Velocity distributions of foraging bumblebees in the presence of predators

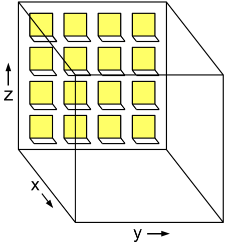
Friedrich Lenz[1], Tom Ings[2], Lars Chittka[2], Aleksei V. Chechkin[3], Holger Kantz[4], Rainer Klages[1]
lenz@maths.qmul.ac.uk

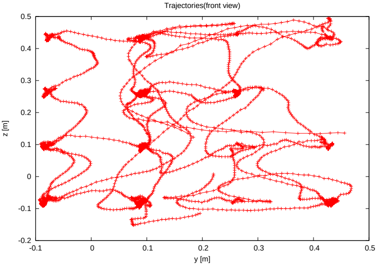
1 Queen Mary University of London, School of Mathematical Sciences
2 Queen Mary University of London, School of Biol. and Chem. Sci.
3 Inst. for Theo. Physics, NSC KIPT, Kharkov, Ukraine
4 Max Planck Institute for the Physics of Complex Systems, Dresden

March 9, 2010

Introduction to the Bumblebee Experiment

- Small experiment in a cube of $\approx 75cm$ side length:
  - Advantages: systematic variation of the environment possible; easier than tracking bumblebees on large scales
  - Disadvantage: no typical free flight of bumblebees
- Two cameras track the position of a bumblebee (50fps)
- Artificial yellow flowers: 4x4 grid on one wall





VARIATION OF ENVIRONMENT

B C

- Two sorts of artificial spiders: white (easily visible) and yellow (cryptic)
- #bumblebees=30 , #data per bumblebee for each stage $\approx$ 7000

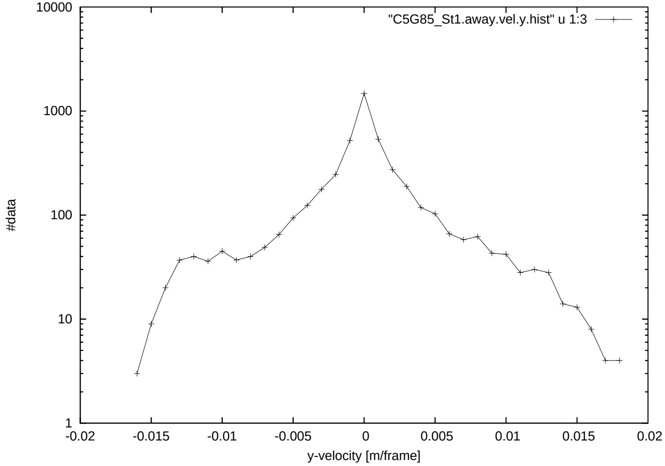
7 experimental stages:

1. Pretraining
2. Training
3. Neutral
4. Midterm-Memory Test
5. Reinforcement Training
6. Remotivation a day later
7. Longterm-Memory Test

Experiment done by:
Thomas C. Ings and Lars Chittka. Current Biology, 18(19):1520-1524 (2008)

HISTOGRAM OF VELOCITY (SEMILOG)

Size of each bin is $5cm/s$



Maximum-Likelihood-Estimation

Given *n* datapoints $\{x_j\}$ and a $pdf_\lambda(x)$ (where λ is a vector of k parameters) the *Likelihood* is:

$$L(\lambda|data) = \prod_{x_j \in data} pdf_\lambda(x_j)$$

MAXIMUM-LIKELIHOOD-ESTIMATION

Given *n* datapoints $\{x_j\}$ and a $pdf_\lambda(x)$ (where λ is a vector of k parameters) the *Likelihood* is:

$$L(\lambda|data) = \prod_{x_j \in data} pdf_\lambda(x_j)$$

$$\ln L(\lambda|data) = \sum_{x_j \in data} \ln pdf_\lambda(x_j)$$

Maximum-Likelihood-Estimation

Given n datapoints $\{x_j\}$ and a $pdf_\lambda(x)$ (where λ is a vector of k parameters) the *Likelihood* is:

$$L(\lambda|data) = \prod_{x_j \in data} pdf_\lambda(x_j)$$

$$\ln L(\lambda|data) = \sum_{x_j \in data} \ln pdf_\lambda(x_j)$$

-> maximize $\ln L$ w.r.t. λ e.g. with Nelder-Mead (MC to find global maximum)

MAXIMUM-LIKELIHOOD-ESTIMATION

Given n datapoints $\{x_j\}$ and a $pdf_\lambda(x)$ (where λ is a vector of k parameters) the *Likelihood* is:

$$L(\lambda|data) = \prod_{x_j \in data} pdf_\lambda(x_j)$$

$$\ln L(\lambda|data) = \sum_{x_j \in data} \ln pdf_\lambda(x_j)$$

-> maximize $\ln L$ w.r.t. λ e.g. with Nelder-Mead (MC to find global maximum)

For pre-binned data:

$$\ln L(\lambda|data) = \sum_{x_j \in data} \ln \int_{min(bin(x_j))}^{max(bin(x_j))} pdf_\lambda(x')dx'$$

$$\ln L(\lambda|data) = \sum_{b \in bins} hist[b] \ln \int_{min(b)}^{max(b)} pdf_\lambda(x')dx'$$

Candidate Velocity-Distributions

1. Exponential

$$pdf_{\lambda}(x) = ce^{-\lambda x}, x \geq cutoff$$

2. Powerlaw

$$pdf_{\mu}(x) = cx^{-\mu}, x \geq cutoff$$

3. Mixture of two normal distributions

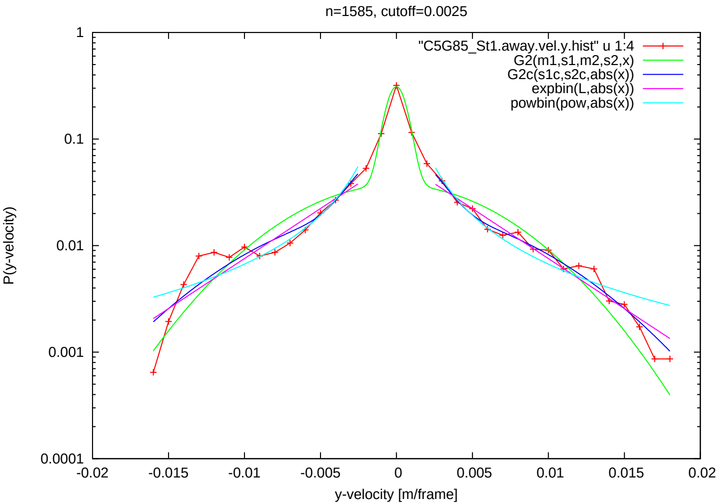
$$pdf_{a,\sigma_1,\sigma_2}(x) = aN(\sigma_1) + (1-a)N(\sigma_2), x \geq cutoff$$

where $N(\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$

4. Mixture of two normal distributions *without* cutoff

$$pdf_{a,\sigma_1,\sigma_2}(x) = aN(\sigma_1) + (1-a)N(\sigma_2)$$

ESTIMATED DISTR.FOR A BUMBLEBEE (SEMILOG)



Akaike information criterion (AIC)

To find the preference between models $i \in \{1, \ldots, 4\}$ whose likelihoods L_i are maximized at λ_i^{max}

$$AIC_i = -2\ln(L_i(\lambda_i^{max}|data) + 2k_i$$

Best model $*$: $AIC_* = \min_i(AIC_i)$

Akaike weights:

$$w_i = c\,\mathrm{e}^{-(AIC_i - AIC_*)/2}$$

where c normalizes the weights: $\sum_i w_i = 1$

$$w_i = c\,\frac{L_i(\lambda_i^{max}|data)}{L_*(\lambda_*^{max}|data)}\mathrm{e}^{k_* - k_i}$$

AKAIKE INFORMATION CRITERION (AIC)

To find the preference between models $i \in \{1, \ldots, 4\}$ whose likelihoods L_i are maximized at λ_i^{max}

$$AIC_i = -2\ln(L_i(\lambda_i^{max}|data) + 2k_i$$

Best model $*$: $AIC_* = \min_i(AIC_i)$

Akaike weights:

$$w_i = c\mathrm{e}^{-(AIC_i - AIC_*)/2}$$

where c normalizes the weights: $\sum_i w_i = 1$

$$w_i = c\frac{L_i(\lambda_i^{max}|data)}{L_*(\lambda_*^{max}|data)}\mathrm{e}^{k_* - k_i}$$

$\Rightarrow$ Mixture of 2 Gaussians best for all bumblebees in all stages

Bayesian information criterion (BIC)

To find the preference between models $i \in \{1, \ldots, 4\}$ whose likelihoods L_i are maximized at λ_i^{max}

$$BIC_i = -2\ln(L_i(\lambda_i^{max}|data) + \ln(n)k_i$$

Best model $*$: $BIC_* = \min_i(BIC_i)$

Akaike weights:

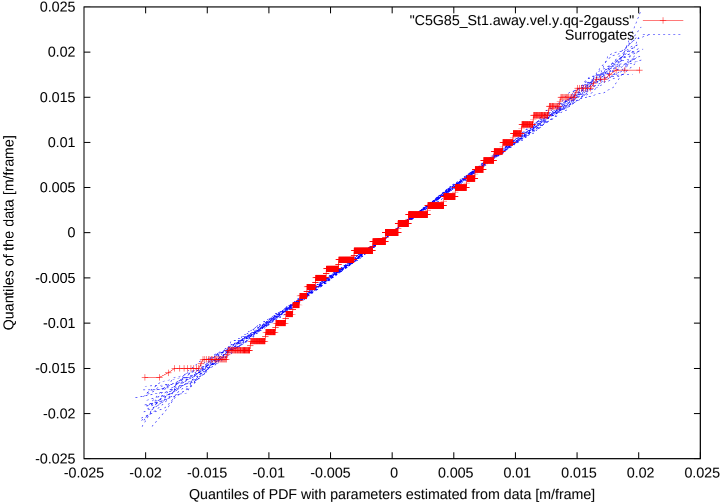
$$w_i = c\mathrm{e}^{-(BIC_i - BIC_*)/2}$$

where c normalizes the weights: $\sum_i w_i = 1$

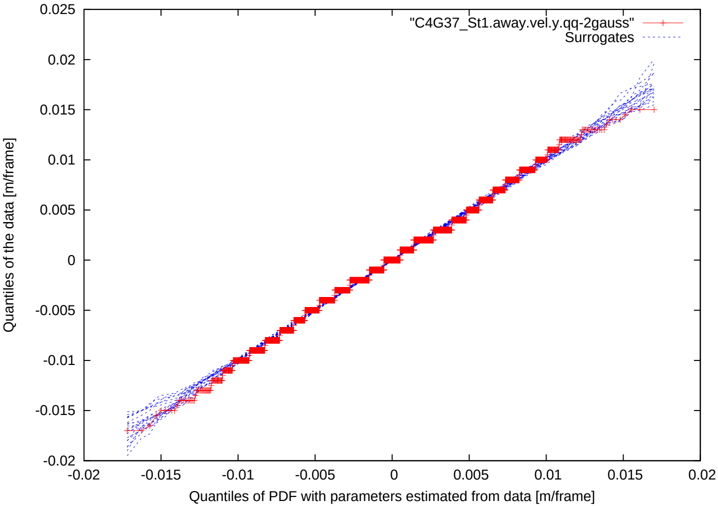
$$w_i = c\frac{L_i(\lambda_i^{max}|data)}{L_*(\lambda_*^{max}|data)}\mathrm{n}^{\frac{k_* - k_i}{2}}$$

$\Rightarrow$ Mixture of 2 Gaussians for 90% of the datasets
For large cutoffs: Preference of Exponential for 10%

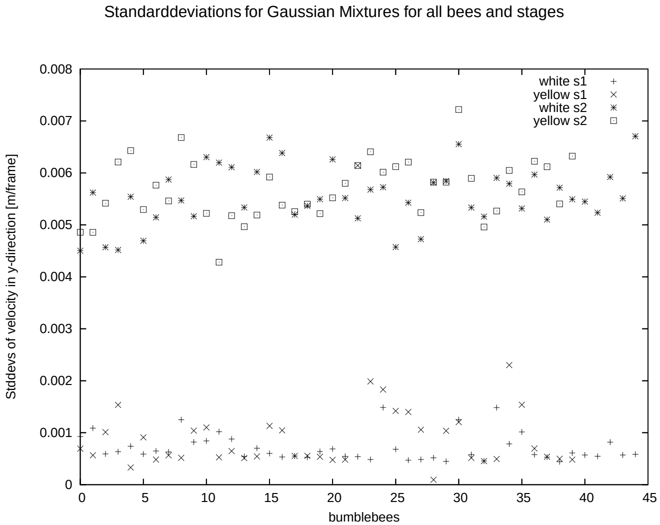
QUANTILE-QUANTILE-PLOT: GAUSSIAN MIXTURE



QQ-PLOT WITH SURROGATES: BEE WITH MORE DATA



PARAMETERS IN DIFFERENT ENVIRONMENTS



Summary

- Mixtures of two Gaussians are best approximations for the velocity distributions
- Deviations for medium velocities are quite large for some bumblebees (with small n). Reason is autocorrelation?
- Powerlaw tails were the worst model followed by exponential tails
- No correlation between parameters of the best models and environmental parameters found: velocities not affected by the presence of spiders.

Outlook

- Compare to other distributions: e.g. Normal+Exponential, Normal+Powerlaw, Lévy-stable
- Describe differences in the distributions for the 3 directions
- Analysis of distances between turning points and coarse grained velocities