Ecological Modelling 484 (2023) 110476

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Towards early response to desert locust swarming in eastern Africa by estimating timing of hatching

Tobias Landmann [a,*], Komi M. Agboka [a], Igor Klein [b], Elfatih M. Abdel-Rahman [a,1], Emily Kimathi [a,1], Bester T. Mudereri [a], Benard Malenge [a], Mahgoub M. Mohamed [c,d], Henri E.Z. Tonnang [a]

[a] *International Centre of Insect Physiology and Ecology, Nairobi, Kenya*
[b] *German Aerospace Center (DLR), German Remote Sensing Data Center, Oberpfaffenhofen, 82234, Weßling, Germany*
[c] *Faculty of Agricultural Studies, Sudan University of Science and Technology, Khartoum North, Sudan*
[d] *Plant Protection Directorate (PPD), Ministry of Agriculture and Forests, Khartoum North, Sudan*

A R T I C L E I N F O

Keywords:
Migratory insect pests
Locust hatching
Ecological modeling
Fuzzy logic
Africa

A B S T R A C T

Desert locust (*Schistocerca gregaria*) plagues threaten agricultural production, food security and the environment across Africa, the Middle East, and Southwest Asia. Control methods targeting adult desert locusts present significant challenges and financial costs. Recognizing this, we developed a ground-breaking fuzzy set Mamdani type inference model that provides an innovative solution for early warning alerts. The model aids in predicting the juvenile stages of locust development, thereby preventing wide-scale locust swarming and mitigating its extensive damages and socioeconomic costs. The novelty of our approach lies in our unique application of environmental variables relevant for locust breeding to estimate the timing and location of desert locust hatching. Additionally, we improved the algorithmic handling of these variables, with localized desert locust bands data used as a proxy for hatching timing with a temporal offset of 35 days. The model's boundary conditions were determined using a training area in Sudan, where comprehensive ground data was available. This rule set was subsequently applied to Turkana County in Kenya, a data-scarce region, demonstrating the model's applicability and success in different contexts. The model's accuracy, assessed by data from the Sudan training site, demonstrated a remarkable score of 82% for true predictions. Furthermore, the model correctly identified the months of highest hatching probabilities in Turkana during 2020, demonstrating its real-world effectiveness and practical value. A correlation analysis affirmed that hatching was associated with increases in chlorophyll levels and precipitation accumulations. Our study marks a significant advancement in predicting the timing of hatching using fuzzy logic in data-scarce environments. By operationalizing more targeted early responses to desert locust infestations, our model facilitates more effective locust control. This study stands as an important contribution to locust management strategies, with substantial implications for agricultural production and food security in affected regions.

1. Introduction

Desert locusts (*Schistocerca gregaria;* family *Acrididae*) are a locust species, that occur in semi to arid regions in Africa, the Middle East, and Southwest Asia. As swarms they sometimes invade countries and areas far away from their breeding grounds. The life cycles of desert locusts are initiated with egg laying in deep sandy soil, with incubation periods being 10 to 65 days (Cressman, 2021). Hatching and molting subsequently occur, and as they mature, individuals can transform from a harmless solitary state to the destructive gregarious phase, whereby hoppers (juveniles in their early, wingless stages) march together in bands. Desert locust development phases vary due to the phenotypic plasticity of the pest that can be abruptly triggered by environmental changes (Nevo, 1996), such as successive above-average rainfall (Symmons and Cressman, 2001). The polymorphisms of the species, thus, influences localized timing of hatching and subsequent unified

* Corresponding author.
E-mail address: tlandmann@icipe.org (T. Landmann).

[1] These authors contributed equally

https://doi.org/10.1016/j.ecolmodel.2023.110476
Received 31 May 2023; Received in revised form 12 July 2023; Accepted 31 July 2023
Available online 9 August 2023

swarming of the adults, even across continents (Pflüger and Bräunig, 2021).

Swarms can travel up to 120 kms per day and near completely consume croplands, rangelands and natural vegetation, within a short time (Rowell and Hemp, 2017). As such, desert locusts are often considered one of the most devastating and unpredictable pests in drylands (Van Huis, 2007). A total of 65 countries in Africa and Asia have been found to be affected by desert locust plagues (Skaf et al., 1990). Desert locusts are a risk factor for agro-pastoral communities in semi-arid to arid regions that are already vulnerable to climate stressors (Spinage, 2012).

Widespread plagues can be intensified by extreme weather events because of climate change and climate variability. For instance, the tropical cyclone “Laban”, which occurred off the coast of eastern Africa in 2019, resulted in excessive and persistent seasonal rainfall (>400% over the long-term mean), not experienced over several decades. This caused breeding and swarming of desert locusts across the region (Chen et al., 2020). In Kenya, about 107,000 km^2 (or 20% of the total land surface area of Kenya) was impacted by the 2019–2021 locust plagues (Kimathi et al., 2020).

The development of hopper bands in eastern Africa is amplified by the continuous availability of dense “green” vegetation and suitable soil conditions because of rainfall anomalies over various seasons. However, the spatio-temporal interactions between climate and vegetation, and how changes in biotic factors and land use dynamics affect this interaction are not well understood (Tratalos, 2001). This poses challenges for early warning and risk forecasting models.

Recent strides have been made to discern outbreak and breeding areas (e.g., Kimathi et al., 2020; Klein et al., 2022) and characterize phase dynamics (changes from solitary to harmful gregarious state) (Topaz et al., 2012). However, there are still many specific factors and unknowns that are yet to be identified to improve real time detection and early warning. For instance, one of the key challenges is detecting early states of swarming which is still impossible in remote and vast areas where locusts breed (i.e., desert, and arid areas in Africa and the Middle East). This is exacerbated by political insecurity in most regions where the desert locust occurs and breeds as well as the deficiency of on-site expertise needed for continuous, and long-term monitoring and early detection (Gay et al., 2020). As a result, the United Nations (UN) Food and Agriculture Organization (FAO) initiated the Desert Locust Information Service (DLIS) platform to provide early warning of swarming (Symmons, 1992). However, the DLIS provides information at a coarse scale (0.25° grid cell resolution). Yet ground-based breeding sites monitoring systems are needed that make use of explicit information encompassing the early life cycle stages of the pest (Meynard et al., 2020). This entails making use of spatially and temporally explicit environmental variables that are known to sustain hatching over larger areas. Therefore, characterizing early life cycle stages of the pest (i.e., hatching) using spatial and temporal variables explaining their life cycle dynamics, would enable the early anticipation of swarming (even before field teams “observe” hopper bands in the field) (Salih et al., 2020). Hence, spatially modeling hatching regions and timing will help to offset the already constrained limited resources needed to conduct ground surveys (Meynard et al., 2020).

Some local to regional studies were designed to predict hatching triggers and tie them to the time of hatching. However, due to the complexity of the ecological processes, some ambiguity and inconsistency were observed in other modeling work in this regard. For instance, some studies have shown contradictory results in terms of variables importance (Rhodes and Sagan, 2021). While soil temperature and moisture were observed as an important variable (Nishide et al., 2017), it can also inhibit hatching when excessive rainfall occurs (Cressman, 2013). Essentially, from previous studies, it can be summarized that over vast tracts of largely inaccessible land, solitary hoppers and bands are difficult to consistently monitor using field surveys. Hoppers also often exhibit a sedentary behavior, before swarming spontaneous and rapidly occurs (Brader et al., 2006). Early life cycle stages of the pests (i.e., bands and hoppers) can be easily missed. Current desert locust monitoring is largely associated with assessing periods of rainfall and green vegetation upsurges that sustain bands or hoppers (Mongare et al., 2023; Wang et al., 2021). There is clearly a gap, in that geospatial desert locust occurrence monitoring routines have evidently, thus far, excluded considering modeling hatching suitability, before the bands and hoppers life cycle stages.

In view of the above gaps, the objective of this study was to use spatio-dynamics variables and a data driven knowledge-based fuzzy logic model to predict timing and location of desert locust hatching. The model rules are applied to Turkana County in Kenya to predict egg hatching occurrence probabilities in time and space. In targeting the early life stages, we hope to contribute to the development of early warning systems, important for remote areas within semi-arid and arid regions of Africa.

The model choice (fuzzy logic) is tailored to the data characteristics found in the study areas, i.e., deducing rules from available desert locust ground observations (here we used bands data as a proxy for hatching, given a temporal offset). Moreover, we used fuzzy logic rules over other modeling approaches because of the ability of this method to deal with data gaps, in our case data gaps on bands and hatching in the Kenyan target area. The developed rules use straightforward ecological principles related to the knowledge of the life cycle changes of the pest (Ricotta, 2000). Zadeh (1965) developed the first fuzzy logic model and theory based on linguistic principles based on vagueness or fuzziness and expressed in membership functions.

2. Method

2.1. Study site(s)

The primary study site used to develop the model rules was northeastern Sudan, specifically the coastal area of the Red Sea that stretches between Port Sudan (19.42°N/37.26°E) and Aqig (17.79°N/38.42°E) (Fig. 1a). This study site was selected because it provided the most consistent and highest multi-temporal field data on desert locusts' bands spanning the period 2013 to 2021.

The Sudan desert locust recession and training area is characterized by frequent occurrence, breeding and outbreaks of the pest. The Sudan study site covers a spatial area of 70.900 km^2 and is characterized by a large spatial variability in rainfall amounts, from 39 mm per annum in the inland areas to 164 mm on the southern coast with an annual mean of approximately 111 mm (El Gamri et al., 2009). Rainfall occurs in the “winter” months, from November to February. The daily maximum temperatures are often > 30 °C, with daytime averages from 20 to 27 °C. The area is characterized by semi-desert vegetation, that readily responds to rainfall events (Halwagy, 1961). These areas contain different desert locust habitats, such as plains and wadis. Furthermore, these habitats vary in soil type (fine silty, sandy dunes and sometimes inter-mixed with fine gravel), vegetation density (low, medium and dense). Moreover, this zone is quasi-permanent habitats for desert locust breeding and development and considered as most frequent for desert locust outbreaks during the rainy season. The study area is dominated by several of the main annual vegetation species preferred by desert locust such as *Tribulus spp.*, *Dipterygium glaucum, Heliotropium spp.*, *Pennisetum typhoideum*) and some perennials e.g. *Panicum turgidum, Acacia spp.*, *Suaeda spp.* and *Callotropis procera*.

Turkana county in Kenya was the target or invasion area (with infrequent outbreaks), used to project the developed rule set to (Fig. 2). Turkana county covers an area of approximately 68.500 km^2. This area is characterized by a large inter-annual rainfall variability with two rainy seasons, from April and July (long rains) and then between October and November (short rains). The annual average is 200 mm, and average daytime temperatures are between 20 and 35 °C (MoALFC, 2021). The vegetation is more diverse than in the Sudan area, and is

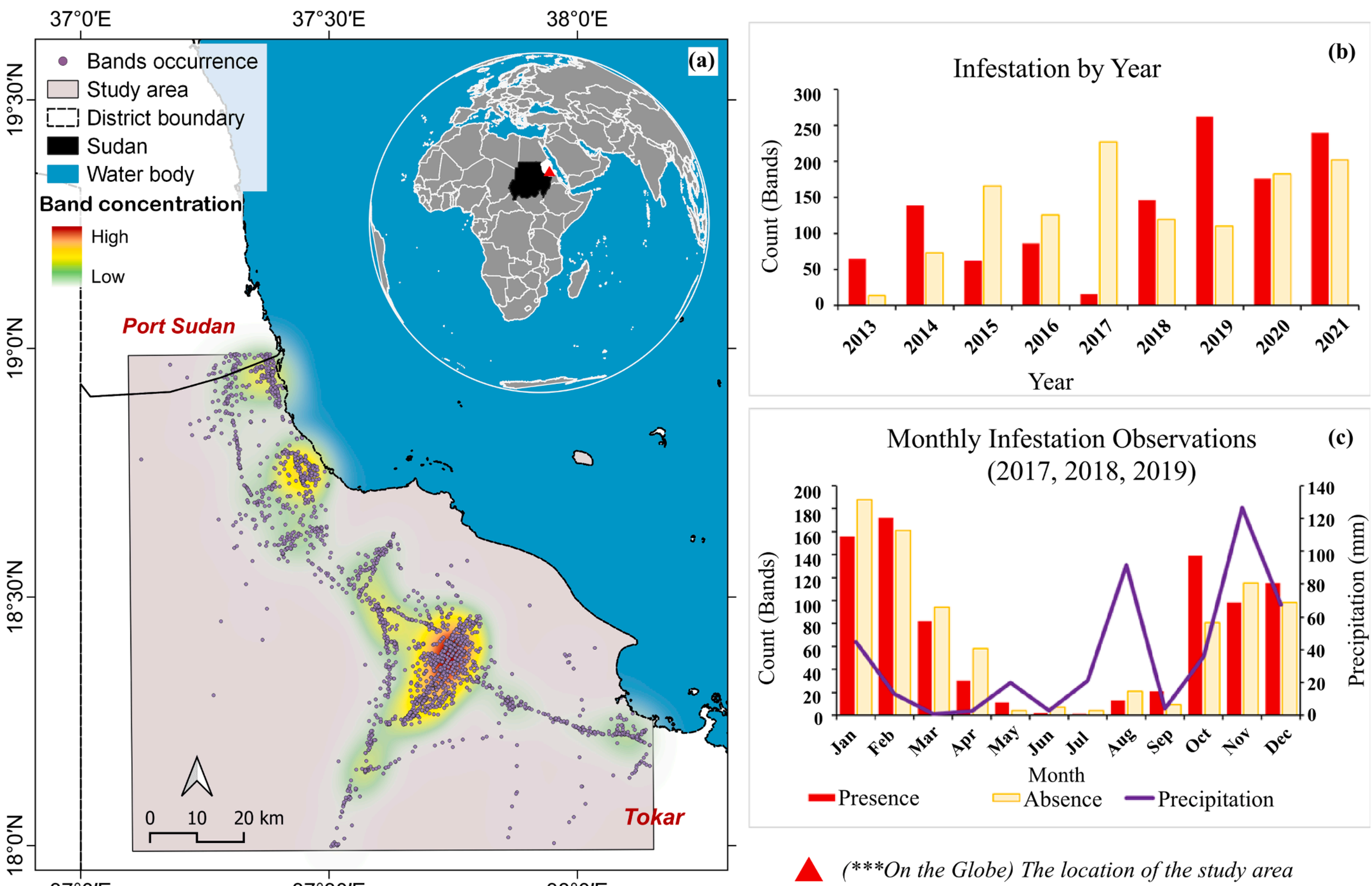

Fig. 1. (**a**) Location of study area in Sudan (East Africa). The spatial and temporal distribution of locust bands on the ground in the Sudan "training" area is used for rule development. The heat map shows concentrations, that is densities of hopper bands that were observed over the sampling period (2013–2021). Brownish-reddish areas exhibit high data densities. (**b**) Infestation (bands counts) presence (red) and not present (orange) over each year in the 8-year period. (**c**) Monthly bands observations for the years 2017- 2019 (priority period), according to observed presence (red) and non-present (light yellow). Precipitation (blue line) denotes the mean annual precipitation (mm) over the same period.

characterized by savanna and semi-desert vegetation, mostly annual grasses (mostly genera *Aristida*) and dwarf shrubs, with large *Acacia* trees in some places. The area is dominated by haplic xerosols soils with sandy loam or loamy sand texture in some places (MoALFC, 2021).

2.2. Datasets

The data used in this study are field occurrences of hopper bands, which were collected in Red Sea area in Sudan, as daily observations from 2013 to 2021. The collection was conducted by the desert locust survey and control officers of the Plant Protection Directorate (PPD) of Sudan. This activity is part of a regular runtime activity to conduct ground survey operations in summer and winter desert locust breeding areas. These well-trained field observers collect data on the desert locust presence, habitat type, breeding sites, locust density, growth phases, stages and relative behavior at the time of observation. The surveys data are saved in reconnaissance and monitoring system of the environment of Schistocerca (RAMSES). This is a computerized GIS-based platform (i.e., RAMSES), through which desert locust occurrence observations at a national scale can be retrieved (Cressman and Hodson, 2009).

We retrieved monthly georeferenced observation points from 2013 to 2021 from the RAMSES platform. The data were cleaned (i.e., absence observations and other hopper stages were excluded) and used in the modeling experiments. The 8-year band data showed some data gaps in certain months due to non-collection of data, however, annual patterns could be established in areas where bands were present and not present. The highest number of band observations were made in the last five years, that is, 2017 to 2021 (i.e., infestation by year, Fig. 1b). The highest presence was observed in 2019 (255) and the lowest in 2017 (15). Observations corresponding to 2018 had the second highest presence midway between 2017 and 2019 (145) (Table 1, supplementary material). Hence, the data obtained in these consecutive years were analyzed using descriptive statistics (Fig. 1–33 in the supplementary material section). These results facilitated the development of the rules embedded into the model.

Furthermore, the obtained data for the priority period (2017–2019) presented intra-annual patterns demonstrating that the highest presence of bands was observed between January to March corresponding to the 1st quarter of the year and from October to December (4th quarter) in at least 2 of all 3 years (Fig. 1c). In the 2nd and 3rd quarters in each year, respectively, occurrence was lower. The rule-based relationships between infestation and the climate and environmental variables were thus developed for these priority years (2017–2019). The data from the most relevant quarters of each year were used to establish the relationships between the bands and the climate and environmental variables.

The Terra-Climate (Terra-Clim) data provided at ~4-km grid/pixel resolution (https://www.climatologylab.org/) were used as the main climate data (the variables used under terraclim) (Table 1). The "greenness" data were obtained from Vlaamse instelling voor technologisch onderzoek (VITO) at a pixel size of 250-meter. The "greenness" datasets were resampled to 1-km pixel resolution, and the monthly means for the priority period were also produced at this resolution. The Terra-Clim grid data over Africa show a good match with station data

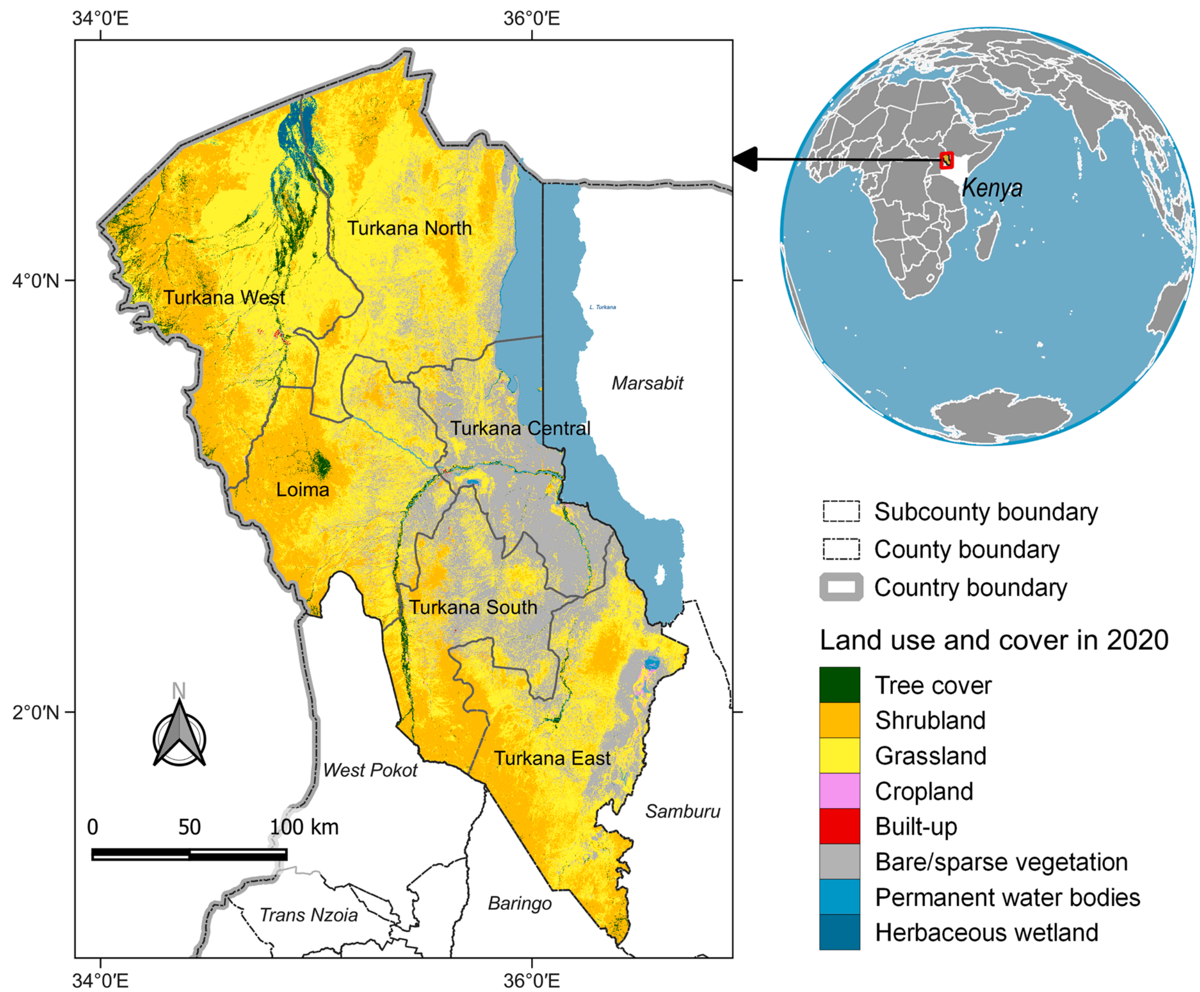

Fig. 2. The location of Turkana County, i.e., the model implementation target area in Kenya. Land use data source: © ESA WorldCover project / Contains modified Copernicus Sentinel data (2020) processed by European Space Agency (ESA) WorldCover consortium.

Table 1
Input data summary overview, showing data source and pixel/grid cell resolution, temporal resolution, data type, measurement unit and data description.

Data (Unit)	Source	Spatial Resolution	Year	Description
Precipitation (mm)	https://www.climatologylab.org/	1/240 (~4 km)	1958–2020	Precipitation, monthly total
Temperature (Min/Max) (°C)	https://www.climatologylab.org/	1/240 (~4 km)	2013–2021	Mean temperature, monthly average
Sand content (cm)	https://data.isric.org	250m	2021	Sand in depth of 5–15 cm (g/kg)
Soil moisture (cm)	https://www.climatologylab.org/	1/240 (~4 km)	1958–2020	Soil moisture, total column - at end of month
Greenness (Daily)	http://irid.ldeo.columbia.edu/SOURCES/.EU/.VITO/.DevCoCast/.greenness	250m	2013–2020	Dekads since vegetation onset

over various biomes (Abatzoglou et al., 2018; Jain et al., 2022). The vegetation “greenness” data (categorized in decade since greening onset) is a compilation of middle infrared, near infrared and red wavebands information, from various satellite platforms, that is cloud corrected and available as 10-day composites (Pekel et al., 2010). The “greenness” variable was specifically developed for arid and semi-arid biomes and suitable for vegetation dynamics in relation to locust breeding (Kimathi et al., 2020). For the implementation site (Turkana), the variables were computed as monthly means from 2017 to 2019.

2.3. Key assumptions

- The absence of observation represents periods of non-optimal conditions for bands occurrence
- Hatching occurs 35 days prior to band formation. This assumption is based on the fact that the duration of the sequence of favorable environmental conditions, that are known to proliferate hatching prior to the sighting of bands in the field, is usually set at 35 days (Symmons and Cressman, 2001).

2.4. Overall modeling approach

Overall, we applied Fuzzy Mamdani to estimate the hatching timing through a combination of membership functions. These are based on environmental variables with hyperparameters derived from rules derivation and descriptive analysis of the available datasets (see supplementary). The overall approach is illustrated in Fig. 3, and each specific step is described below.

Step 1 consisted of the identification of the problem, which, in this case, is predicting the location and timing of desert locust hatching

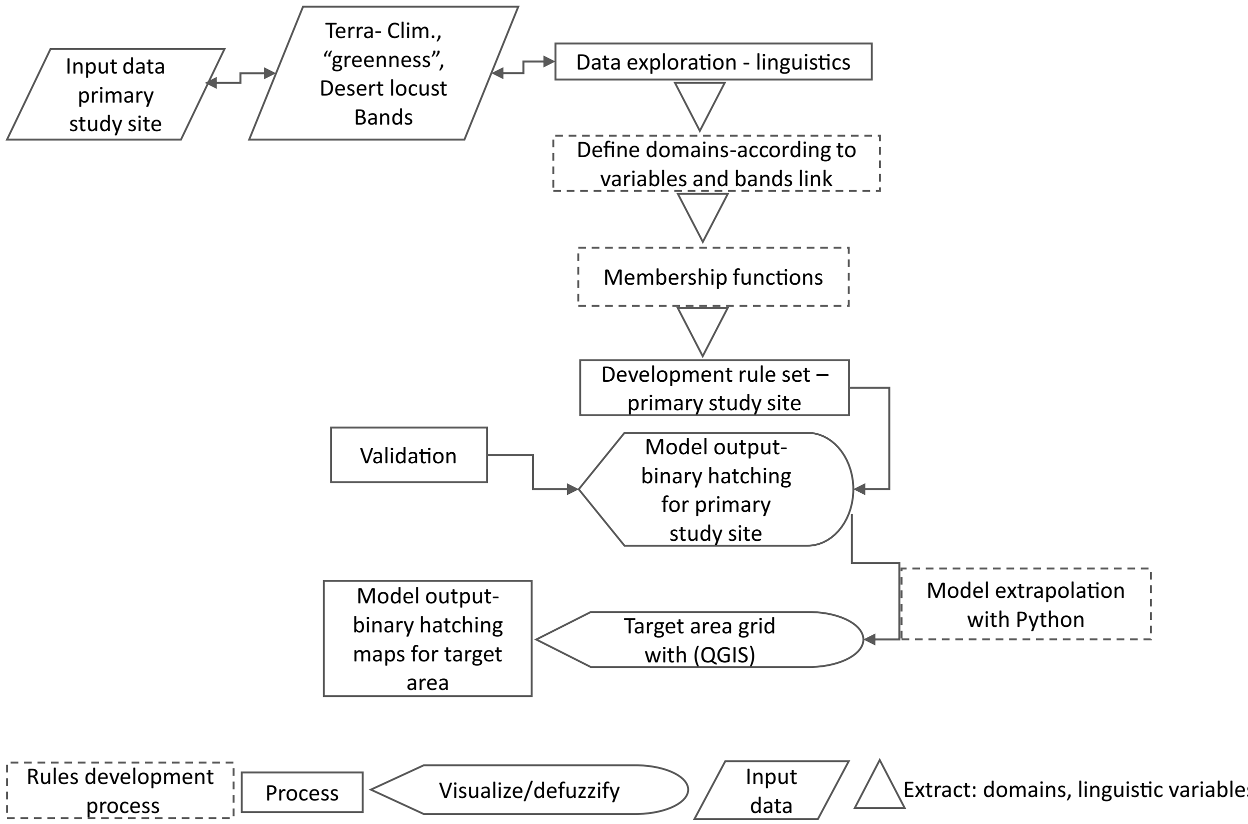

Fig. 3. Overall workflow summarizing the different steps used in the model development.

probability. We further identified climatic and vegetation variables (Table 1). These variables were chosen for analysis because they have been reported by earlier studies to trigger egg laying, hatching and sustain the development of the band phases (Gomez et al., 2019; Kimathi et al., 2020).

In *Step 2*, we defined the linguistic variables and their membership functions (Fig. 4a-e). Sigmoid, gaussian, trapezoid and triangular functions were selected to convert the input variables into fuzzy sets. Our selection was based on how best the function represents the input variables. Each membership function was iterated with a combination of

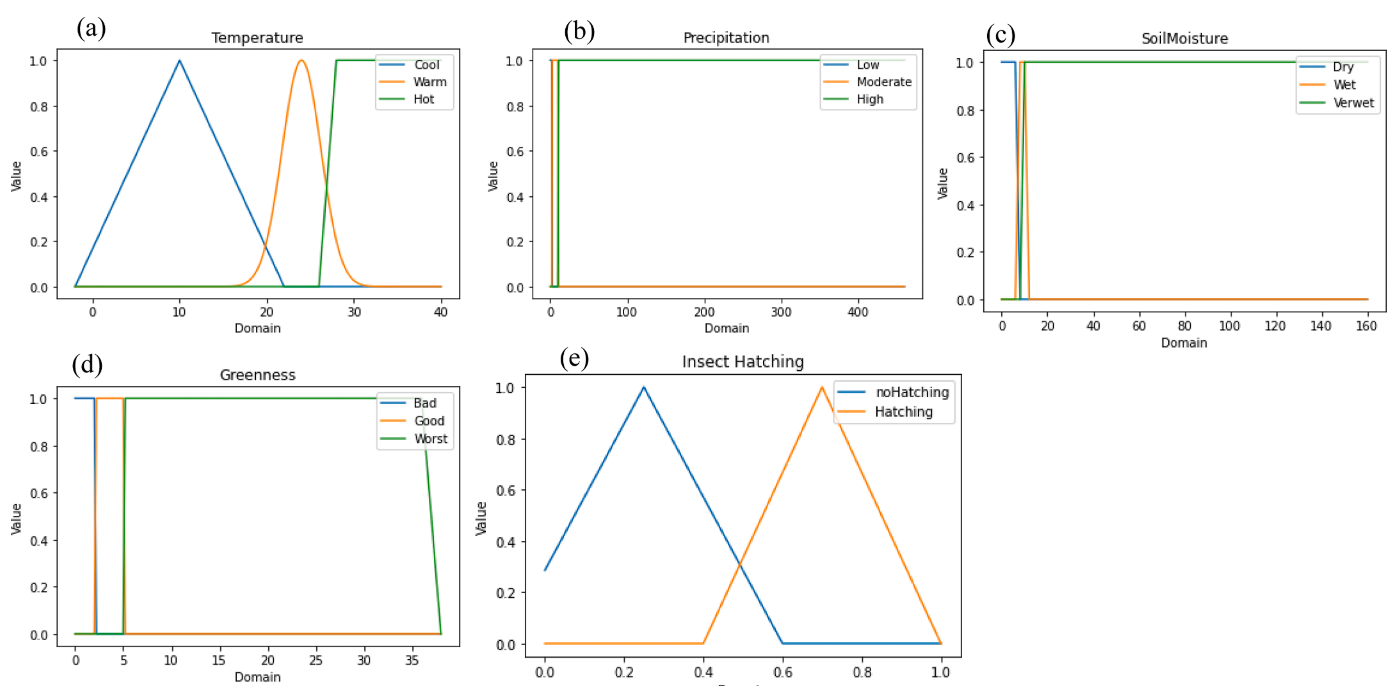

Fig. 4. Membership function curves for variables/bands data for 0–1 memberships, and linguistic categories using sigmoid, gaussian, trapezoid and triangular memberships. Membership grades increase from the left to the right for each domain/variable.

different membership functions. The optimum combination that gives the highest accuracy was adopted for each variable. For the priority years, the monthly value ranges, means, and medians were calculated for each climatic and environmental variable (Fig. 5). These statistics were used to define and adjust the hyperparameters of the membership functions, such as the center and width.

Step 3 entailed the creation of the fuzzy rules, which use the set relationship between the input variables and the output variables. They are expressed in the form of "if-then" statements, where the antecedent (if-part) contains one or more fuzzy sets defined on the input variables, and the consequent (then-part) contains a fuzzy set defined on the output variable. An example of fuzzy rule is, "*if the temperature is high and the soil moisture is low, then the insect hatching probability is likely to be low*". A total of 81 rules were developed to account for the different combinations of the linguistic language rules. Once developed, the rules were combined using fuzzy logic operators. The most common operators are AND, OR, and NOT; however, we opted for AND operator as it represents best the relationships between our selected input variables and the output variable.

Step 4 was made of 3 sub-steps: the mapping of the input crisp values to fuzzy values using the membership functions defined above (such as "very high", "high", "medium","low," and "very low"); then applying the fuzzy rules to the fuzzy inputs to obtain a fuzzy output; followed by the conversion of the fuzzy output into a crisp value. Although this last sub-step can be done using various methods such as centroid, maximum, or mean of maximum, for simplicity we used the centroid approach for the conversion.

The model was tested and validated (*step 5*) by comparing its output with the expected output for a set of input values i.e., 2020 that were not used while developing the rules. We expressed these values in the form of probability in which the probability of hatching = 1 and no hatching = 0. When the results were not satisfactory, they were refined by adjusting the membership functions or the fuzzy rules.

The developed rule-based model was applied (*step 6*) to provide a spatial visualization of the probability of hatching timing for a given month and year.

2.5. Mathematical computing of the fuzzy logic model

The Fuzzy Logic system script was written in Python using the fuzzylogic library. It reads input data related to various environmental conditions and predicts whether or not hatching will occur.

The mathematical model were defined using fuzzy logic membership functions and fuzzy logic rules. The key components are as follows:

Domains and Membership Functions: The program defines a number of domains, which are essentially the environmental variables with ranges and resolutions. For each domain, there are several membership functions. These functions assign a degree of membership to each value in the domain. They are used to define fuzzy sets. For example, the temperature domain was defined as:

The fuzzy sets for the temperature domain were "cool", "warm", and "hot". The degree of membership to each of these sets was determined by the respective membership functions (Fig. 4). Other variables domains are provided in the supplementary files (Supplementary Figures 29–34).

Fuzzy Logic Rules: These rules map combinations of fuzzy sets in the input domains to fuzzy sets in the output domain. The goal of these rules is to predict whether locusts will hatch or not based on the input conditions. For example, one of the rules was:

This rule stated that if the temperature is "cool", the precipitation is "low", the soil moisture is "dry", and the greenness is "bad", then there will be no hatching probability.

As stated, there were 81 such rules in the script (see supplementary materials section on computing rules), each addressing a different combination of input conditions and resulting in a prediction of either "no hatching" or "hatching".

In the computing, the rules are not explicitly defining a traditional "if-then" logic but rather creating a multidimensional mapping between the input conditions and the output prediction, which can be thought of as a kind of multidimensional interpolation or pattern recognition.

Prediction: The script reads in data from a csv file "**terra_green200.csv"**, transforms some data to a numerical format, and then utilizes this data along with the established fuzzy logic rules to make predictions.

Plotting: The script contains functions to plot the membership functions of each domain. These plots help to visualize how the fuzzy sets are defined within each domain.

3. Results

The FIS model developed exhibited an accuracy of 82% of true prediction based on the model result for the Sudan training area, and the validation data from the year 2020. Using the centroid approach for the conversion of the developed rules (Fig. 4), the model was applied to the target area, Turkana County. The spatio-temporal visualizations for Turkana were produced as monthly binary maps for the priority period (2017–2020) (Fig. 6), depicting hatching probability. Areas of high hatching probability or high risk are shown as reddish colored pixels (or "true") and blue illustrates "false" or no risk. The northern part of Turkana has overall the highest hatching probability in each year, especially in an average to below average rainfall year such as 2017. From November to December in 2019 onwards, there is a marked increase in hatching probability throughout the Turkana target area. This trend is sustained until May 2020. This period corresponds to the severe desert locust outbreak in Kenya in 2020.

Based on the binary outputs for the priority period, a 5-km grid cell frequency of hatching events "heat map" could be produced for Turkana County (Fig. 7, left image). These maps help to visualize hatching suitability risk areas. The aggregated frequency "heat map" map (0–1) shows large frequencies of predicted desert locust egg hatching over the northern and the southern parts of Turkana County, respectively. The pixels with a dark redish fill and light green boundaries indicate locations with a frequency of monthly hatching >0.5. These locations can provide suitable zones hatching at medium to high probabilities and make up >40% of the total land surface area of Turkana County. "High hatching suitability" or high hatching risk areas (>0.75 frequency scores) were found to make up 8% of the total area of the county.

Three randomly selected "high frequency" (>0.75 frequency scores) hatching probability pixels (IDs 834, 1011, 1914) were extracted from the frequency heat map (Fig. 7, left). Graphs were produced to illustrate timeline profiles for egg hatching probability for each of the three pixels (IDs) (Fig. 7, right). The three randomly selected IDs are located within the north, center and south of the Turkana area. For the two northerly pixels (IDs 834 and 1011; north to center of the study area), hatching suitability/probability was predicted to be highest between March to May over all years, while in the south (pixel ID 1914), hatching suitability was predicted to occur mostly between October and November for all years. Temporal frequencies were also more varied in the southern part. For the outbreak year 2020, however, high hatching suitability's were consistently found between January and May, over all IDs.

The associations between hatching suitability and rainfall and "greenness" were well established by the rule set using the Sudan bands data, especially over the priority period 2017–2019 (Fig. 4, and Figs. 1 and 36 in the supplementary materials). Specifically, the highest rainfall skewness and distribution upwards (29.9 mm rainfall) was found to match the highest bands occurrence and with that hatching, given the used time lag of 35 days. Similarly, cumulative "greenness" over the eight-year and over the three-year priority period, because of higher rainfall in December, was found to be a factor for bands occurrence and hence hatching (e.g., Figs. 1, 35 and 36 in the supplementary materials).

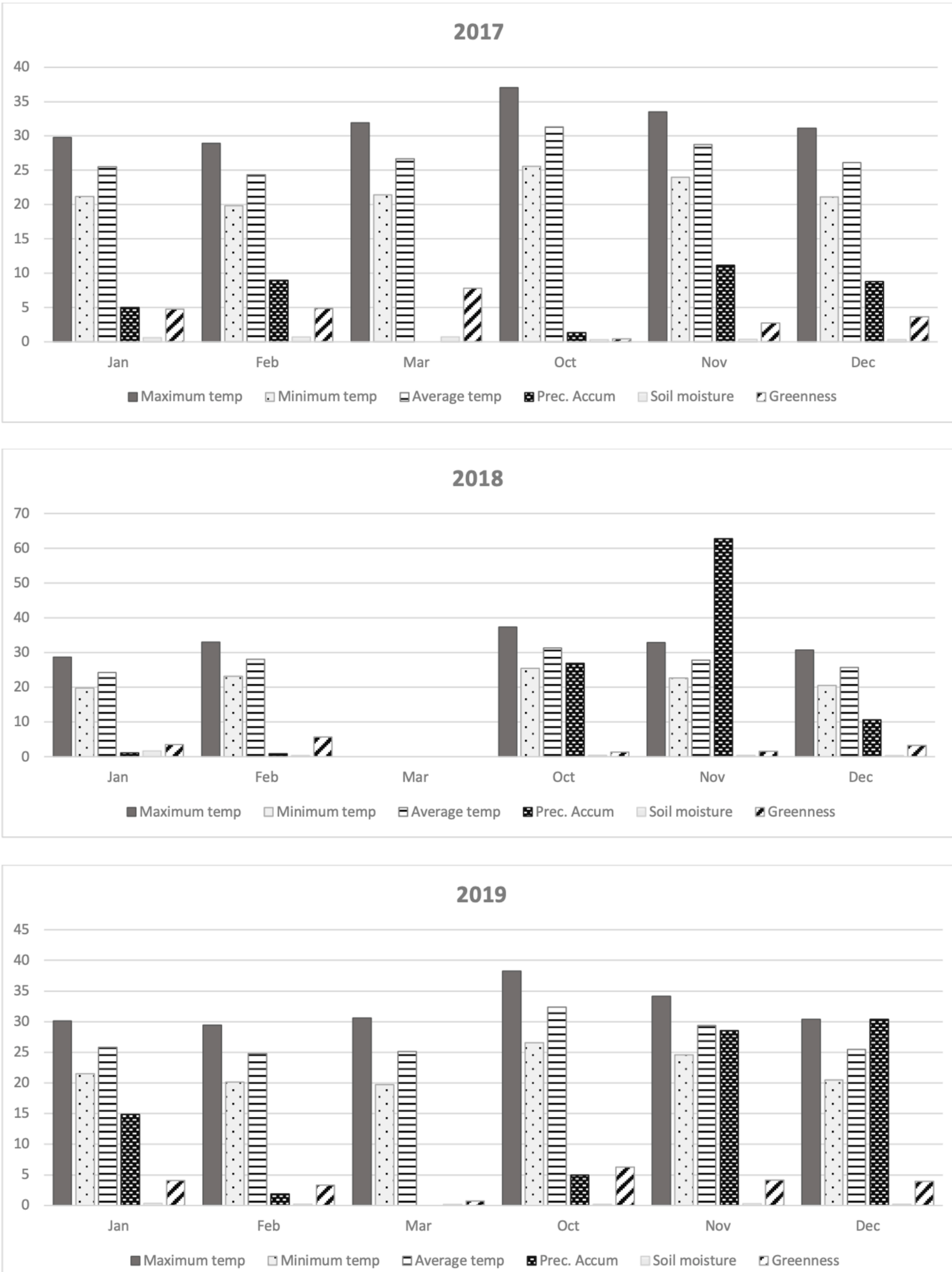

Fig. 5. Mean statistics of terra-climatic and "greenness" variables from 2017 to 2019 (priority period) according to months indicated as numbers (1–12) that are within the 1st and the 4th quarters. These values were used to develop the hyperparameters for the model.

temperature = Domain ("Temperature", −2, 40, res=0.1) temperature.cool = triangular(−2, 22) temperature.warm = gauss(24, 0.1) temperature.hot = trapezoid (26, 28, 40, 42)

R1 = Rule ({(temperature.cool, precipitation.low, soilMoisture.dry, greenness.bad): insect.noHatching})

4. Discussion

4.1. Novelty of the study

The novelty of this work is that location and timing of desert locust hatching probability/suitability could be modeled, which is a critical variable needed for locust early warning and early response, especially in data scarce environments (Cressman, 2013). Hatching events over longer periods are indicative for subsequent swarming and large-scale desert locust outbreaks, e.g., as was the case in eastern Ethiopia in 2019 (Salih et al., 2020). In other similar work, such as by Kimathi et al. (2020), hatching was predicted by using varied ecological variables to monitor seasonal desert locust breeding grounds. This included the hoppers and other adult stages. However, monitoring specifically juvenile life stages and hatching has not been yet systematically performed or considered (Maeno et al., 2020). The spatial and temporal patterns we found for hatching (Figs. 7 and 9) correspond well to the lagged presence of hopper bands; Sun et al. (2022) found that May 2020 exhibited the highest abundance of bands (that this is hatching activities) throughout north-western Kenya.

Another novelty was the implementing of a data-driven fuzzy logic model that encapsulates data science and knowledge related to the ecology of the pest. For instance, the membership functions (Fig. 4) are determined based on the spatial-temporal dynamics of the environmental variables in relationship to the presence and non-presence of bands using data exploration (bands were used as a proxy for hatching, given a temporal offset of 35 days). These domain specific relationships can be linear or rather follow a gaussian membership function (Babuška, 2012). Here, we assumed that phase dynamics are related to hatching, that is the presence of bands. The Fuzzy rule-based system is ecologically stable from location to location because they do not rely on specific data points or hard-coded rules. Instead, they use fuzzy logic to create rules based on linguistic variables that are expressed in natural language. These rules are designed to capture the uncertainty and imprecision that are often present in real-world systems, which renders them more adaptable to different environments (Hagras et al., 2012). Furthermore, fuzzy rule-based systems are better suited for situations where data is scarce or unreliable because they do not require precise input data (Rahim et al., 2018). Instead, they can operate with approximate or incomplete data, making them more robust in situations where data are difficult to obtain, or the quality of the data is poor. This is because fuzzy logic allows for degrees of truth, rather than binary true/false decisions, which allows the system to make decisions even when there is uncertainty or ambiguity in the data. Overall, fuzzy rule-based systems are a powerful tool for modeling complex, uncertain systems, and they are well-suited for applications where data are sparse or unreliable.

The hatching probability frequency in Turkana ("heat map", Fig. 7) is driven and triggered by persistent rainfall and associated and lagged vegetation "greenness" (to a lesser degree by soil moisture, e.g., Fig. 5 and membership function in Fig. 4).These three variables have been confirmed by prior studies to be of utmost relevance for biological processes to induce hatching and breeding in eastern Africa (Nishide et al., 2017). Although extended flooding (> 14 days) has been found to adversely affect hatching (Woodman, 2015), this has been controversial. In yet other arid regions with sandy and clay soils, extended flooding was profoundly associated with high hatching activities and subsequent swarming (Gomez et al., 2019). Moreover, using primarily rainfall as a proxy for hatching can lead to erroneous model outcomes since landform, run off and evaporation may locally affect moisture collection and egg hatching as a life cycle responds (Dinku et al., 2010). "Greening" due to vegetation chlorophyll activity and density, even if patchy, interrupted and fine scaled, seems to be more profound determinant for monitoring hatching location and timing and the phenotypic plasticity of the species (Salih et al., 2020). As shown in the associations of variables with locusts bands, time-line remote sensing observations can be effectively integrated to indicate and monitor "greening" events over large areas (Klein et al., 2022). The availability of vegetation further sustains the development of hopper bands since, at this life stage, they require vegetation for nutrition and shelter (Sun et al., 2022).

The general monthly binary hatching probability patterns (Fig. 6) illustrated that, just like the frequency map, the center of Turkana County is to a lesser degree affected by hatching within the modeled period. In inspecting a reference vegetation map for the region (Alkhalil et al., 2020), it became clear that high hatching probability areas (red pixels) are largely characterized by sparse vegetation (i.e., grasslands or open shrublands). Denser vegetation and very sparse to bare areas, as found in the center of Turkana County, were modeled as being low risk areas (Figs. 6 and 7). This is corroborated by Dong et al. (2023) who stated that desert locust life cycles in arid areas, including hatching, are supported by sparsely vegetated areas (i.e., grasslands or open shrublands) given appropriate edaphic conditions. The hatching patterns, moreover, largely follow rainfall and "greening" patterns. From December to April rainfall amounts are enhanced compared to the latter part of the year (Kinyua, 2021). Consequently, hatching probability is most prevalent from January to April which is preceded by a period of higher annual precipitation.

Pest ecology principles and how environmental variables determine ecological processes are characterized by heuristic relationships and require intrinsic knowledge (Nestel et al., 2004). Because of this, and since our timeline data is short (from 2013 to 2020) while the bands occurrence data were not well spread over the training region, a fuzzy logic model was most appropriate (Kecman, 2001). As we expect the pest ecology to be stable from location to location, extrapolating the rules to the target areas is thus appropriate despite the data scarcity (bands data) in Turkana County. Rules based on pest ecological principles provide more ecologically stable and replicable outputs, than outputs based on machine learning modeling; these require "big data" to create probabilistic outputs from random statistical relationships (Zhou et al., 2017). This means that statistical machine learning model outputs are best suited to that specific training data only (and their statistical relationships) and do necessitate data to be accurately extrapolated over other areas under consideration (Landmann et al., 2020). This limits the transferability of the model to other areas where training data were/are not available. Conversely, the advantage of knowledge representation algorithms (i.e., fuzzy logic models) is that ecological processes can be considered using data science and knowledge about the pest's bio-ecology, and we expected that the results would not change significantly

Fig. 6. Monthly binary hatching probability for the observation period 2017 to 2020 (5-km grid data). High hatching susceptibility (red) is found towards the north and south of Turkan, and most notably in early 2020.

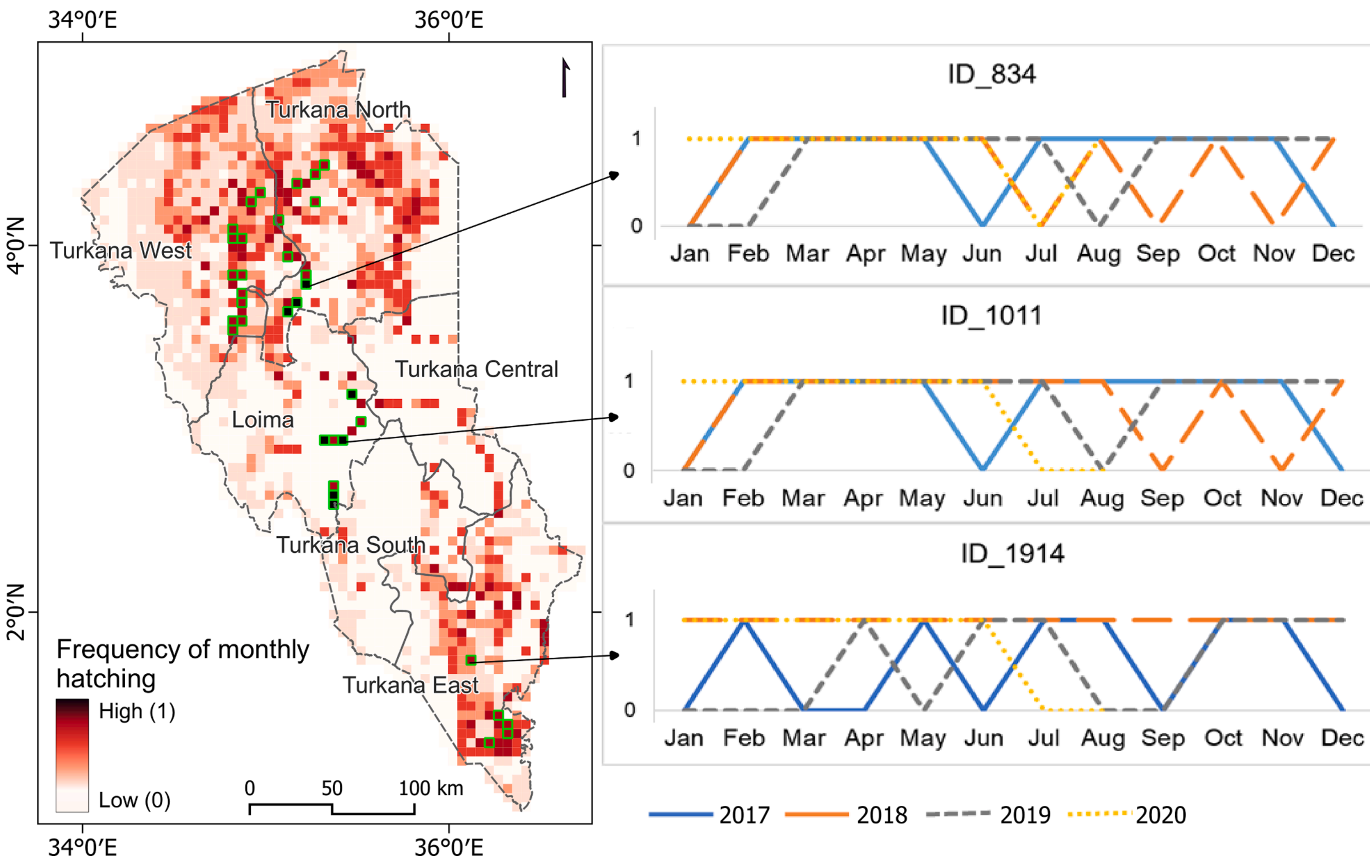

Fig. 7. Aggregated 5-km resolution "heat map" of hatching frequency suitability over Turkana in Kenya (left). Dark reddish and green outlined pixels illustrate high frequency of hatching suitability, over the priority time frame in which major outbreaks occurred (1 = predicted hatching, 0 = no hatching). Per pixel monthly hatching timelines are shown for randomly selected pixels as graphs for 2017 (continuous gray fade line), 2018 (orange long-dashed fade line), 2019 (blue short-dashed), and 2020 (red dashed, validation year). Every pixel in the frequency maps was also uniquely identified by a number, to demarcate a precise location on ground (i.e., ID_XXXX, where X stands for 1 to 2500 pixels within the gridded Turkana extent).

over space. Therefore, a fuzzy approach is more appropriate since it better mimics ecological processes compared to conventional data-driven machine learning, especially in the African context where political instability renders difficult the assess to remote areas. Our results showed that even if the based data are collected elsewhere and rules are local specific (Sudan), these can be accurately applied to other areas albeit this prerequisites similar agro-ecological conditions in both the rule development (Sudan) and the rule implementation sites (Turkana).

To instigate an early warning system for locust hatching, fuzzy rules should be developed for quasi real time and spatially explicit soil moisture, "greenness" and rainfall data (i.e., Mohanty et al., 2017).

4.2. Limitations of the study and potential improvements

Given the inherent uncertainty and fuzziness in environmental and ecological data, there are several potential improvements that could enhance the performance of specifically the Fuzzy Mamdani system in this context. Firstly, enhancing the robustness of the system against noisy and incomplete data could significantly improve its performance. This could be achieved through the integration of adaptive noise filtering techniques within the fuzzification process. Secondly, the use of advanced optimization algorithms for tuning the parameters of the membership functions could result in more accurate models. Genetic algorithms, particle swarm optimization, or gradient descent methods could be employed for this purpose.

Furthermore, boundary conditions in a Fuzzy Mamdani system play a pivotal role as they define the range and scope of the membership functions. Adjusting these boundary conditions can significantly impact the performance of the model. For instance, overly broad boundary conditions can lead to models that are too general and lack specificity. On the other hand, excessively narrow boundary conditions might result in overfitting, where the model performs well on training data but poorly on unseen data. Adjusting the boundary conditions based on the data distribution and the specific characteristics of the ecological or environmental system under study can lead to improved performance as observed in this study. The optimal range for these boundary conditions can often be found using cross-validation techniques or through sensitivity analysis depending on the phenomena under study.

Overall the proposed improvements and modifications can serve as a foundation for further studies. Future research could focus on implementing and evaluating these enhancements within a Fuzzy Mamdani system for ecological and environmental management. Experimental or simulation studies could be designed to analyze the impacts of these modifications on the model's performance. For instance, researchers could compare the performance of models with and without adaptive noise filtering techniques, or between models that use different optimization algorithms for parameter tuning. By systematically adjusting the boundary conditions, researchers could perform sensitivity analyses to determine their optimal ranges. Experimental studies could then be carried out to compare the performance of models with different boundary condition settings. These studies would not only provide valuable insights into the impacts of these improvements but could also further our understanding of how best to apply Fuzzy Mamdani systems within the field of ecology and environmental management. The results of these studies could then serve to guide future research and applications of fuzzy modeling approaches in this field.

5. Conclusions

In conclusion, this study has provided valuable insights into various ecological factors and their influence on spatio-temporal hatching dynamics of desert locusts. We showed that the juvenile stages of desert locusts in eastern Africa can be monitored, explicitly the probability of their hatching timing and location in remote and inaccessible areas, and over large tracts of land. Using data exploration, we developed a fuzzy rule set by linking desert locust occurrence data (bands data linked to hatching), to associated ecological factors (accumulated rainfall, soil moisture, temperature regimes and vegetation greenness) and their dynamics. Since bands data were readily available for Port Sudan in Sudan, we used data from this site for rule development. A true prediction accuracy of 82% could be achieved using the Sudan training data.

Through the data exploration part, high occurrence of bands (hatching) could be statistically linked to preceding rainfall and vegetation greenness responds (from 250-meter satellite data). Based on the developed rules in the training areas, the model was successfully implemented for Turkana and monthly hatching probability and overall (2017–2020) hatching frequency maps could be developed. The probability map indicated high hatching activity in early 2020; this period also corresponds to the desert locust swarming and outbreak period in Kenya. Our fuzzy logic model showed how sparse data for one site can be effectively used for rule set implementation (mapping) in other ecologically similar sites. Mapping of a subtle ecological phenomena such as desert locust hatching will help to instigate timely control and “on the ground” monitoring responses. This is a first step in operationalizing a more targeted early response to desert locust infestations for their more effective control.

Further research is needed to explore the use of other high-resolution factors such as soil moisture dynamics (when available in future), that may also influence hatching rates, as well as photoperiod and genetic variations. Ideally, in future research, a system dynamics modeling approach should be tested to comprehensively understand individual feedback between various hatching risk dimensions, including long term climate change.

Ethics approval and consent to participate

Not applicable.

CRediT authorship contribution statement

Tobias Landmann: Conceptualization, Formal analysis, Funding acquisition, Project administration, Writing – original draft, Writing – review & editing. **Komi M. Agboka:** Investigation, Data curation, Methodology, Software, Writing – review & editing. **Igor Klein:** Writing – review & editing, Funding acquisition. **Elfatih M. Abdel-Rahman:** Data curation, Writing – review & editing. **Emily Kimathi:** Supervision, Writing – review & editing. **Bester T. Mudereri:** Supervision, Validation. **Benard Malenge:** Investigation, Methodology, Visualization. **Mahgoub M. Mohamed:** Data curation, Validation. **Henri E.Z. Tonnang:** Conceptualization, Supervision.

Declaration of Competing Interest

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

Data availability

Data will be made available on request by sending an email to: tlandmann@icipe.org.

Acknowledgements

The authors gratefully acknowledge the financial support for this research by the German Aerospace Center (DLR), through the project Remote sensing application for better locust management in Kenya (award D/565/67305815). Further, we are thankful to the following organizations and agencies: Center for International Migration and Development (CIM), Fund for Human Capacity Development with Partners from International Agricultural Research (PIAF), Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH (18.2085.1–003.03); the Swedish International Development Cooperation Agency (Sida); the Swiss Agency for Development and Cooperation (SDC); the Australian centre for International Agricultural Research (ACIAR); the Federal Democratic Republic of Ethiopia; and the Government of the Republic of Kenya.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.2023.110476.

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958 to 2015. Sci. Data 5 (1), 1–12.

Alkhalil, A., Kadaoure, I., Kouadio, M., 2020. An evaluation of 20-m ESA-CCI S2 prototype LC product. Front. Sustain. Food Syst. 4, 504334.

Babuška, R., 2012. Fuzzy Modeling For Control, 12. Springer Science & Business Media.

Brader, L., DJibo, H., Faya, F.G., Ghaout, S., Lazar, M., Luzietoso, P.N., Ould-Babah, M. A., 2006. Towards a more effective response to desert locusts and their impacts on food security, livelihood and poverty. Multilateral Evaluation of the 2003–05 Desert Locust Campaign. Food and Agriculture Organisation, Rome, Italy, pp. 1–42.

Chen, C., Qian, J., Chen, X., Hu, Z., Sun, J., Wei, S., Xu, K., 2020. Geographic distribution of desert locusts in Africa, Asia and Europe using multiple sources of remote-sensing data. Remote Sens. (Basel) 12 (21), 3593.

Cressman, K., 2013a. Climate change and locusts in the WANA Region. Clim. Change and Food Security in West Asia and North Africa 131–143.

Cressman, K., 2021. Technical innovations in global early warning in support of desert locust area-wide management. Area-wide Integrated Pest Management. CRC Press, pp. 765–778.

Cressman, K., 2013b. Role of remote sensing in desert locust early warning. J. Appl. Remote Sens. 7 (1), 075098-075098.

Cressman, K., Hodson, D., 2009. Surveillance, information sharing and early warning systems for transboundary plant pests diseases: the FAO experience. Arab J. Plant Protection 27 (2), 226–232.

Dinku, T., Ceccato, P., Cressman, K., Connor, S.J., 2010. Evaluating detection skills of satellite rainfall estimates over desert locust recession regions. J. Appl. Meteorol. Clim. 49, 1322–1332.

Dong, Y., Zhao, L., Huang, W., 2023. Monitoring of Desert Locust in Africa and Asia. Springer Nature.

El Gamri, T., Saeed, A.B., Abdalla, A., 2009. Rainfall of the Sudan: characteristics and prediction. J. Faculty of Arts 27. University of Khartoum.

Gay, P.E., Lecoq, M., Piou, C., 2020. The limitations of locust preventive management faced with spatial uncertainty: exploration with a multi-agent model. Pest Manag. Sci. 76 (3), 1094–1102.

Gómez, D., Salvador, P., Sanz, J., Casanova, C., Taratiel, D., Casanova, J.L., 2019. Desert locust detection using Earth observation satellite data in Mauritania. J. Arid Environ. 164, 29–37.

Hagras, H., Wagner, C., 2012. Towards the wide spread use of type-2 fuzzy logic systems in real world applications. IEEE Comput. Intell. Mag. 7 (3), 14–24.

Halwagy, R., 1961. The vegetation of the semi-desert north east of Khartoum. Sudan. Oikos 12 (1), 87–110.

Jain, S., Tiwari, V., Thapa, A., Mangla, R., Jaiswal, R.K., Kumar, V., Pandey, K., 2022. Performance evaluation of Google earth engine based precipitation datasets under different climatic zones over India. Remote Sensing in Earth Syst. Sci. 1–14.

Kecman, V., 2001. Learning and Soft computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models. MIT press.

Kimathi, E., Tonnang, H.E., Subramanian, S., Cressman, K., Abdel-Rahman, E.M., Tesfayohannes, M., Kelemu, S., 2020. Prediction of breeding regions for the desert locust Schistocerca gregaria in East Africa. Sci. Rep. 10 (1), 11937.

Klein, I., van der Woude, S., Schwarzenbacher, F., Muratova, N., Slagter, B., Malakhov, D., Kuenzer, C., 2022. Predicting suitable breeding areas for different locust species–A multi-scale approach accounting for environmental conditions and current land cover situation. Int. J. Appl. Earth Observation and Geoinf. 107, 102672.

Kinyua, I., 2021. National agriculture rural inclusive growth project (NARIGP) climate risk profiles validation workshop presentation.

Landmann, T., Dubovyk, O., Ghazaryan, G., Kimani, J., Abdel-Rahman, E.M., 2020. Wide-area invasive species propagation mapping is possible using phenometric trends. ISPRS J. Photogrammetry and Remote Sensing 159, 1–12.
Maeno, K.O., Piou, C., Ghaout, S., 2020. The desert locust, Schistocerca gregaria, plastically manipulates egg size by regulating both egg numbers and production rate according to population density. J. Insect Physiol. 122, 104020.
MoALFC, 2021. Kenya County Climate Risk Profile: Turkana County. Ministry of Agriculture, Livestock, Fisheries and Co-operatives (MoALFC), NairobiKenya, p. 32.
Mohanty, B.P., Cosh, M.H., Lakshmi, V., Montzka, C., 2017. Soil moisture remote sensing: state-of-the-science. Vadose Zone J. 16, 1–9.
Mongare, R., Abdel-Rahman, E.M., Mudereri, B.T., Kimathi, E., Onywere, S., Tonnang, H. E., 2023. Desert Locust (Schistocerca gregaria) invasion risk and vegetation damage in a key upsurge area. Earth 4 (2), 187–208.
Meynard, C.N., Lecoq, M., Chapuis, M.P., Piou, C., 2020. On the relative role of climate change and management in the current desert locust outbreak in East Africa. Glob. Chang. Biol. 26 (7), 3753–3755.
Nestel, D., Carvalho, J., Nemny-Lavy, E., 2004. The spatial dimension in the ecology of insect pests and its relevance to pest management. Insect Pest Manag.: Field and Protected Crops 45–63.
Nevo, D., 1996. The desert locust, Schistocerca gregaria, and its control in the land of Israel and the Near East in antiquity, with some reflections on its appearance in Israel in modern times. Phytoparasitica 24, 7–32.
Nishide, Y., Suzuki, T., Tanaka, S., 2017. Synchrony in the hatching of eggs in the desert locust Schistocerca gregaria (Orthoptera: acrididae): egg condition influences hatching time in the laboratory and under simulated field temperatures. Appl. Entomol. Zool. (Jpn.) 52, 599–604.
Pekel, J.F., Ceccato, P., Vancutsem, C., Cressman, K., Vanbogaert, E., Defourny, P., 2010. Development and application of multi-temporal colorimetric transformation to monitor vegetation in the desert locust habitat. IEEE J. Selected Topics in Appl. Earth Observations and Remote Sensing 4 (2), 318–326.
Pflüger, H.J., Bräunig, P., 2021. One hundred years of phase polymorphism research in locusts. J. Comparative Physiol. A 207 (3), 321–326.
Rahim, N.F., Othman, M., Sokkalingam, R., Kadir, E.A., 2018. Forecasting crude palm oil prices using fuzzy rule-based time series method. IEEE Access 6, 32216–32224.
Ricotta, C., 2000. From theoretical ecology to statistical physics and back: self-similar landscape metrics as a synthesis of ecological diversity and geometrical complexity. Ecol. Modell. 125 (2–3), 245–253.
Rhodes, K., Sagan, V., 2021. Integrating remote sensing and machine learning for regional-scale habitat mapping: advances and future challenges for desert locust monitoring. IEEE Geosci. Remote Sensing Magazine 10 (1), 289–319.
Rowell, Hugh, Hemp, Claudia., 2017. Jago's Grasshoppers of East and North East Africa, 2.
Salih, A.A., Baraibar, M., Mwangi, K.K., Artan, G., 2020. Climate change and locust outbreak in East Africa. Nat Clim Chang 10 (7), 584–585.
Spinage, C.A., 2012. Locusts the forgotten plague Part I: locusts and their ecology. Afr. Ecol.: Benchmarks and Historical Perspectives 481–532.
Skaf, R., Popov, G.B., Roffey, J., 1990. The desert locust: an international challenge. Philosophical Trans. Royal Society of London. B, Biol. Sci. 328, 525–538, 1251.
Sun, R., Huang, W., Dong, Y., Zhao, L., Zhang, B., Ma, H., Li, X., 2022. Dynamic forecast of desert locust presence using machine learning with a multivariate time lag sliding window technique. Remote Sens. (Basel) 14 (3), 747.
Symmons, P.M., 1992. A model to predict the incubation and nymphal development periods of the desert locust. Schistocerca gregaria (Orthoptera: Acrididae) 1. Bullet. Entomol. Res. 82 (4), 517–520.
Symmons, P.M., Cressman, K., 2001. Desert Locust guidelines: Biology and Behaviour. FAO, Rome, pp. 1–42.
Tratalos, J.A., 2001. Doctoral dissertation. University of Greenwich.
Van Huis, A., 2007. Strategies to control the desert locust Schistocerca gregaria. Area-Wide Control of Insect Pests: From Research to Field Implementation. Springer, Netherlands, pp. 285–296.
Topaz, C.M., D'Orsogna, M.R., Edelstein-Keshet, L., & Bernoff, A.J. , 2012. Locust dynamics: behavioral phase change and swarming.
Wang, L., Zhuo, W., Pei, Z., Tong, X., Han, W., Fang, S., 2021. Using long-term earth observation data to reveal the factors contributing to the early 2020 desert locust upsurge and the resulting vegetation loss. Remote Sens. (Basel) 13 (4), 680.
Woodman, J.D., 2015. Surviving a flood: effects of inundation period, temperature and embryonic development stage in locust eggs. Bull. Entomol. Res. 105 (4), 441–447.
Zadeh, L.A., 1965. Fuzzy sets. Inf. Control 8, 338–353.
Zhou, L., Pan, S., Wang, J., Vasilakos, A.V., 2017. Machine learning on big data: opportunities and challenges. Neurocomputing 237, 350–361.