each factor, speculating why our analysis results show non-significant relationships.

Soil characteristics considered at the sub-watershed scale offer a convenient means to organize field measurements, enabling an assessment of the role played by each in predicting greenhouse farming. Soil attributes, for example soil moisture, are highly variable in time and particularly at the watershed scale [75]. Even so, obtaining continuous field data to assess a vast area is challenging. Our approach used a static soil data layer [43], found useful to gain a general idea and baseline understanding of importance of soil parameters in predicting greenhouse farming. Our results show minimal non-significant relationship between observed greenhouses and considered soil parameters. Despite our finding, the relative importance of soil characteristics cannot be overlooked. Studies by [30,64] indicate that soils in the study area have formed as a result of volcanic activity of the Rift Valley. They are fertile, deep, well drained clays that support diverse agricultural activities. We attempt to explain our finding based on available literature and knowledge of horticulture farming. Increasing use of superior soilless media [75], e.g., peat, moss, that has better physical properties compared to soil could lend ordinary soil parameters less influential in siting greenhouses. Soilless media is preferred because it is free of disease, pest and weeds contamination, is relatively inexpensive and environmentally friendly providing an overall less cost of production. It also allows easy adjustment of plant nutrients, depending on crop needs. Nevertheless, not all farmers use soilless substrate. Increased availability of fertilizers and soil dressers provide cheaper alternatives for investors to manipulate soils elements to desirable status irrespective of soil status of a location. Farmers view soil fertility as a dynamic process that integrates a soil’s chemical and physical characteristics, agricultural necessities, and factors in the surrounding environment [76]. More importantly, farmers are influential in the process of increasing soil fertility or degrading soil [77].

Road infrastructure enhances a region’s suitability for not only horticulture but also numerous other commercial activities. Fresh horticultural products are highly perishable requiring readily available means of transportation to the market. In the absence of a good road network post-harvest losses [78] can occur which constrain production and inability of investors to recover investment costs. Even though large scale companies have refrigerated trucks to transport produce, the costs attached to such transportation can be efficiently reduced in the presence of a good road network. Farm proximity to
roads network could offer faster access to the airport. However, growing concern indicates that traffic pollution [79] poses a threat of produce contamination, particularly on farms clustered in proximity to roads [80]. These may explain the poor ability of road density in predicting spatial spread of greenhouses. It is not farfetched to reason that in effort to avoid such risks farms choose to concentrate in interior rural areas where a dense road network is absent. Farms in the interior areas of sub-watersheds face a challenge in transporting commodities to the airport for international markets. Night transportation with less traffic congestion seems a convenient alternative. This spells a need to increase road amenities in productive areas, but in a sustainable way. A note of concern is that our approach to use road density as a key variable could be adjusted to consider distance of farms from main roads or road quality as the variable influencing location choice. The role of road infrastructure in fueling greenhouse horticulture must be established in order to improve our understanding of spatial spread of horticulture and its impacts on environmental resources.

Our assumption that regions with denser river and streams network would likely attract more greenhouses was not supported by our findings. Results show a weak non-significant prediction, a probable indication that presence of dense network of rivers, may not necessarily assure quantity and quality of irrigation water, or its reliability, which is desirable in continuous intensive greenhouse farming.

4.3.3. Suitability of Google Earth and Downloaded Data

Google Earth provides a powerful tool for viewing global imagery with integrated GIS data and well organized tools, maps and graphics, e.g., borders, place labels, roads, ability to zoom in/out, viewing historical archived data among other tools [35,36]. In our assessment, its use as a tool to generate acreage under greenhouses in Kenya is especially suitable for studies in areas with limited access to aerial imagery. To the best of our judgment, the background Landsat imagery of medium spatial resolution in Google Earth, alongside other images from DigitalGlobe and GeoEye satellites with minimal cloud coverage enabled candid spatial assessment of trends in horticultural greenhouses in Kenya. The clarity of greenhouses, as displayed in Google Earth allowed precise area measurement, around the features of interest.

On the contrary, available archived Google Earth data for the region were limited to 2000–2003. Images of greenhouses, within this period were therefore grouped as the starting period. While greenhouse horticulture in Kenya dates back to the 1990s, Google Earth, maybe limited in assessing trends in area under greenhouses or the number of farms during that time. Other archives of remote or aerial data would be required. Interestingly, the soaring number of Google Earth data usage in scientific work, and across disciplines, indicates its growing applicability.

The soil dataset used in our study was recently used by [27]. This evidence prompted its use in our model, and found it of suitable representation of the general soil characteristics in the area. Furthermore, obtaining extensive and detailed soil data is difficult, because recent soil studies are localized, and the data generated are limited to specific local scales of sampling that answers specific research questions. There is need for current, broader scale soil datasets that can be used by the wider research community.
The use of administrative boundaries, the census data (2009), and the sub-watershed data layers in a GIS environment enabled the extraction of population information that was employed in the regression model.

Other data layers of rivers and streams, rainfall, the road network, and the DEM have been compiled by the International Livestock Research Institute group, and are available in public domain. These data were considered suitable for the purpose of our study, and have been used in numerous studies and regional reports.

Though the coarse resolution 250 m DEM [31] used in the study could affect the accuracy of derived slope values, we assume that measures taken, e.g., averaging the slope, would minimize likely errors. For the purpose of the study, and in our best judgment, the static data layers provided useful information, not easily gathered for remote study areas in developing countries.

5. Conclusion

In the current study, we examined spatial distributions patterns and the significance of selected demographic, infrastructure and topo-edaphic factors in influencing observed greenhouse location in 84 sub-watersheds using available datasets and Google Earth archives. To the best of our knowledge, the current work is the first attempt to map the spread and dynamics of greenhouse horticulture production at sub-watersheds level in Kenya. Sub-watersheds are chosen as study units since, as a unit, they balance well between data collection and production of meaningful results and feedback that would serve overall best watershed management plans in the future. The study focused on the central highlands of Kenya due to its diverse agro-ecological (humid to very dry arid areas) and sharp rainfall gradients that prevail between the uplands and lowlands.

Results indicate a non-uniform spread occurring as high clusters in a few sub watersheds and low clusters in other. An increase in number of greenhouses and area under cultivation is evident between 2000–2011, possibly due to better trade terms, wider international market, diversification and increased fresh produce demand. We also find four statistically significant factors: population density, the average rainfall, average slope and dams that accurately predicted the observed spatial spread across area of study. While mainstream literature notes that the country has good climatic condition and soils, the ability of these factors to solely explain the observed patterns is debatable. Mapped distribution highly suggests dependency of greenhouse locations on multiple factor combinations rather than a single factor in selecting sites for production.

Mapping greenhouse distribution and deriving the area under cultivation in Kenya using accessible, online tools such as Google Earth is novel, since these data are currently unavailable, neither documented among other publically available export statistics. Having established the extent of distribution, the next goal would be to investigate environmental impacts of such farming on sub-watersheds, examining highly vulnerable regions.

Derived quantitative data on production area is useful in modeling environmental impacts of agricultural practices such as the application of pesticides and fertilizers, irrigation and product transport. The results are also applicable in decision making and priority management such that permitting regulations governing intensity of commercial farming in a given area, considering other ecosystem stressors are emphasized. Sub-watersheds vary greatly in size, endowed resource, number
of stakeholders sharing resources, etc. These are factors for consideration when stipulating where to allow permits; where to increase amenities and prioritization of infrastructural development; provision of better sanitation, etc. A potential shortcoming of the approach used is that all mapped greenhouses accounted for are assumed to carry out commercial horticulture production. From experience working in this sector, only flowers were found in greenhouses. The findings are more or less region specific since climatic conditions and infrastructure driving export farming can vary diversely.

The study extends knowledge on spatial patterns of greenhouse farming and clarifies the significance of chosen key parameters in influencing choice of greenhouse sites. Further studies can explore land use dynamics in hot spots zones as they can give an indication of watershed resource status and level of sustainability in greenhouse farming.

Acknowledgments

We gratefully acknowledge the extensive public GIS data provision by International Livestock Research Institute (ILRI). We would also want to thank colleagues Rocio, D. Onoro and Catherine, O. Alexander, for their assistance in providing helpful suggestions on the manuscript.

Author Contributions

All authors have contributed substantially to conception of this paper. They designed the study and developed the methodology. Faith Justus collected the data, performed the analysis and wrote the manuscript with guidance, directions and detailed edits from Danlin Yu.

Conflicts of Interest

The authors declare no conflict of interest.

References


41. Tavakkoli, E.; Rengasamy, P.; McDonald, G.K. High concentrations of Na⁺ and Cl⁻ ions in soil solution have simultaneous detrimental effects on growth of faba bean under salinity stress. J. Exp. Bot. 2010, 61, 4449–4459.


**Appendix**

**Figure A1.** A picture of rose plants growing directly on slightly raised soil beds in a greenhouse. The vegetative bending forms a base from which upright shoots grow to stems, and are harvested as cut flowers (the picture taken by Kimani M. David, on 10 November 2013 in Kenya).

![Figure A1](image1)

**Figure A2.** A picture of hypericum plants directly growing in the ground, on low beds that are covered by plastic under a greenhouse. At the foreground is support for drip line used for irrigation (the picture taken by Kimani M. David, on 9 February 2014 in Kenya).

![Figure A2](image2)