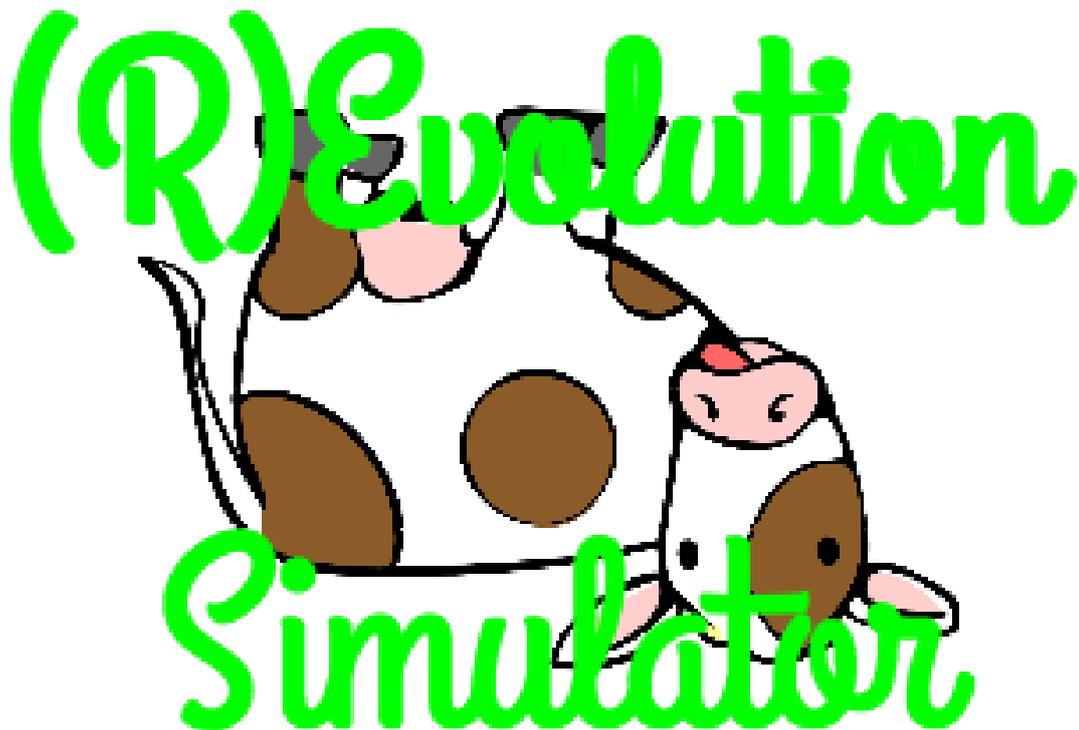


(R)Evolution Simulator

Federico Malnati, Corso Sebastopoli 306, Torino (TO)
Matteo Palmieri, Strada Caldano 59/61, Caselle Torinese (TO)
Alessandro Sosso, Corso Vittorio Emanuele II 20, Torino (TO)



Liceo Scientifico Galileo Ferraris
Docenti di riferimento: Andrea Doveri, Annalisa Gratteri



Indice

1	Introduzione	3
2	Esperimenti	3
2.1	Descrizione della progettazione	3
2.1.1	I chunk	3
2.1.2	Le creature	4
2.1.3	Interfaccia grafica	5
2.2	Materiale e procedura	5
2.3	Le variabili misurate	5
2.3.1	Variabili delle creature	6
2.3.2	Variabili dei chunk	6
3	Discussione	6
3.1	Caratteristiche genetiche e selezione naturale	6
3.1.1	Caratteristiche genetiche numeriche: agility, bigness, speed, fertility	6
3.1.2	Caratteristiche genetiche mendeliane: resistenza alla temperatura	7
3.1.3	Caratteristiche genetiche di controllo: equilibrio di Hardy-Weinberg e Genetic Drift	7
3.2	Distribuzione delle creature sul territorio secondo il cibo	7
3.3	Andamento demografico	8
3.3.1	Ciclo di nascite e morti	8
4	Conclusioni	8
5	Ringraziamenti	8
A	Immagini e grafici	9

1 Introduzione

In tutti i campi della ricerca é in crescita esponenziale l'utilizzo di processori, mezzi di calcolo potenti che permettono di simulare ambienti naturali, in cui é possibile osservare piú facilmente fenomeni particolari al variare di parametri. Mediante i computer é possibile simulare voli spaziali, moti astronomici, ecosistemi, anche sotto forma di videogioco. Il nostro progetto ha previsto la creazione di un programma (scritto in linguaggio Python) che simula all'interno di un ambiente virtuale i processi di evoluzione di una specie. Ogni creatura si comporta come nel mondo reale, ricercando l'ambiente piú favorevole per vivere, nutrendosi, riproducendosi e morendo se si trova in un ambiente troppo sfavorevole. Analizzando i dati si osserverá quindi come essi rispecchiano ciò che accade realmente in natura, come avviene la selezione naturale e come in un certo ambiente sopravvivono gli animali piú adatti a vivere lí.

Attraverso questo progetto sará possibile osservare i meccanismi studiati dalla genetica delle popolazioni, la trasmissione dei caratteri genetici su larga scala e la selezione naturale degli individui della specie. Verranno analizzati i vari caratteri (alcuni numerici e alcuni di tipo mendeliano) e il loro comportamento a confronto con l'equilibrio di Hardy-Weinberg e con la deriva genetica.

Le creature da noi considerate possono ben rappresentare una qualunque specie di erbivori, i quali, pur avendo un comportamento piuttosto semplice, daranno la possibilitá di analizzare l'andamento sinusoidale della popolazione e la distribuzione differenziata sul territorio. L'interfaccia grafica, realizzata in Python, permette di visualizzare il tutto in modo diretto ed efficace.

2 Esperimenti

2.1 Descrizione della progettazione

Il territorio virtuale nel quale le creature vivono ha forma rettangolare ed é suddiviso in piccole porzioni quadrate denominate *chunk*, mentre il tempo é misurato in unitá chiamate *tick*. Ad ogni *tick* vengono aggiornate le caratteristiche di tutte le creature vive e di tutti i *chunk*.

2.1.1 I chunk

Ogni *chunk* é caratterizzato da: una posizione nel mondo; una quantitá massima di erba (tra 0 e 100); un coefficiente di crescita dell'erba, direttamente proporzionale alla quantitá massima di erba nel *chunk*; una temperatura, che varia tra -100 (freddo) e 100 (caldo), costante nel tempo. Sia la temperatura che la quantitá iniziale dell'erba sono distribuite nel mondo tramite il Rumore Semplice di Perlin, una funzione che assegna valori vicini a variabili vicini (un "randomizzazione" piú omogenea, come si puó osservare nella Figura 1 in appendice).

Nella Figura 2 in appendice si possono osservare due immagini rappresentanti il mondo in assenza delle creature, nella prima dal punto di vista termico (le zone sono piú calde tendono al colore rosso, mentre quelle piú fredde tendono al colore blu), mentre nella seconda dal punto di vista della distribuzione della quantitá massima di cibo (le zone verde scure hanno poco cibo, mentre le zone verde chiaro hanno tanto cibo). Come si puó osservare, la temperatura varia gradualmente tra un *chunk* e un altro vicino, grazie appunto al Rumore Semplice di Perlin. Lo stesso discorso vale per la distribuzione della quantitá cibo. Tuttavia, ad un certo istante, la distribuzione del cibo presente in un chunk puó variare, in base a dove le creature hanno mangiato. Nel caso di un'eventuale morte di tutte le creature, il cibo ricrescerebbe tornando alla situazione in Figura 2.

2.1.2 Le creature

Le creature, ad ogni *tick*, cercano in una zona quadrata di 7×7 *chunk* attorno a loro, dove conviene andare per mangiare (in base alla quantità di cibo di ogni *chunk* considerato e alla distanza di esso dal punto in cui si trovano). Una volta raggiunta la destinazione le creature mangiano, aumentando la loro energia vitale e diminuendo la quantità di cibo presente nel *chunk*. Ad ogni *tick* le creature perdono una quantità di energia proporzionale a quella che possiedono.

La riproduzione avviene in modo casuale tra individui nella zona quadrata 7×7 *chunk* di sesso opposto. In particolare se due individui hanno abbastanza energia, si riproducono generando una nuova creatura che avrà caratteristiche ereditate dai genitori e mediate da mutazioni. Durante la riproduzione, i genitori perdono una certa quantità di energia che cedono al figlio. [Diagramma di flusso per il comportamento delle creature, Figura 3].

La morte delle creature può essere causata da tre fattori:

- Vecchiaia (*old age*): la vita massima viene calcolata alla nascita in base a un numero casuale con andamento gaussiano di media 1000 *tick* e deviazione standard 200 *tick*;
- Temperatura (*temperature*): ogni creatura ha un proprio genotipo riguardante la sua capacità di resistere alla temperatura (assimilabile alla presenza di pelo). In particolare esistono tre diversi alleli: un allele dominante (*N*), senza particolari caratteristiche; due alleli recessivi (*l* e *c*), rispettivamente adatti a vivere in clima freddo e caldo. I possibili genotipi sono quindi omozigoti (*NN*, *ll*, *cc*) o eterozigoti (*Nl*, *Nc*, *lc*). I genotipi *NN*, *Nl*, *Nc*, *lc* corrispondono ad un fenotipo *N*, adatto a vivere in zone temperate (di colore bianco in Figura 2) e muore più facilmente muovendosi verso un clima estremo (caldo o freddo). I genotipi *ll* e *cc* corrispondono rispettivamente ai fenotipi *l* e *c*. Il primo è più adatto a vivere in climi freddi (di colore blu in Figura 1) e ha una probabilità di morire più facilmente in climi caldi. Il secondo, al contrario, vive più facilmente in climi caldi (di colore rosso in Figura 1). A ogni *tick* la creatura ha una certa probabilità di morire in base alla temperatura del *chunk* dove si trova e al suo genotipo [Figura 3]; La probabilità di morte per temperatura, sull'asse *y* in Figura 4, ha come valore massimo 0.02 poiché viene calcolata ad ogni *tick*. Un valore massimo corrispondente a 1 porterebbe ad un numero di morti troppo elevato e non realistico.
- Fame (*starvation*): se una creatura non riesce a trovare abbastanza cibo entro un certo numero di cicli, l'energia scende sotto un valore minimo e la creatura muore

Le creature sono caratterizzate da diversi geni, che sono trasmessi nella riproduzione. I geni possono essere di tre tipi distinti:

1. I geni numerici sono caratterizzati da un numero, che ne esprime anche il fenotipo. Nella riproduzione, i geni numerici del figlio vengono calcolati uno per uno a partire dal gene numerico di uno dei due genitori, che viene moltiplicato per un certo numero casuale con andamento gaussiano di media 1. I geni numerici sono:
 - *agility*, l'attitudine a muoversi della creatura;
 - *bigness*, le dimensioni della creatura;
 - *fertility*, esprime il numero di *tick* minimi che devono passare tra una riproduzione e l'altra, ma è inversamente proporzionale all'energia minima necessaria per potersi riprodurre;
 - *num_control*, un gene numerico utilizzato come gene di controllo (non ha effetti fenotipici).

2. I geni mendeliani sono caratterizzati da un genotipo formato da una coppia di alleli, e da un fenotipo, calcolato secondo le leggi della genetica a partire dal genotipo. Nella riproduzione i due alleli di ogni gene mendeliano vengono trasmessi selezionandone uno da ciascun genitore. I geni mendeliani sono:
 - *temp_resist*, la capacità della creatura di resistere ad una certa temperatura; gli alleli sono *c*, *l* (recessivi) ed *N* (dominante);
 - *mndl_control*, un gene mendeliano utilizzato come gene di controllo (non ha effetti fenotipici); gli alleli sono *a* (recessivo) ed *A* (dominante).
3. I geni secondari sono geni numerici non più trasmessi nella riproduzione, ma calcolati a partire da altri geni numerici della creatura. I geni secondari sono:
 - *speed*, definita come $2 \cdot \text{agility}/\text{bigness}$.

2.1.3 Interfaccia grafica

Nell'interfaccia grafica, i chunk sono rappresentati come quadratini (visibili in Figura 2), mentre le creature come cerchi. La dimensione dei cerchi può rappresentare aspetti diversi della creatura, impostabili nel programma. Nelle immagini in Figura 5 le creature di dimensioni maggiori hanno un'energia maggiore in quell'istante perché hanno trovato una maggiore quantità di cibo nei *tick* precedenti. Anche il colore delle creature può rappresentare aspetti diversi, come osservabile in Figura 5:

- l'abbinamento azzurro/giallo rappresenta la distinzione dei due sessi;
- l'abbinamento rosso/grigio-bianco/blu rappresenta il fenotipo delle creature in relazione alla loro capacità di resistere alla temperatura. In particolare, il rosso un fenotipo *c* (preferenza per il caldo), il blu un fenotipo *l* (preferenza per il freddo), il grigio un fenotipo *N* eterozigota e bianco un fenotipo *N* omozigota (preferenza per climi temperati).

2.2 Materiale e procedura

Il programma è stato scritto inizialmente con Processing estensione Python per poter rappresentare facilmente l'ambiente dal punto di vista grafico e ottenere così un riscontro immediato per le eventuali modifiche al codice. È stato quindi riadattato e portato nel linguaggio Python 3.6 per una maggiore flessibilità. Abbiamo realizzato il simulatore in più versioni, implementando in ciascuna le varie dinamiche di vita (temperatura, morte per vecchiaia, riproduzione ecc.) e variando di volta in volta i coefficienti per bilanciare il sistema.

Ottenuta la versione definitiva, abbiamo raccolto i dati utilizzando simulazioni lunghe 10000 *tick*, in un territorio di grandezza 60×45 *chunk*. Nella nostra visualizzazione grafica il *chunk* ha dimensioni 10×10 pixel. Durante la simulazione le informazioni e le variabili di tutte le creature vengono salvate in diversi file da noi analizzati.

2.3 Le variabili misurate

Verranno osservati tutti i caratteri presenti in ciascun individuo (la velocità, la grandezza, la lunghezza del pelo ecc.) e si osserverà la loro variazione con il passare delle generazioni e all'interno di ambienti con caratteristiche diverse.

2.3.1 Variabili delle creature

- *ID*, un numero progressivo che la identifica
- *birth_tick*, il *tick* di nascita (per le creature già presenti sul territorio alla creazione del mondo, è un numero negativo casuale)
- *parents_ID*, cioè l'ID dei genitori
- *sex*, che è 0 o 1
- *death_tick*, il *tick* di morte
- *death_cause*, la causa della morte
- *genes*, i geni che caratterizzano la creatura.

2.3.2 Variabili dei chunk

- *foodmax*, cioè la quantità massima di cibo che può essere presente (da 0 a 100)
- *growth_coef*, coefficiente di crescita dell'erba che è direttamente proporzionale a *foodmax*
- *food*, variabile nella quale viene salvata in ogni ciclo la quantità di erba presente sul territorio
- *temperature*, cioè la temperatura del chunk (varia da -100 a 100)

3 Discussione

3.1 Caratteristiche genetiche e selezione naturale

3.1.1 Caratteristiche genetiche numeriche: *agility*, *bigness*, *speed*, *fertility*

Si è potuto osservare che col passare dei *tick* le creature sono sottoposte a selezione naturale, così che le loro caratteristiche genetiche si evolvono nel tempo (Figura 6): difatti tutti i geni numerici tendono a convergere ad un determinato valore, e la distribuzione percentile dei valori si riduce.

Dal punto di vista evolutivo, le creature devono migliorare dal punto di vista della velocità (quelle più veloci riescono a mangiare prima il cibo, anche a maggiori distanze) e, contemporaneamente, devono diventare più grandi, per mangiare una maggiore quantità di cibo ogni volta. Dato che le creature più veloci hanno una maggiore capacità di sopravvivenza, il grafico che rappresenta *speed* mostra un netto aumento medio con il passare del tempo.

Al contrario la caratteristica *bigness* mostra una diminuzione verso l'inizio della simulazione. Questo è spiegabile mediante il fatto che le creature più grandi sono, per come è stata definita *speed* ($speed = agility/bigness$), più lente, e quindi svantaggiate nella situazione iniziale. Si osserva, però, che *bigness* non scende sotto un certo minimo in quanto le creature più grandi possono mangiare anche una maggiore quantità di cibo dal territorio dove si trovano.

Infine la caratteristica *fertility* aumenta con il passare del tempo, nonostante *fertility* rappresenti il minimo numero di *tick* tra una riproduzione e l'altra. Questo è dovuto al fatto che le creature con maggiori valori di *fertility* necessitano di meno energia per potersi riprodurre, in quanto quest'ultima è inversamente proporzionale al gene stesso.

3.1.2 Caratteristiche genetiche mendeliane: resistenza alla temperatura

Abbiamo suddiviso il territorio in 8 classi di frequenza climatiche e abbiamo osservato quante creature sono presenti in ogni classe, facendo un grafico per ogni fenotipo relativo alla resistenza alla temperatura (c , l e N) [Figura 7]. Le fasce sono colorate con 8 colori che variano dal rosso (clima caldo) al blu (clima freddo).

Si è potuto osservare che le creature con fenotipo c , cioè adatte a sopportare il caldo, si disponevano nei *chunk* con temperatura maggiore, a differenza di quelle con fenotipo l , che si disponevano nei *chunk* con temperatura minore. Infine quelli con fenotipo N e n si disponevano nelle zone meno estreme, con una temperatura intermedia. Nella Figura 8 si può osservare il territorio dal punto di vista climatico e la disposizione delle creature colorate in base al loro fenotipo della temperatura.

La separazione in base alla temperatura tra creature si può considerare, nel caso più estremo, come una speciazione allopatrica, in quanto, sebbene teoricamente in grado di riprodursi, due creature con fenotipo c ed l non potranno incontrarsi, poiché una delle due morirebbe nel caso giungesse nel territorio adatto all'altra.

3.1.3 Caratteristiche genetiche di controllo: equilibrio di Hardy-Weinberg e Genetic Drift

Discorso diverso vale per i due geni che non hanno effetti fenotipici e hanno quindi la funzione di controllo.

Il gene numerico *num_control*, come atteso, non ha un comportamento univoco (sempre crescente, decrescente o costante), ma varia da simulazione a simulazione [Figura 9]. Inoltre non si può osservare la stessa riduzione della variazione della distribuzione percentile come per i geni numerici.

Per il gene di controllo mendeliano *mndl_control*, invece, si nota che la distribuzione di esso sul territorio è omogenea al variare della temperatura, sia per il fenotipo A dominante, sia per quello a recessivo, come si può osservare dalla Figura 10. Inoltre abbiamo controllato se esso rispettasse l'equilibrio di Hardy-Weinberg. Le condizioni per questo equilibrio sono: assenza di mutazioni, accoppiamenti casuali, popolazione di grandi dimensioni, assenza di flusso genico e mancanza di selezione naturale in base a quel gene. Abbiamo osservato che il carattere considerato non rispecchia questo equilibrio, in quanto la popolazione è di dimensioni abbastanza ridotte (circa 150 creature nel mondo in un certo istante) ed è soggetta ad un andamento incostante dovuto al casuale consumo eccessivo di cibo in alcune aree nell'arco di alcuni *tick*. Possiamo quindi affermare che questo carattere è soggetto a deriva genetica.

3.2 Distribuzione delle creature sul territorio secondo il cibo

Come fatto per la temperatura, abbiamo diviso il territorio in 8 classi in base alla quantità di cibo massimo in ogni *chunk*. Abbiamo quindi rappresentato il numero di creature nelle varie classi con il passare dei *tick*. Le fasce sono rappresentate con una tonalità di verde diversa: quelle più scure hanno meno cibo, mentre quelle più chiare ne hanno di più.

Dal grafico in Figura 11 si può osservare come le creature si trovino principalmente nelle fasce più chiare, ovvero con una maggiore quantità di cibo. Dalla Figura 12 si può invece osservare la distribuzione delle creature sul territorio in un certo istante.

3.3 Andamento demografico

3.3.1 Ciclo di nascite e morti

La popolazione delle creature segue un andamento sinusoidale con un incremento della popolazione di circa 100 individui. (Figura 13)

Questo comportamento è dovuto a un meccanismo di autoregolazione e si verifica quando, essendoci un numero troppo basso di creature rispetto a quello che l'ecosistema può sopportare, queste si riproducono incrementando velocemente il loro numero. A questo punto, però, sono troppe e, non essendoci abbastanza cibo, avviene un forte calo demografico, quindi il ciclo ricomincia. Si osserva quindi che a un picco di nascite segue sempre un picco di morti a cui segue un calo delle nascite e un successivo calo delle morti, per riprendere infine da capo il ciclo.

4 Conclusioni

Abbiamo osservato che il nostro sistema si comporta come un ecosistema naturale, in cui le caratteristiche genetiche tramite il ricambio e la mutazione genica sono sottoposte a selezione naturale; in cui è presente una selezione sul territorio basata sulle sue caratteristiche naturali; in cui sussiste un ciclo di nascite e morti paragonabile a quello reale. Abbiamo quindi dimostrato che con un sistema basato su leggi semplici può fornire una simulazione di sistemi complessi come quelli naturali.

Il nostro progetto permette quindi di analizzare un numero di casi maggiore rispetto a quello osservabile in natura, in un arco di tempo decisamente ridotto. Una simulazione da 20000 tick e 60×45 *chunk* (elaborata dal computer in circa 30 minuti), permette di osservare circa 16000 creature, cosa difficile da fare in natura per la maggior parte delle specie, se non nell'arco di numerose generazioni.

Il programma, utile anche a fine didattico per visualizzare concetti apparentemente astratti, ha una notevole flessibilità e permette numerosi ampliamenti e sviluppi: l'introduzione di una specie di predatori può mostrare il ciclo preda/predatore delle equazioni di Lotka-Volterra; l'introduzione di una malattia contagiosa può mettere in luce le dinamiche di diffusione di un'epidemia e permettere di trovare il minimo numero di individui immuni al contagio affinché questo non si diffonda.

5 Ringraziamenti

Si ringrazia il Professor Riccardi, docente del corso di Python, per i preziosi suggerimenti e gli stimoli creativi circa lo sviluppo del codice.

Riferimenti bibliografici

- [1] D. Shiffman: *The Nature of Code, simulating natural systems with Processing*
- [2] David Sadava, H. Craig Heller, Gordon H. Orians, William K. Purves, David M. Hillis: *Biologia.blu, le basi molecolari della vita e dell'evoluzione*, Zanichelli
- [3] <http://py.processing.org/reference/>
- [4] Wikipedia, L'enciclopedia libera, <http://it.wikipedia.org/>
- [5] <http://evolution.gs.washington.edu/pgbook/pgbook.pdf>, Joseph Felsenstein, "Theoretical evolutionary genetics"

A Immagini e grafici

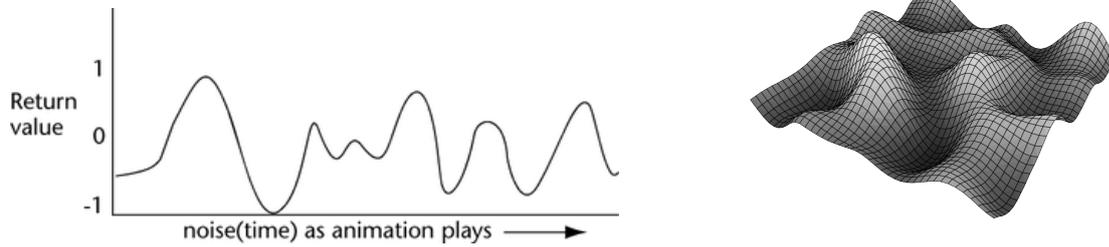


Figura 1: Esempio di grafico della funzione Perlin Noise a una variabile (a sinistra) e a 2 variabili (sulla destra)

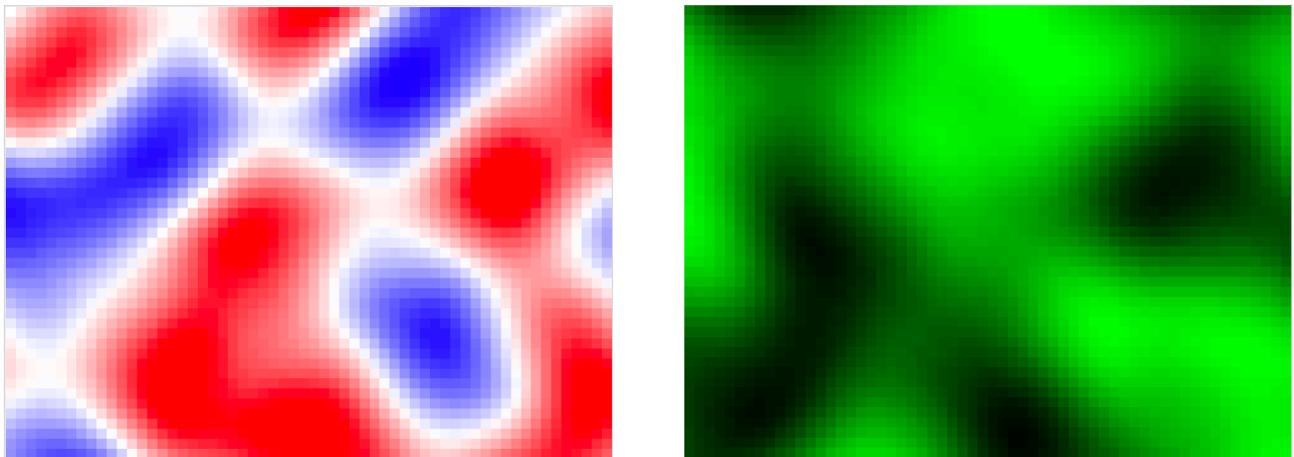


Figura 2: Distribuzione della temperatura e del cibo: nell'immagine a sinistra il territorio é colorato in base alla temperatura di ogni *chunk* (rosso verso il caldo e blu verso il freddo). Sulla destra si vede il cibo: le zone verde scuro hanno una minore quantità di cibo, mentre quelle verde chiaro una quantità maggiore. Entrambe le figure sono state generate con la funzione Perlin Noise in Figura 1.

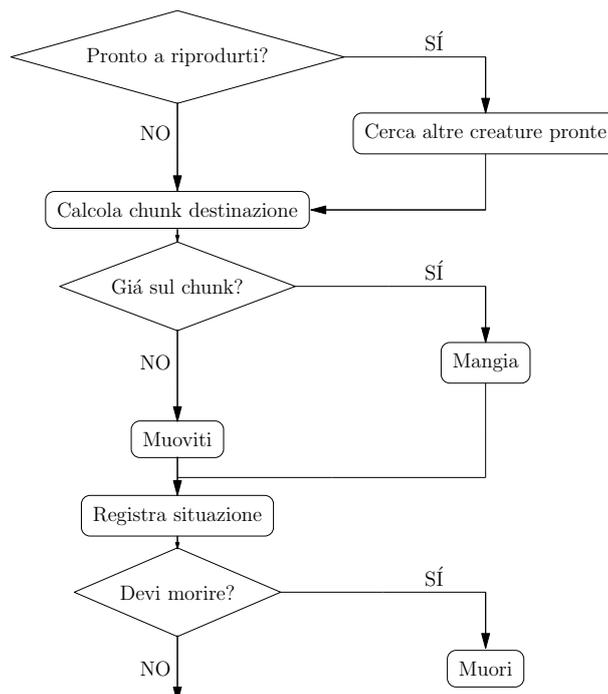


Figura 3: Schema del comportamento delle creature

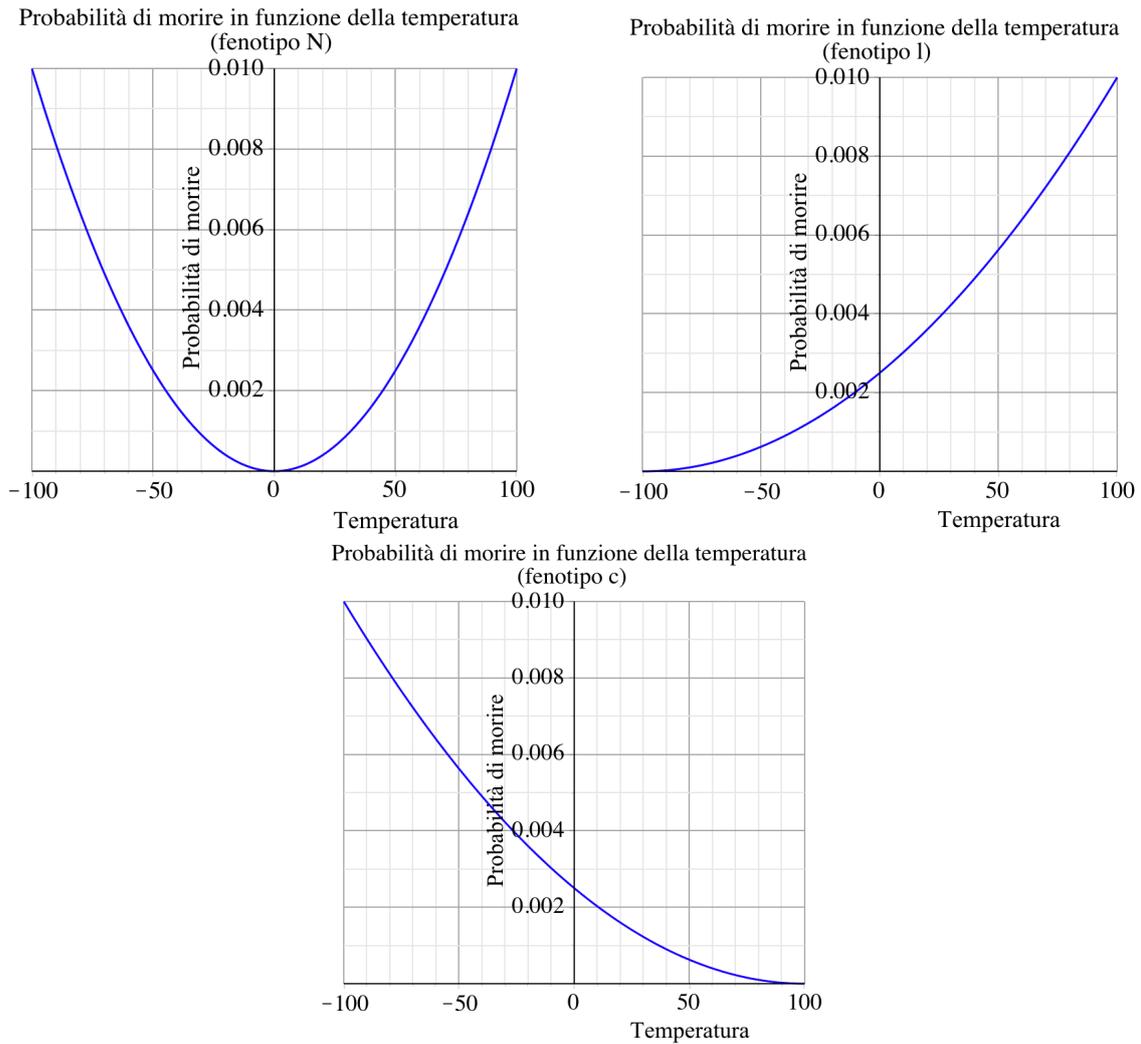


Figura 4: Probabilità di morte per la temperatura (sull'asse y), in relazione alla temperatura (sull'asse delle ascisse) in base ai 3 possibili fenotipi (N , l e c).

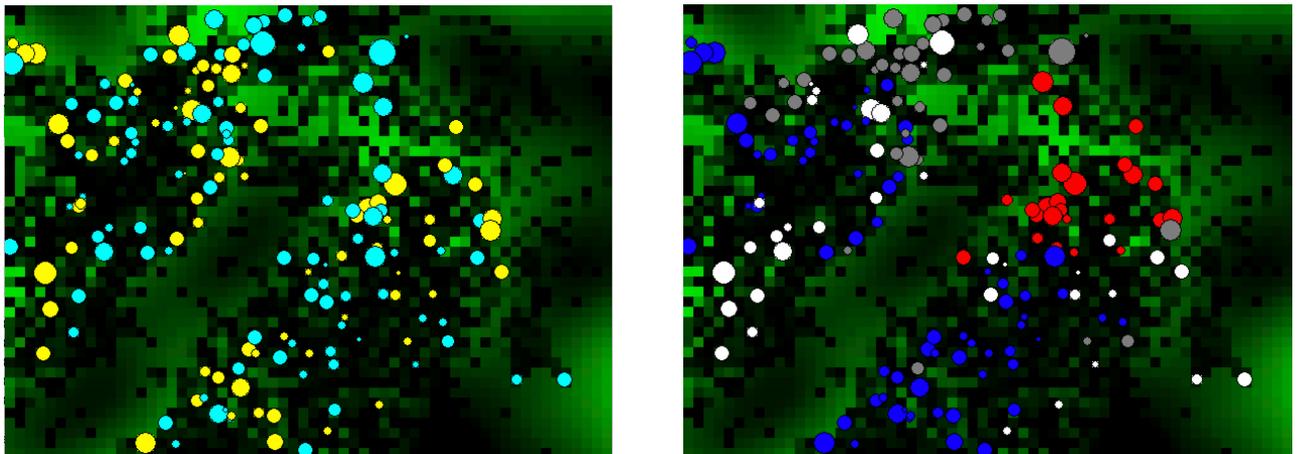


Figura 5: In figura é rappresentato lo stesso istante. A destra le creature hanno un colore che rappresenta la loro capacità di resistere alla temperatura (blu per il freddo, rosso per il caldo, bianco e grigio per climi intermedi). A sinistra, invece, sono rappresentate le creature il base al loro sesso (gialle o azzurre).

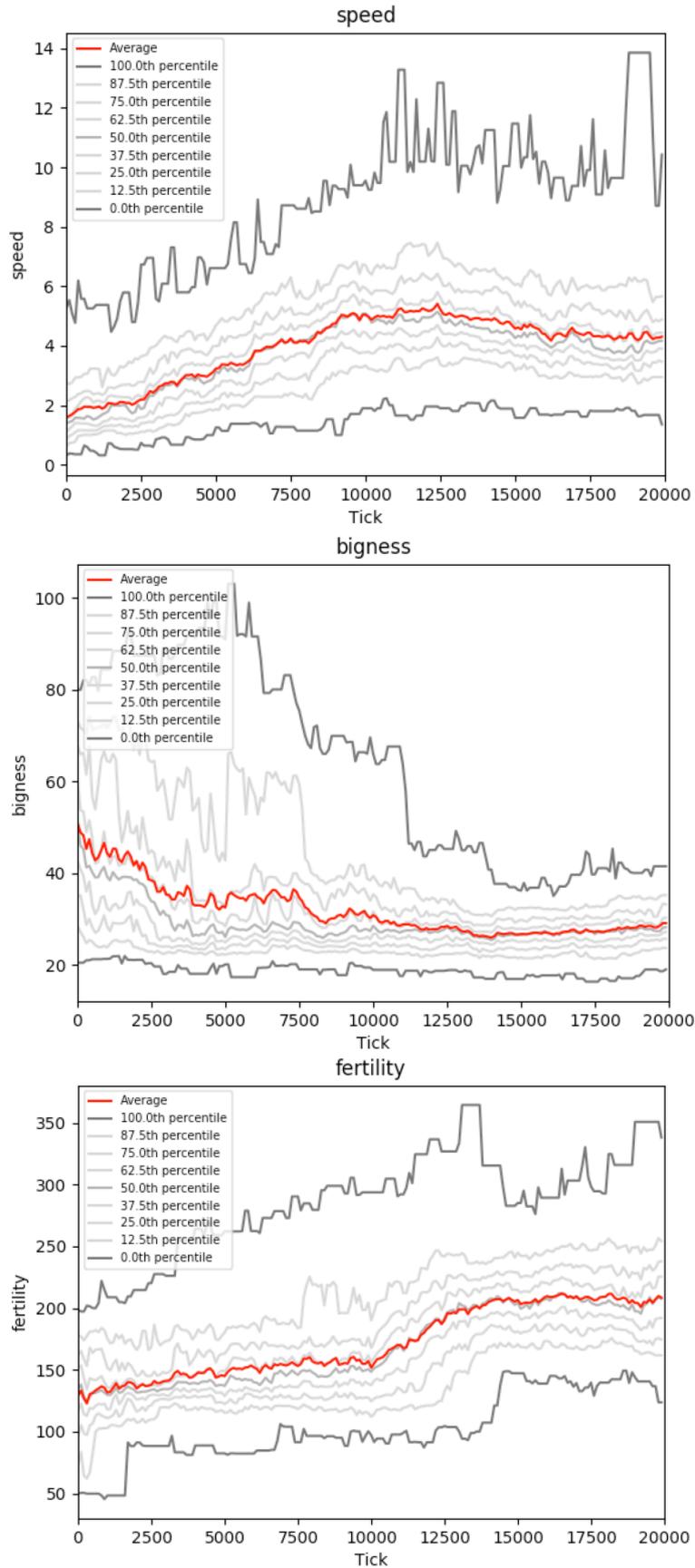


Figura 6: Grafici dell'evoluzione delle caratteristiche genetiche delle creature (*speed*, *bigness* e *fertility*).

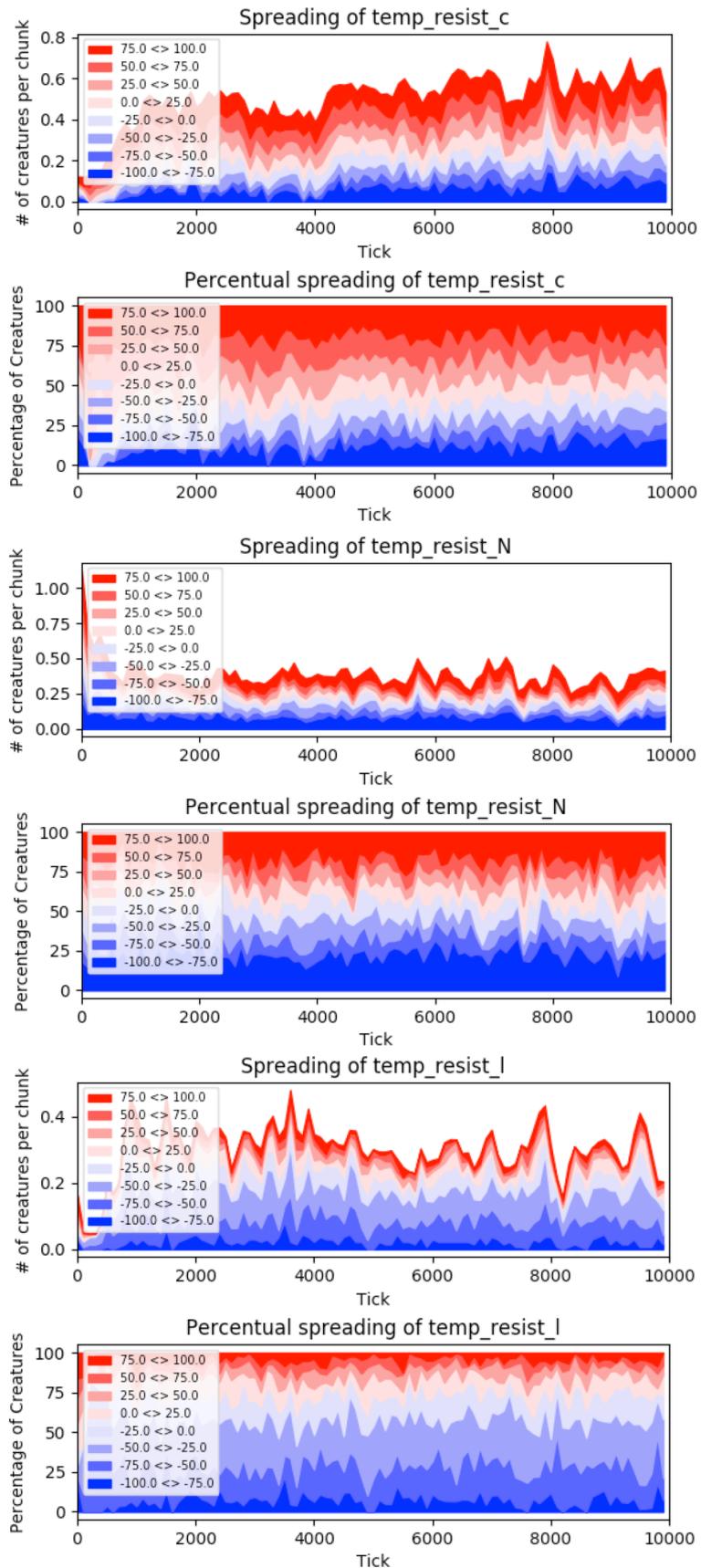


Figura 7: Grafici della distribuzione delle creature in base alla temperatura. I sei grafici rappresentano i fenotipi *c*, *l* e *N* del gene *temp_resist* (dal punto di vista del numero assoluto delle creature e della percentuale)

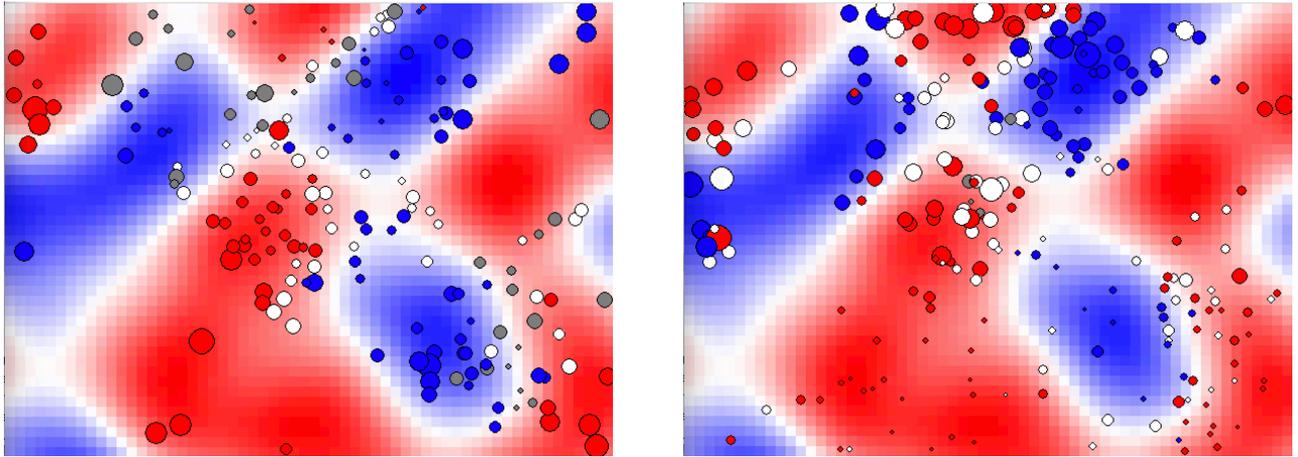


Figura 8: Due istanti di una simulazione in cui si può osservare la distribuzione delle creature sul territorio in base al loro fenotipo e alla temperatura.

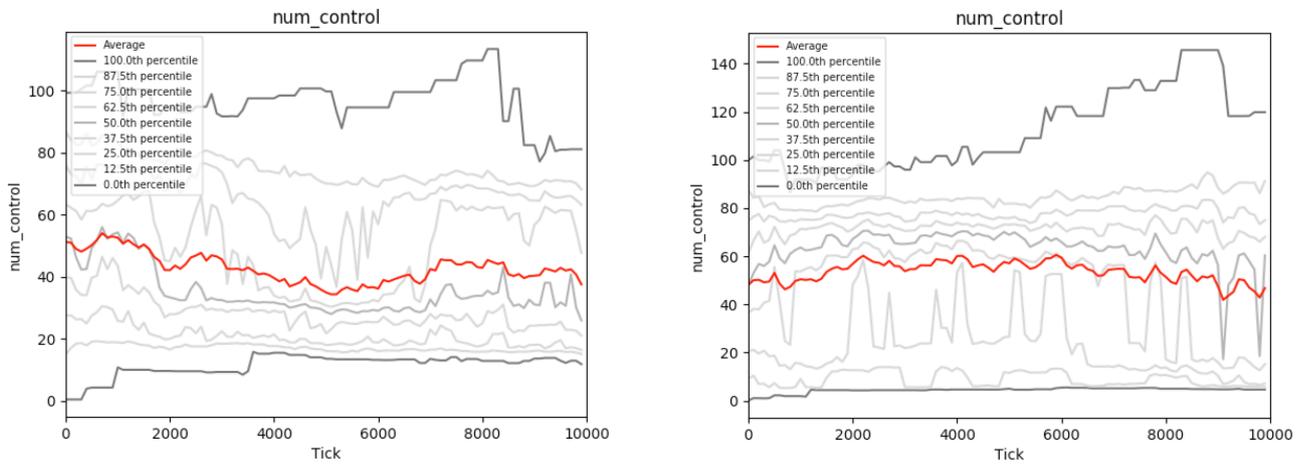
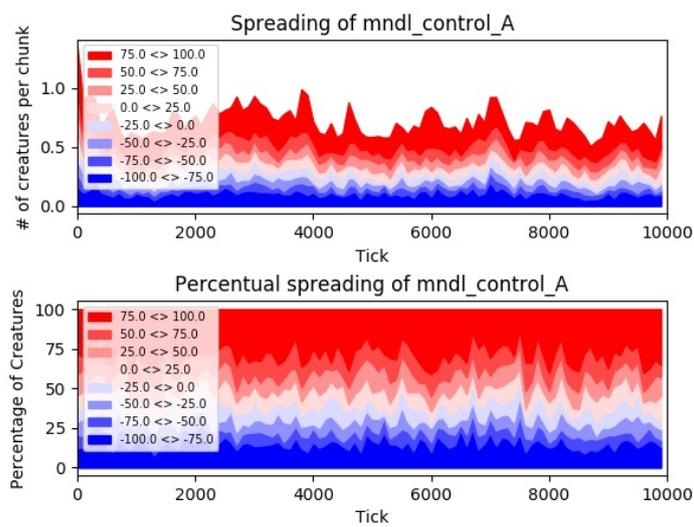


Figura 9: Grafici dell'evoluzione del gene *num_control*



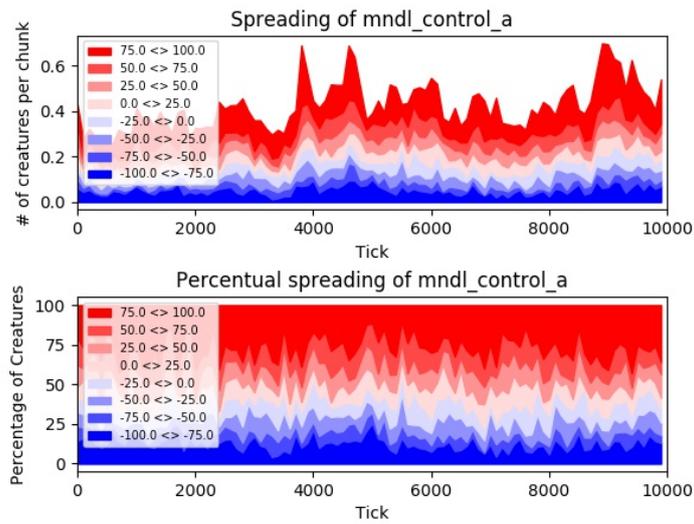


Figura 10: Grafici della distribuzione delle creature in base alla temperatura. I quattro grafici rappresentano i fenotipi *A* ed *a* del gene *mndl_control* (dal punto di vista del numero assoluto delle creature e della percentuale)

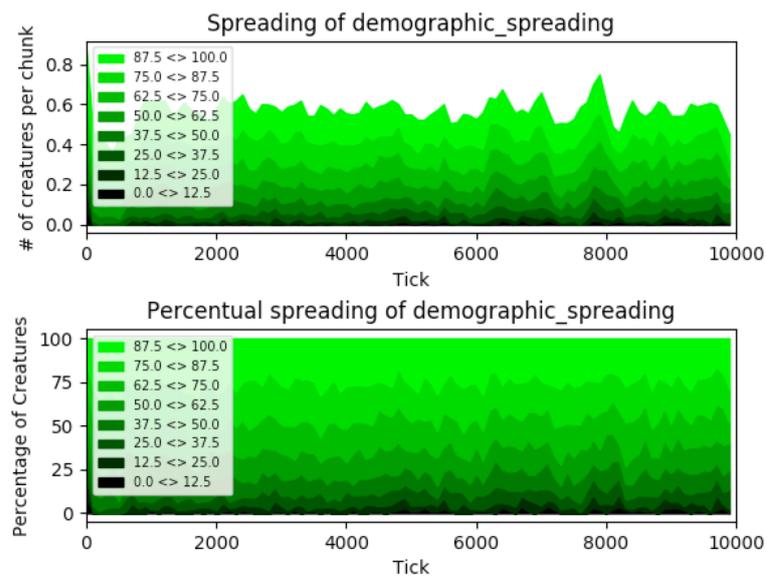


Figura 11: Distribuzione delle creature nelle fasce in cui si trova piú cibo (quelle chiare), rispetto a quelle dove si trova meno cibo (quelle piú scure).

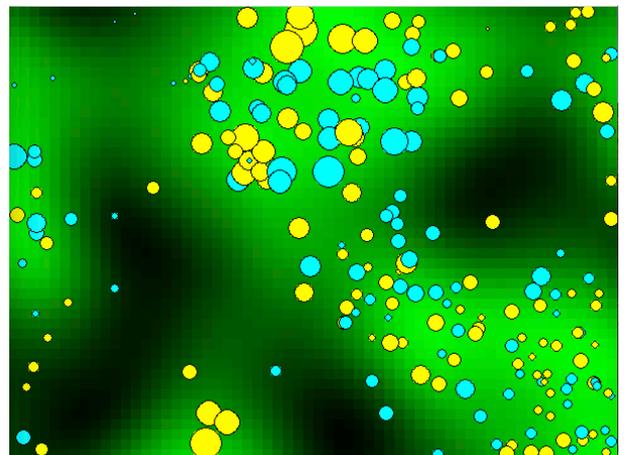
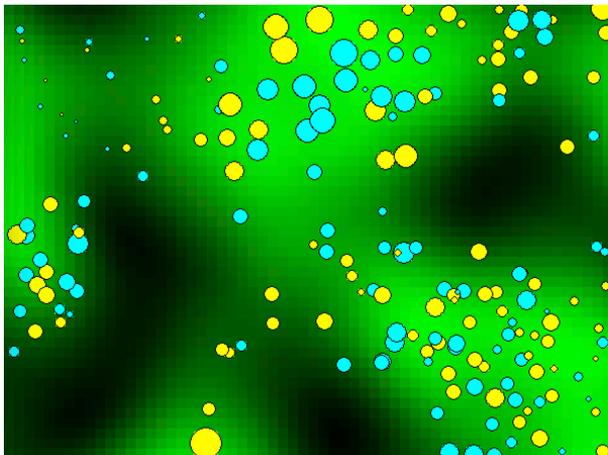


Figura 12: Due istanti di una simulazione in cui si può osservare la distribuzione delle creature sul territorio in base alla quantità massima di cibo.

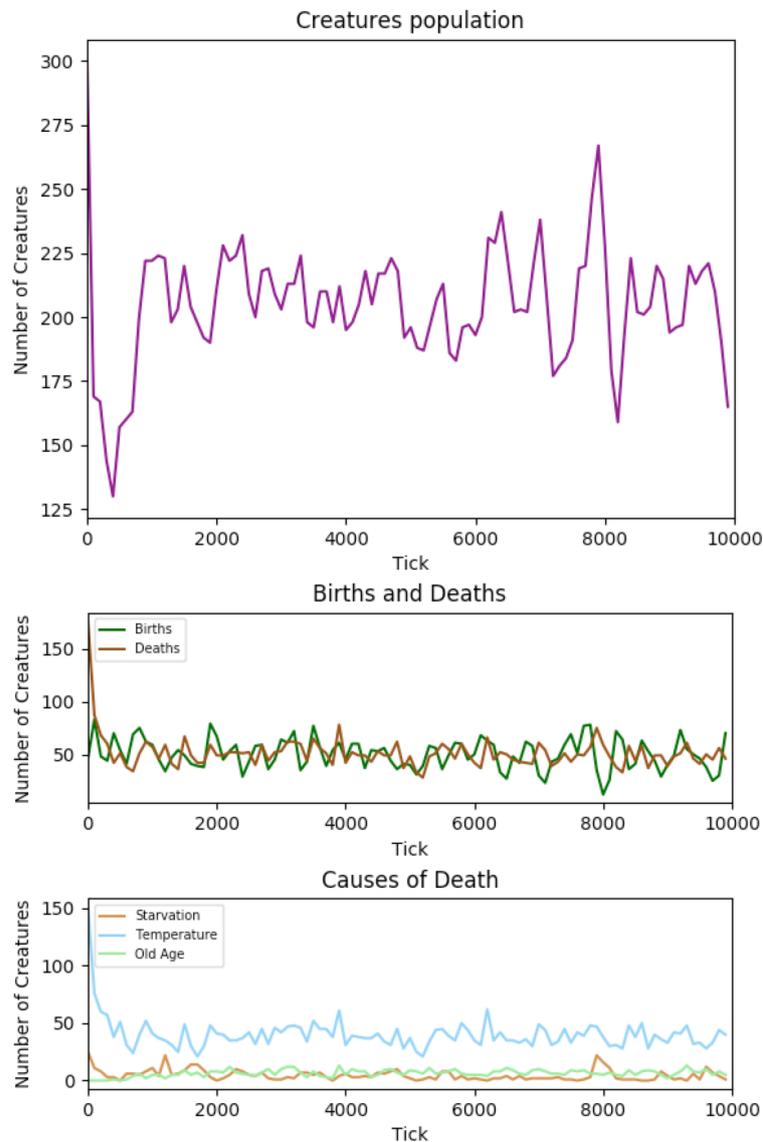


Figura 13: Andamento della popolazione: nel primo grafico si osserva il numero di creature con il passare dei *tick*, con un andamento sinusoidale, mentre il secondo grafico mostra l'andamento di nascite e morti e il terzo l'entità delle varie cause di morte.