Chapter 9 Decision Analysis

- 9.1 Context and basic steps (most important part)
- 9.2 Example
- 9.3 Multistage decision analysis (example)
- 9.4 Hierarchical decision analysis (example)
- 9.5 Personal vs. institutional decision analysis

Bayesian decision theory

- Potential decisions d
  - or actions a

Bayesian decision theory

- Potential decisions d
  - or actions a
- Potential consequences x
  - x may be categorical, ordinal, real, scalar, vector, etc.

Bayesian decision theory

- Potential decisions d
  - or actions a
- Potential consequences x
  - x may be categorical, ordinal, real, scalar, vector, etc.
- Probability distributions of consequences given decisions $p(x \mid d)$
  - in decision making the decisions are controlled and thus $p(d)$ does not exist

Bayesian decision theory

- Potential decisions d
  - or actions a
- Potential consequences x
  - x may be categorical, ordinal, real, scalar, vector, etc.
- Probability distributions of consequences given decisions $p(x \mid d)$
  - in decision making the decisions are controlled and thus $p(d)$ does not exist
- Utility function $U(x)$ maps consequences to real value
  - e.g. euro or expected lifetime
  - instead of utility sometimes cost or loss is defined

Bayesian decision theory

- Potential decisions d
  - or actions a
- Potential consequences x
  - x may be categorical, ordinal, real, scalar, vector, etc.
- Probability distributions of consequences given decisions $p(x \mid d)$
  - in decision making the decisions are controlled and thus $p(d)$ does not exist
- Utility function $U(x)$ maps consequences to real value
  - e.g. euro or expected lifetime
  - instead of utility sometimes cost or loss is defined
- Expected utility $\mathrm{E}[U(x) \mid d] = \int U(x)p(x \mid d)dx$

Bayesian decision theory

- Potential decisions d
  - or actions a
- Potential consequences x
  - x may be categorical, ordinal, real, scalar, vector, etc.
- Probability distributions of consequences given decisions $p(x \mid d)$
  - in decision making the decisions are controlled and thus $p(d)$ does not exist
- Utility function $U(x)$ maps consequences to real value
  - e.g. euro or expected lifetime
  - instead of utility sometimes cost or loss is defined
- Expected utility $\mathrm{E}[U(x) \mid d] = \int U(x) p(x \mid d) dx$
- Choose decision d^*, which maximizes the expected utility

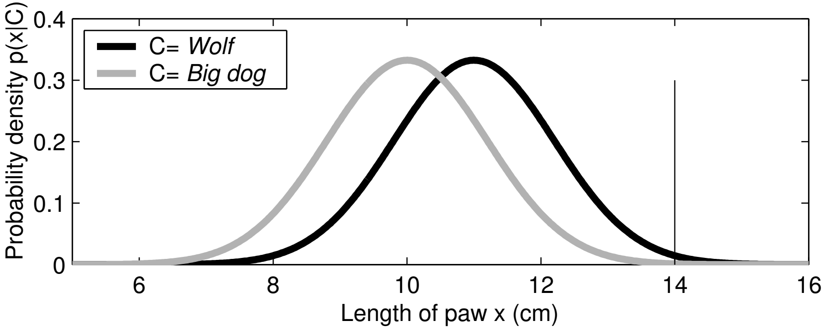
$$d^* = \arg\max_d \mathrm{E}[U(x) \mid d]$$

Example of decision making: 2 choices

- Helen is going to pick mushrooms in a forest, while she notices a paw print which could made by a dog or a wolf

Example of decision making: 2 choices

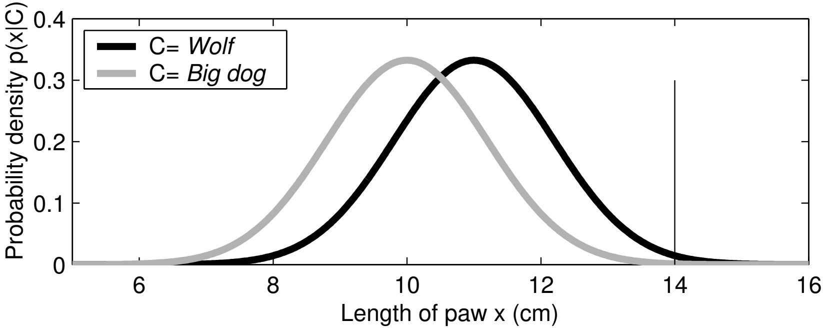
- Helen is going to pick mushrooms in a forest, while she notices a paw print which could made by a dog or a wolf
- Helen measures that the length of the paw print is 14 cm and goes home to Google how big paws dogs and wolves have, and tries then to infer which animal has made the paw print



observed length has been marked with a horizontal line

Example of decision making: 2 choices

- Helen is going to pick mushrooms in a forest, while she notices a paw print which could made by a dog or a wolf
- Helen measures that the length of the paw print is 14 cm and goes home to Google how big paws dogs and wolves have, and tries then to infer which animal has made the paw print



observed length has been marked with a horizontal line

- Likelihood of wolf is 0.92 (alternative being dog)

Example of decision making

- Helen assumes also that in her living area there are about one hundred times more free running dogs than wolves, that is *a priori* probability for wolf, before observation is 1%.

Example of decision making

- Helen assumes also that in her living area there are about one hundred times more free running dogs than wolves, that is *a priori* probability for wolf, before observation is 1%.
- Likelihood and posterior

Animal	Likelihood	Posterior probability
Wolf	0.92	0.10
Dog	0.08	0.90

Example of decision making

- Helen assumes also that in her living area there are about one hundred times more free running dogs than wolves, that is *a priori* probability for wolf, before observation is 1%.
- Likelihood and posterior

Animal	Likelihood	Posterior probability
Wolf	0.92	0.10
Dog	0.08	0.90

- Posterior probability of wolf is 10%

Example of decision making

- Helen has to make decision whether to go pick mushrooms

Example of decision making

- Helen has to make decision whether to go pick mushrooms
- If she doesn’t go to pick mushrooms utility is zero

Example of decision making

- Helen has to make decision whether to go pick mushrooms
- If she doesn't go to pick mushrooms utility is zero
- Helen assigns positive utility 1 for getting fresh mushrooms

Example of decision making

- Helen has to make decision whether to go pick mushrooms
- If she doesn't go to pick mushrooms utility is zero
- Helen assigns positive utility 1 for getting fresh mushrooms
- Helen assigns negative utility -1000 for a event that she goes to the forest and wolf attacks (for some reason Helen assumes that wolf will always attack)

Example of decision making

- Helen has to make decision whether to go pick mushrooms
- If she doesn't go to pick mushrooms utility is zero
- Helen assigns positive utility 1 for getting fresh mushrooms
- Helen assigns negative utility -1000 for a event that she goes to the forest and wolf attacks (for some reason Helen assumes that wolf will always attack)

	Animal	
Decision d	Wolf	Dog
Stay home	0	0
Go to the forest	-1000	1

Utility matrix $U(x)$

Example of decision making

- Helen has to make decision whether to go pick mushrooms
- If she doesn't go to pick mushrooms utility is zero
- Helen assigns positive utility 1 for getting fresh mushrooms
- Helen assigns negative utility -1000 for a event that she goes to the forest and wolf attacks (for some reason Helen assumes that wolf will always attack)

Decision d	Animal Wolf	Dog
Stay home	0	0
Go to the forest	-1000	1

Utility matrix $U(x)$

Action d	Expected utility $\mathrm{E}[U(x) \mid d]$
Stay home	0
Go to the forest	-100+0.9

Utilities for different actions

Example of decision making

- Maximum likelihood decision would be to assume that there is a wolf

Example of decision making

- Maximum likelihood decision would be to assume that there is a wolf
- Maximum posterior decision would be to assume that there is a dog

Example of decision making

- Maximum likelihood decision would be to assume that there is a wolf
- Maximum posterior decision would be to assume that there is a dog
- Maximum utility decision is to stay home, even if it is more likely that the animal is dog

Example of decision making

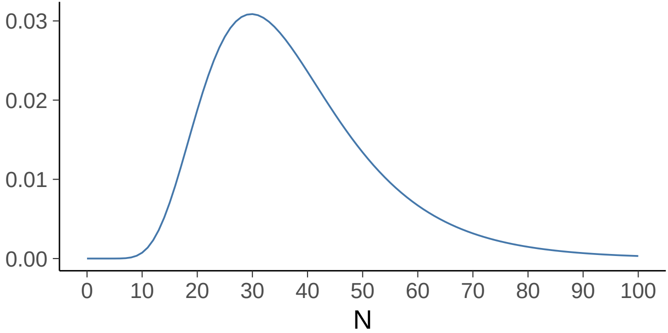
- Maximum likelihood decision would be to assume that there is a wolf
- Maximum posterior decision would be to assume that there is a dog
- Maximum utility decision is to stay home, even if it is more likely that the animal is dog
- Example illustrates that the uncertainties (probabilities) related to all consequences need to be carried on until final decision making

Example of decision making: several choices

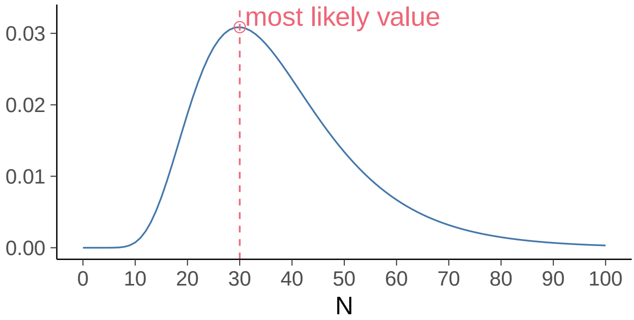
- You decide to earn money by selling a seasonal product
    - You pay 7€ per each, and sell them 10€ each
    - You need to decide how many (N) items to buy

Example of decision making: several choices

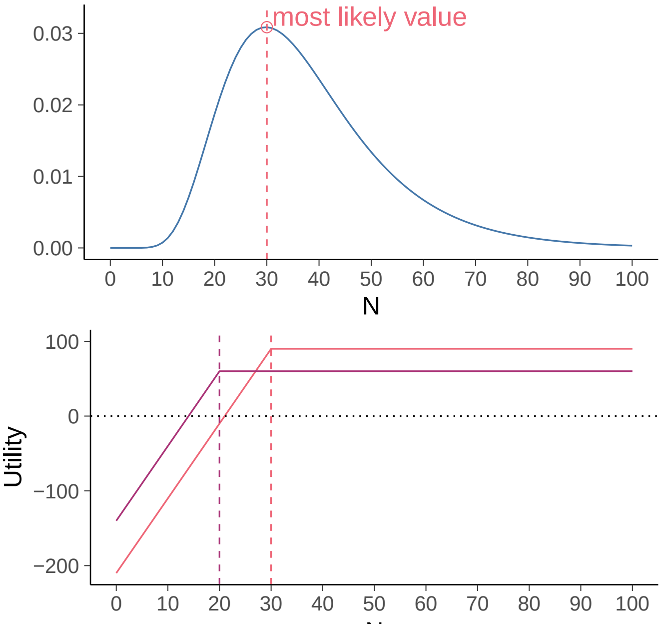
- You decide to earn money by selling a seasonal product
  - You pay 7€ per each, and sell them 10€ each
  - You need to decide how many (N) items to buy
  - You ask your friends how many they used to sell and estimate a distribution for how many you might sell



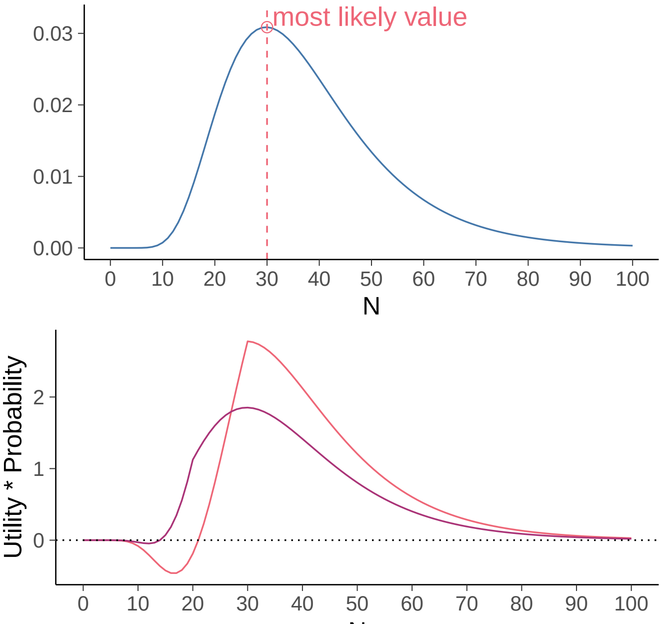
Example of decision making: several choices



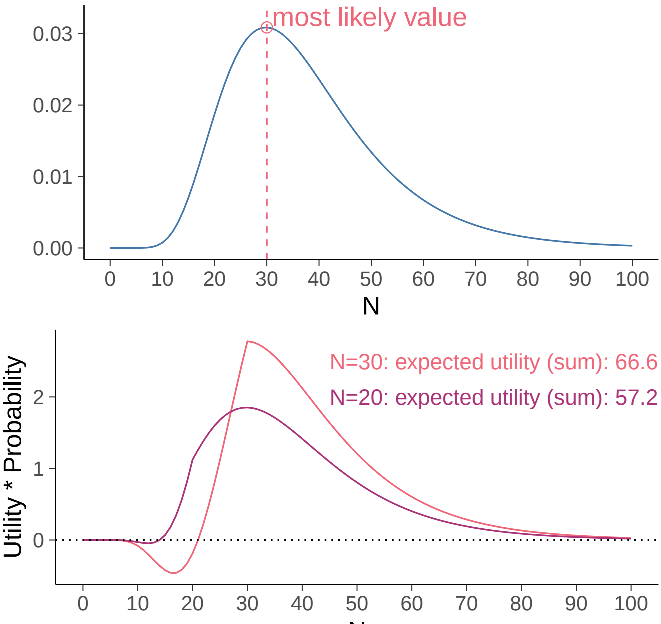
Example of decision making: several choices



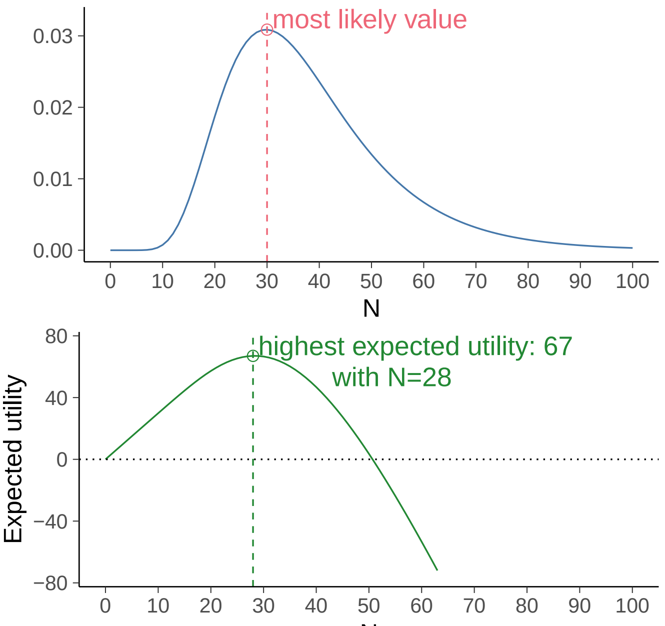
Example of decision making: several choices



Example of decision making: several choices



Example of decision making: several choices



Decision making in sales

- Common task in commerce and restaurants

Challenges in decision making

- Actual utility functions are rarely linear

Challenges in decision making

- Actual utility functions are rarely linear
  - the expected utility is 5€ for
    a) 100% of receiving 5€
    b) 50% of losing 1M€ and 50% of winning 1M€ + 10€

Challenges in decision making

- Actual utility functions are rarely linear
  - the expected utility is 5€ for
    a) 100% of receiving 5€
    b) 50% of losing 1M€ and 50% of winning 1M€ + 10€
  - most gambling has negative expected utility
    (but the excitement of the game may have positive utility)

Challenges in decision making

- Actual utility functions are rarely linear
  - the expected utility is 5€ for
    a) 100% of receiving 5€
    b) 50% of losing 1M€ and 50% of winning 1M€ + 10€
  - most gambling has negative expected utility
    (but the excitement of the game may have positive utility)
- What is the cost of human life?

Challenges in decision making

- Actual utility functions are rarely linear
  - the expected utility is 5€ for
    a) 100% of receiving 5€
    b) 50% of losing 1M€ and 50% of winning 1M€ + 10€
  - most gambling has negative expected utility
    (but the excitement of the game may have positive utility)
- What is the cost of human life?
- Multiple parties having different utilities

Model selection as decision problem

- Choose the model that maximizes the expected utility of using the model to make predictions / decisions in the future

Multi-stage decision making (Section 9.3)

- 95 year old has a tumor that is malignant with 90% prob

Multi-stage decision making (Section 9.3)

- 95 year old has a tumor that is malignant with 90% prob
- Based on statistics
  - expected lifetime is 34.8 months if no cancer

Multi-stage decision making (Section 9.3)

- 95 year old has a tumor that is malignant with 90% prob
- Based on statistics
  - expected lifetime is 34.8 months if no cancer
  - expected lifetime is 16.7 months if cancer and radiation therapy is used

Multi-stage decision making (Section 9.3)

- 95 year old has a tumor that is malignant with 90% prob
- Based on statistics
  - expected lifetime is 34.8 months if no cancer
  - expected lifetime is 16.7 months if cancer and radiation therapy is used
  - expected lifetime is 20.3 months if cancer and surgery, but the probability of dying in surgery is 35% (for 95 year old)

Multi-stage decision making (Section 9.3)

- 95 year old has a tumor that is malignant with 90% prob
- Based on statistics
  - expected lifetime is 34.8 months if no cancer
  - expected lifetime is 16.7 months if cancer and radiation therapy is used
  - expected lifetime is 20.3 months if cancer and surgery, but the probability of dying in surgery is 35% (for 95 year old)
  - expected lifetime is 5.6 months if cancer and no treatment

Multi-stage decision making (Section 9.3)

- 95 year old has a tumor that is malignant with 90% prob
- Based on statistics
  - expected lifetime is 34.8 months if no cancer
  - expected lifetime is 16.7 months if cancer and radiation therapy is used
  - expected lifetime is 20.3 months if cancer and surgery, but the probability of dying in surgery is 35% (for 95 year old)
  - expected lifetime is 5.6 months if cancer and no treatment
- Which treatment to choose?
  - quality adjusted life time
  - 1 month is subtracted for the time spent in treatments

Multi-stage decision making (Section 9.3)

- 95 year old has a tumor that is malignant with 90% prob
- Based on statistics
  - expected lifetime is 34.8 months if no cancer
  - expected lifetime is 16.7 months if cancer and radiation therapy is used
  - expected lifetime is 20.3 months if cancer and surgery, but the probability of dying in surgery is 35% (for 95 year old)
  - expected lifetime is 5.6 months if cancer and no treatment
- Which treatment to choose?
  - quality adjusted life time
  - 1 month is subtracted for the time spent in treatments
- Quality adjusted life time
  - See the book for the multi-stage decision making

Design of experiment

- Which experiment would give most additional information
  - decide values x_{n+1} for the next experiment
  - which values of x_{n+1} would reduce the posterior uncertainty or increase the expected utility most

Design of experiment

- Which experiment would give most additional information
  - decide values x_{n+1} for the next experiment
  - which values of x_{n+1} would reduce the posterior uncertainty or increase the expected utility most
- Example 1
  - biopsy in the cancer example

Design of experiment

- Which experiment would give most additional information
  - decide values x_{n+1} for the next experiment
  - which values of x_{n+1} would reduce the posterior uncertainty or increase the expected utility most
- Example 1
  - biopsy in the cancer example
- Example 2
  - imagine that in bioassay the posterior uncertainty of LD50 is too large
  - which dose should be used in the next experiment to reduce the variance of LD50 as much as possible ?
    - this way less experiments need to be made (and less animals need to be killed)

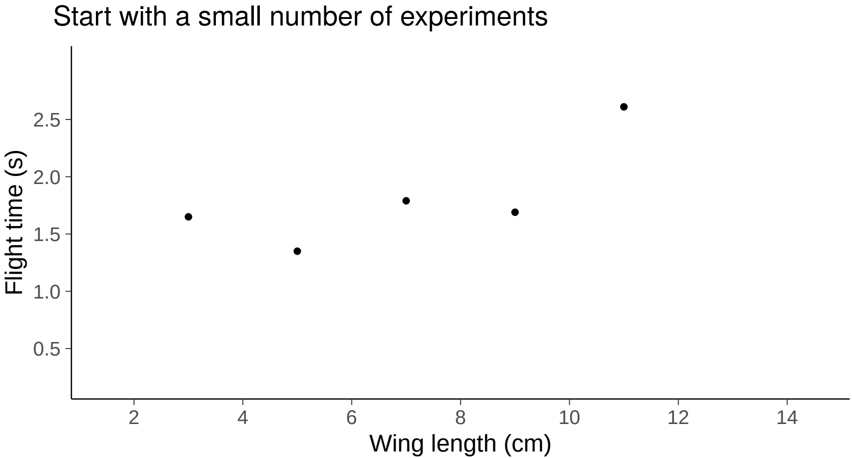
Design of experiment

- Which experiment would give most additional information
  - decide values x_{n+1} for the next experiment
  - which values of x_{n+1} would reduce the posterior uncertainty or increase the expected utility most
- Example 1
  - biopsy in the cancer example
- Example 2
  - imagine that in bioassay the posterior uncertainty of LD50 is too large
  - which dose should be used in the next experiment to reduce the variance of LD50 as much as possible ?
    - this way less experiments need to be made (and less animals need to be killed)
- Example 3
  - optimal paper helicopter wing length

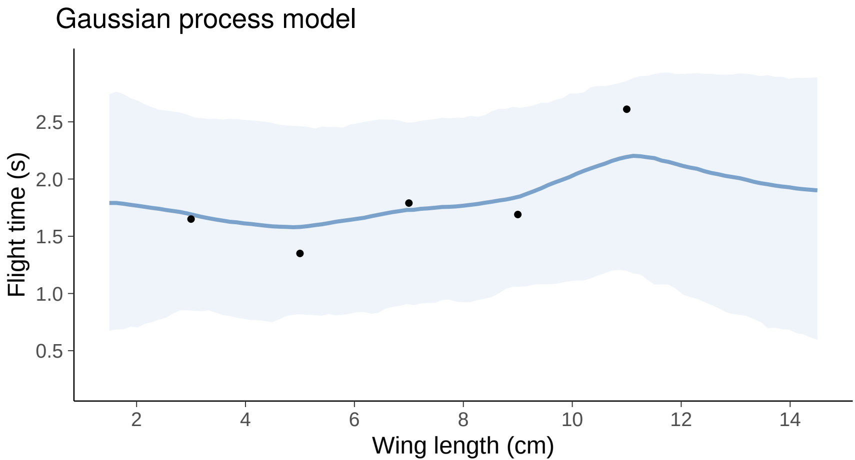
Bayesian optimization

- Design of experiment
- Used to optimize, for example,
  - machine learning / deep learning model structures, regularization, and learning algorithm parameters
  - material science
  - engines
  - drug testing
  - part of Bayesian inference for stochastic simulators

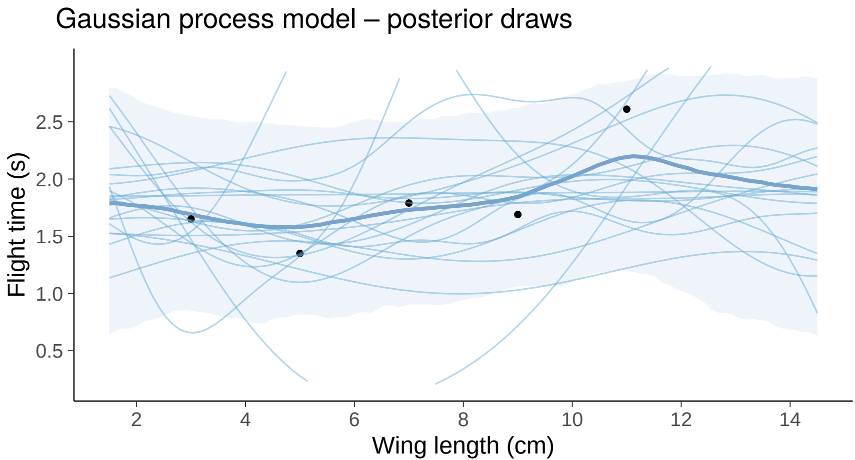
Bayesian optimization of wing length



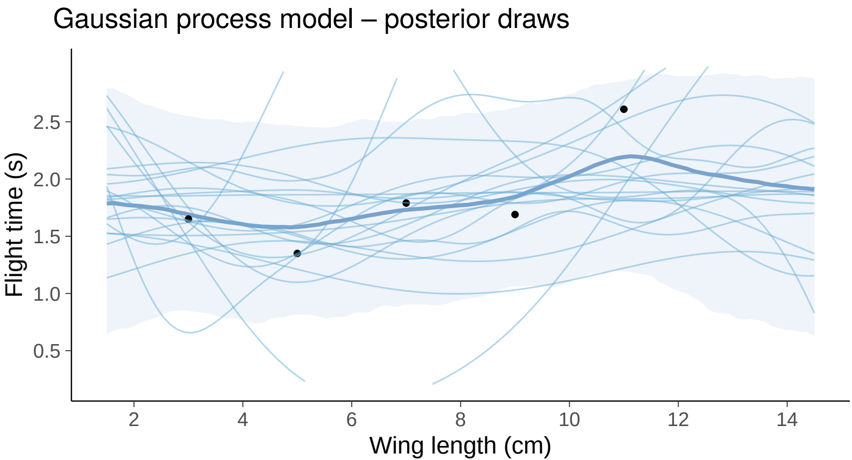
Bayesian optimization of wing length



Bayesian optimization of wing length

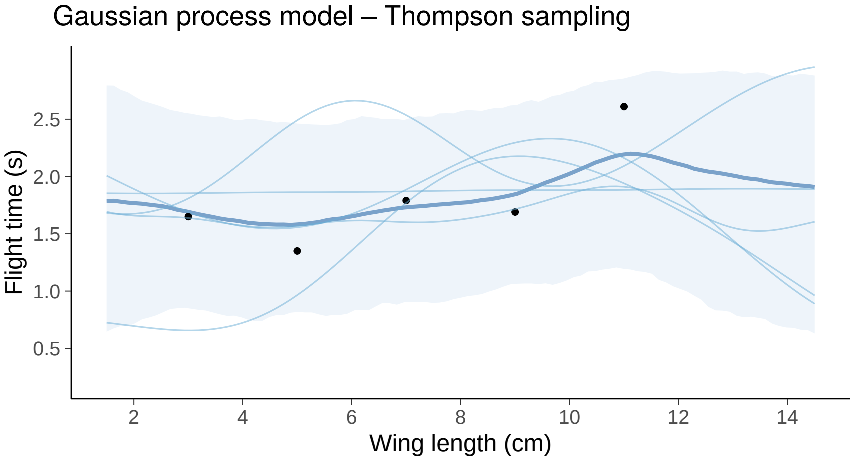


Bayesian optimization of wing length



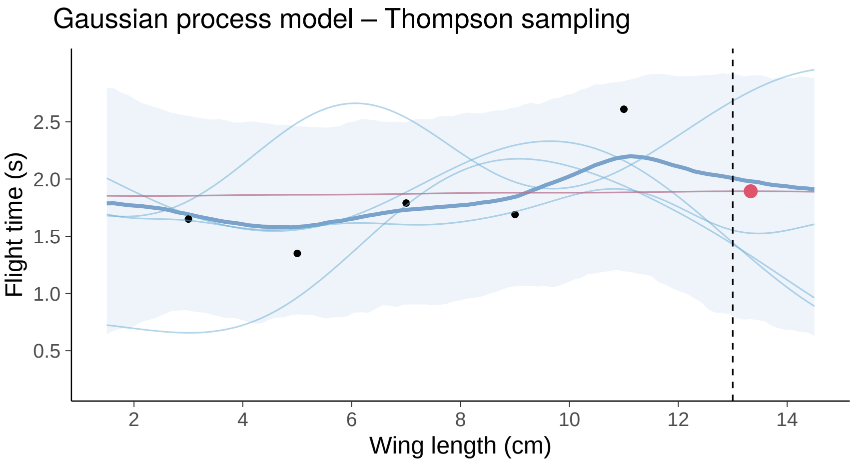
- Thompson sampling:
  - pick one posterior draw (function)
  - find the wing length corresponding to the max. of that draw
  - make the next observation with that wing length

Bayesian optimization of wing length



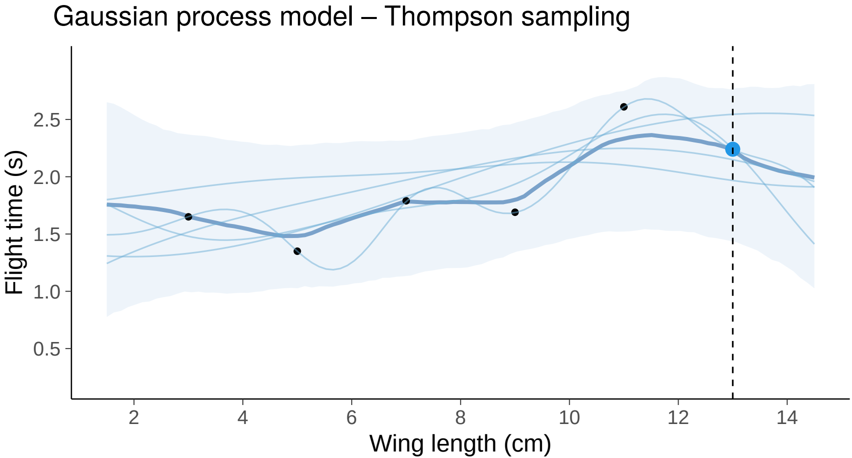
- Thompson sampling:
  - pick one posterior draw (function)
  - find the wing length corresponding to the max. of that draw
  - make the next observation with that wing length

Bayesian optimization of wing length



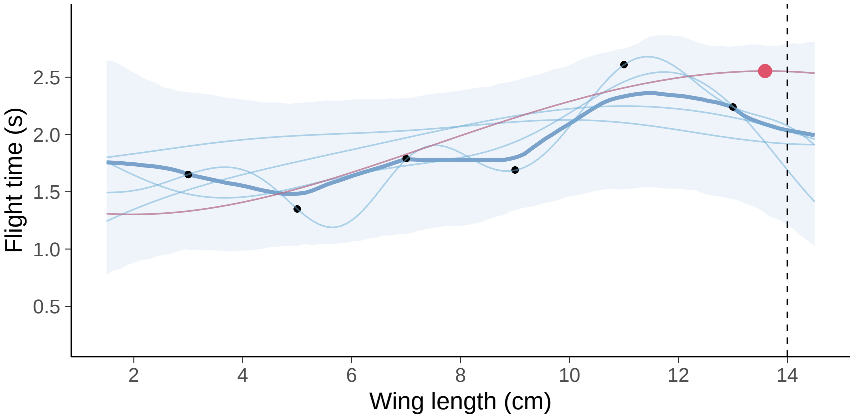
- Thompson sampling:
  - pick one posterior draw (function)
  - find the wing length corresponding to the max. of that draw
  - make the next observation with that wing length

Bayesian optimization of wing length



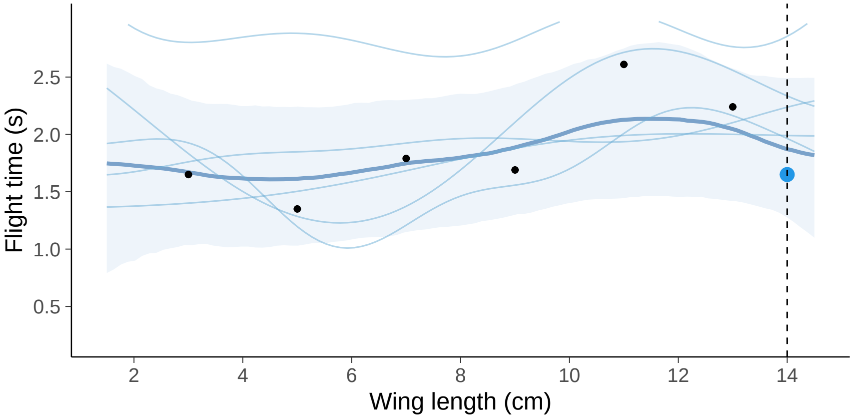
Bayesian optimization of wing length

Gaussian process model – Thompson sampling

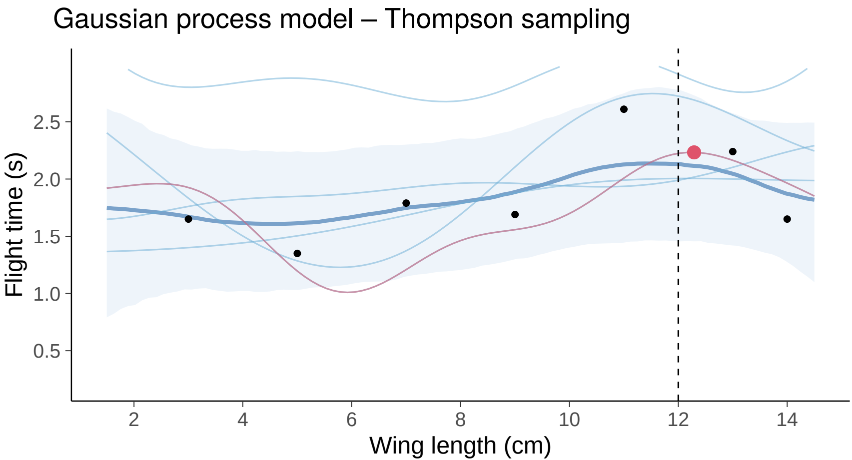


Bayesian optimization of wing length

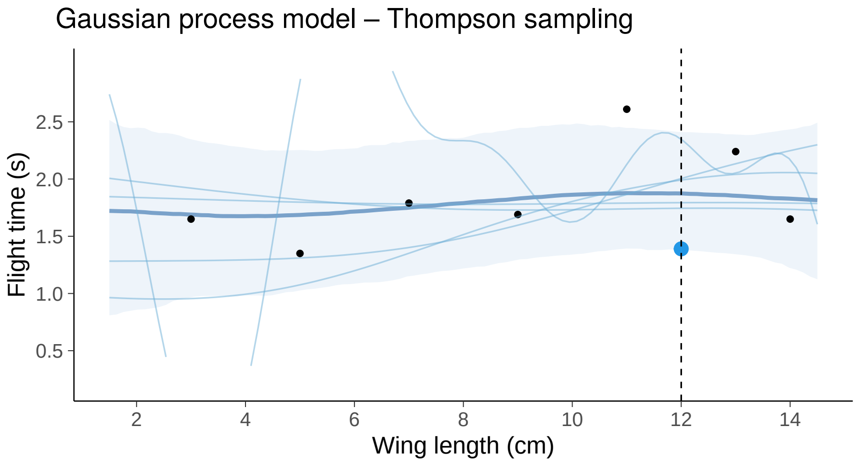
Gaussian process model – Thompson sampling



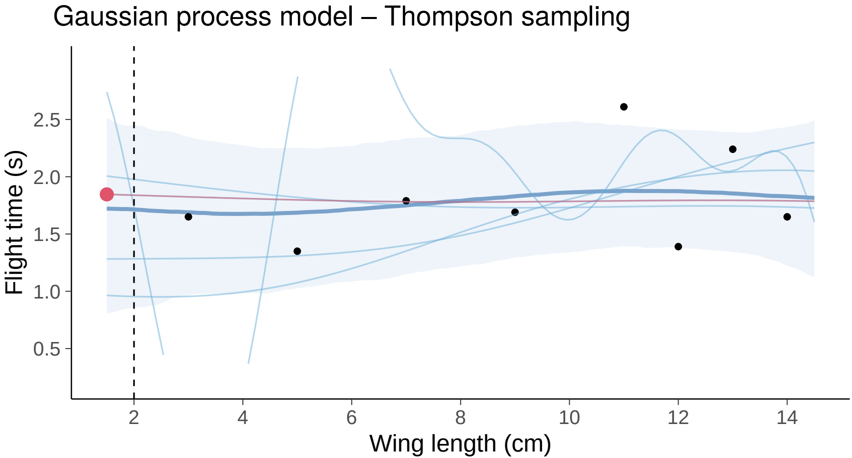
Bayesian optimization of wing length



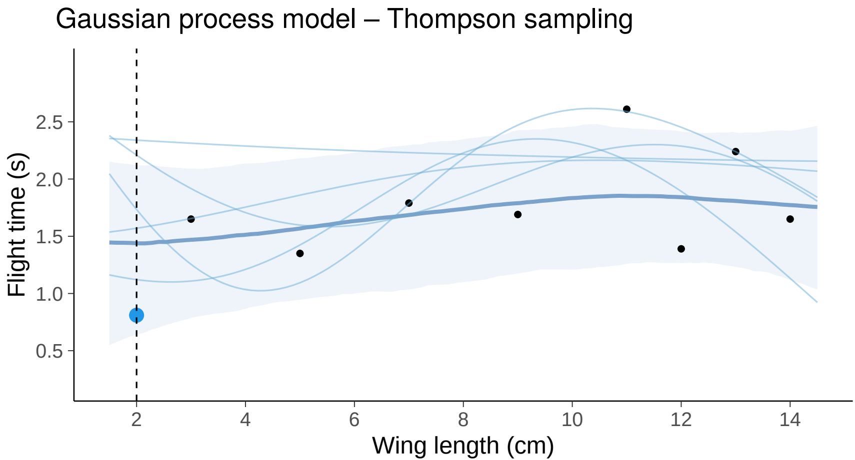
Bayesian optimization of wing length



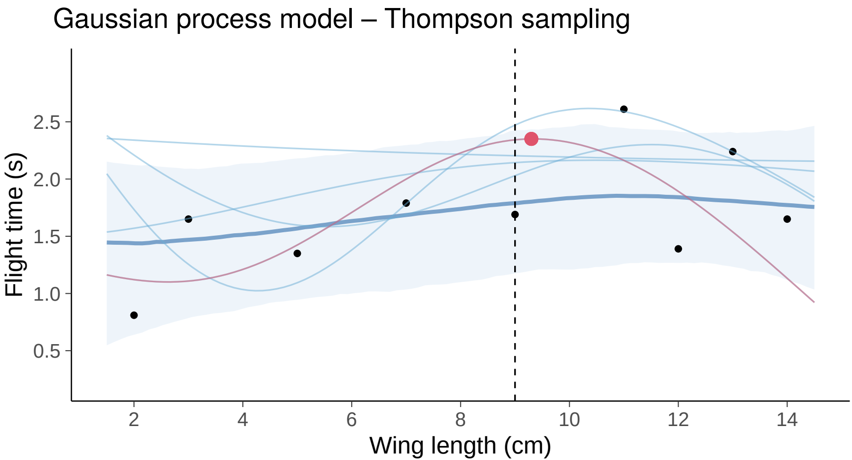
Bayesian optimization of wing length



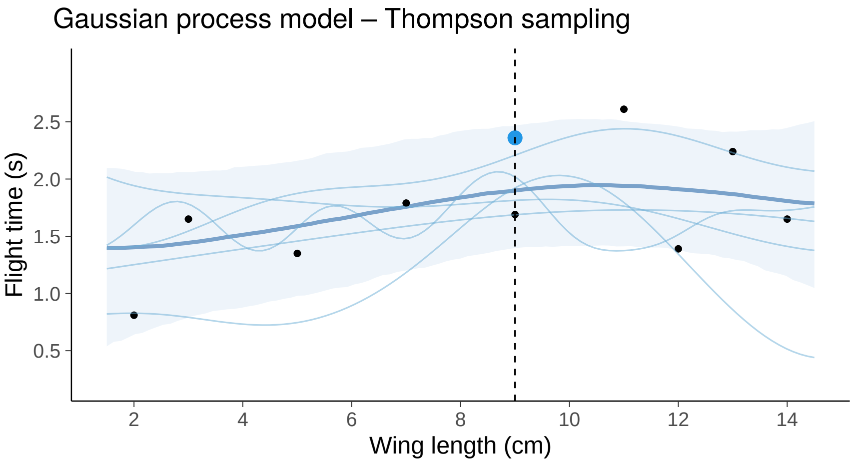
Bayesian optimization of wing length



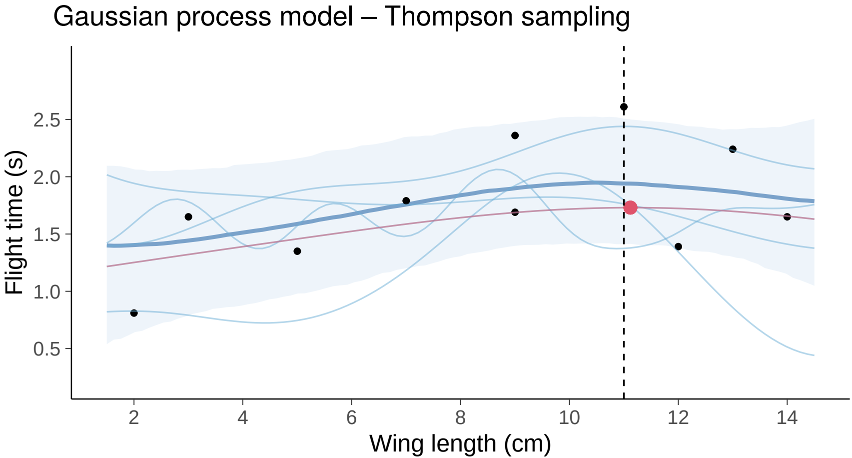
Bayesian optimization of wing length



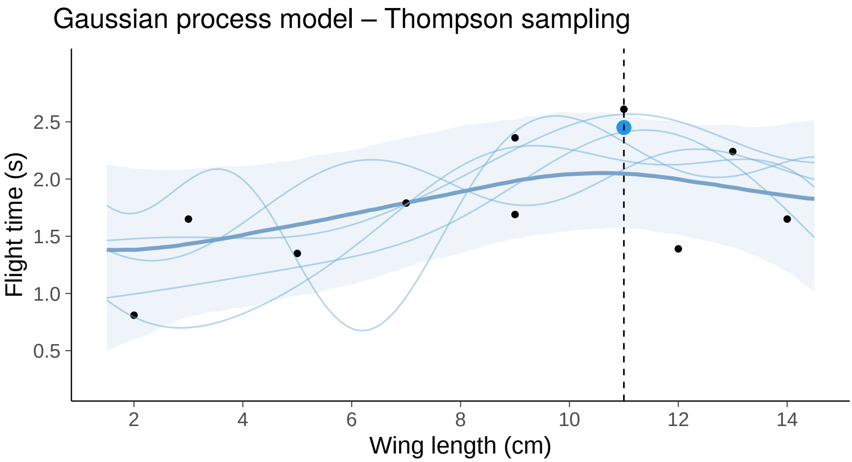
Bayesian optimization of wing length



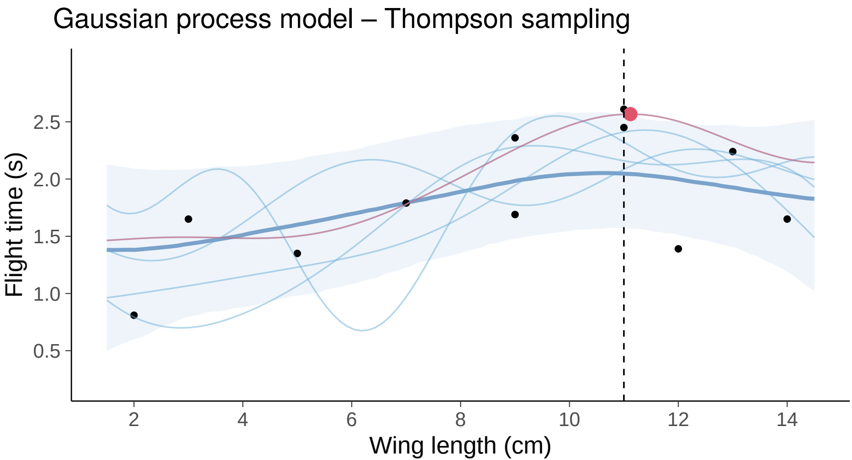
Bayesian optimization of wing length



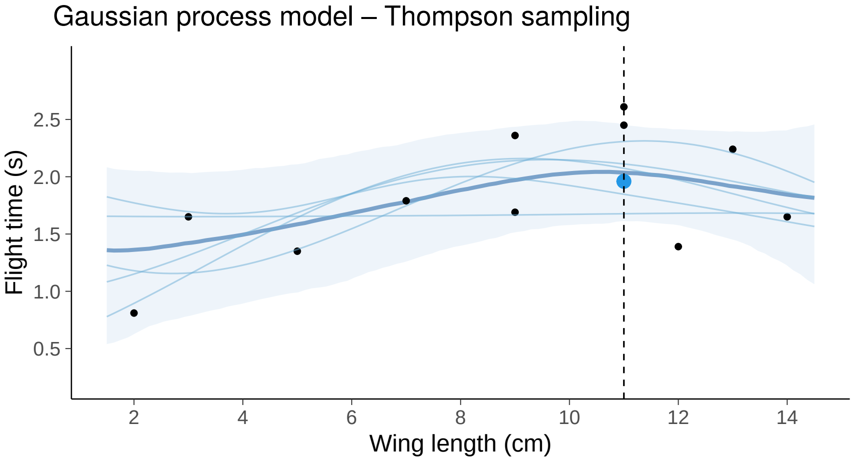
Bayesian optimization of wing length



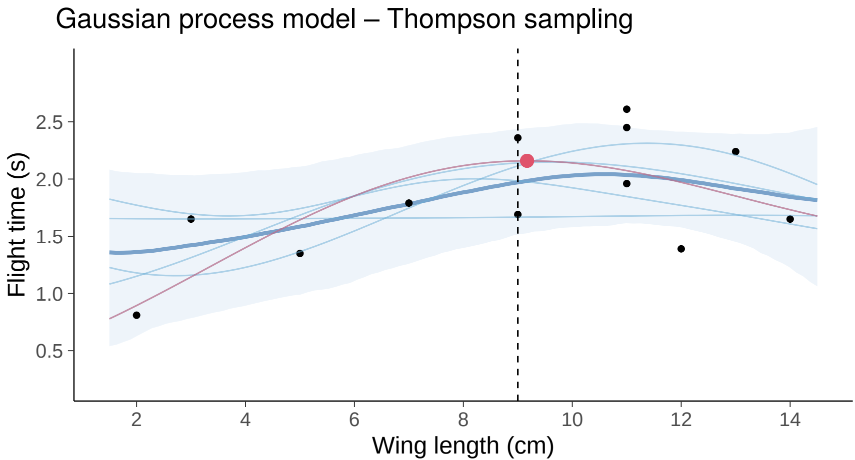
Bayesian optimization of wing length



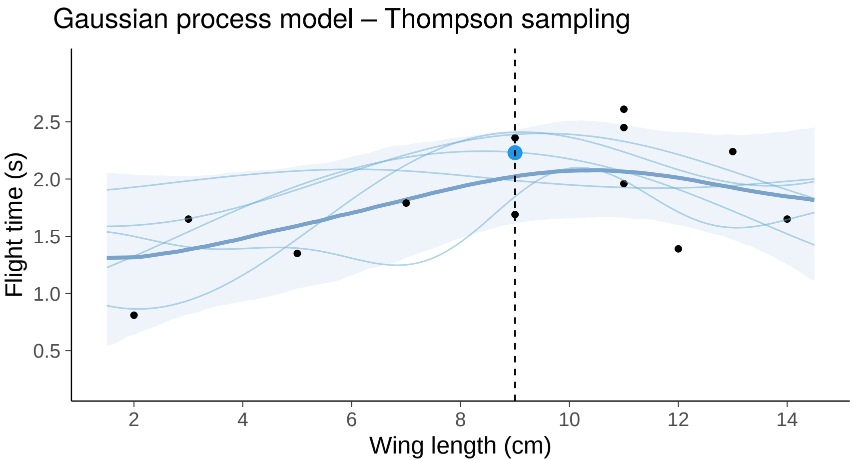
Bayesian optimization of wing length



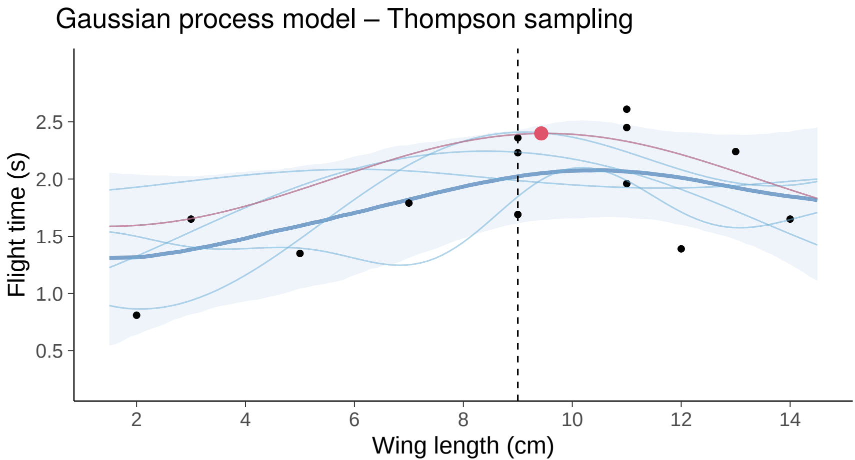
Bayesian optimization of wing length



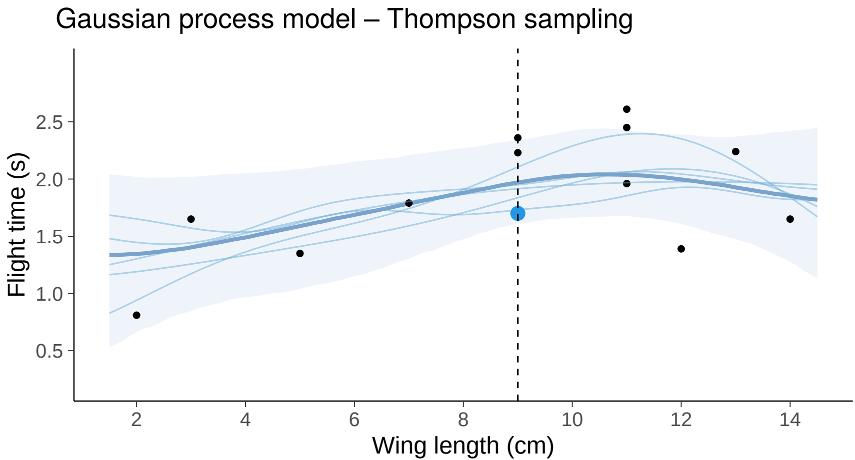
Bayesian optimization of wing length



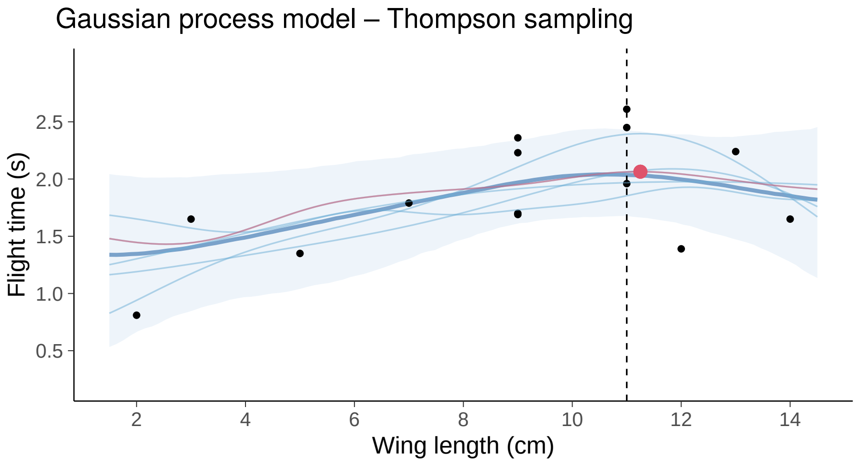
Bayesian optimization of wing length



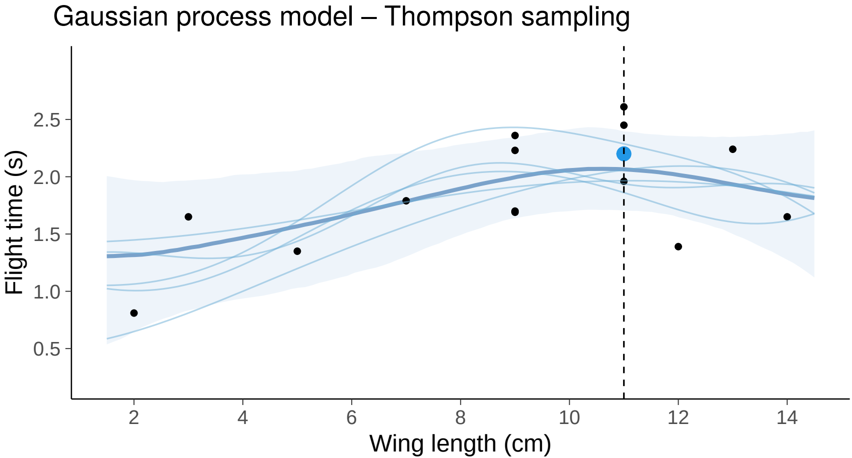
Bayesian optimization of wing length



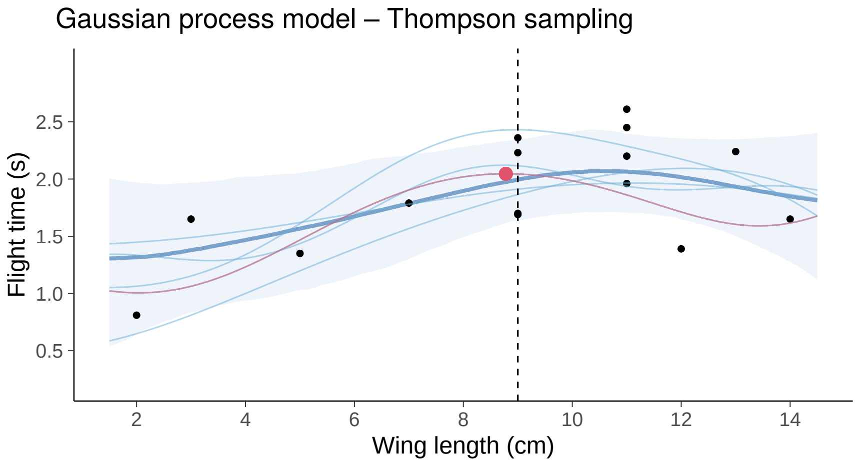
Bayesian optimization of wing length



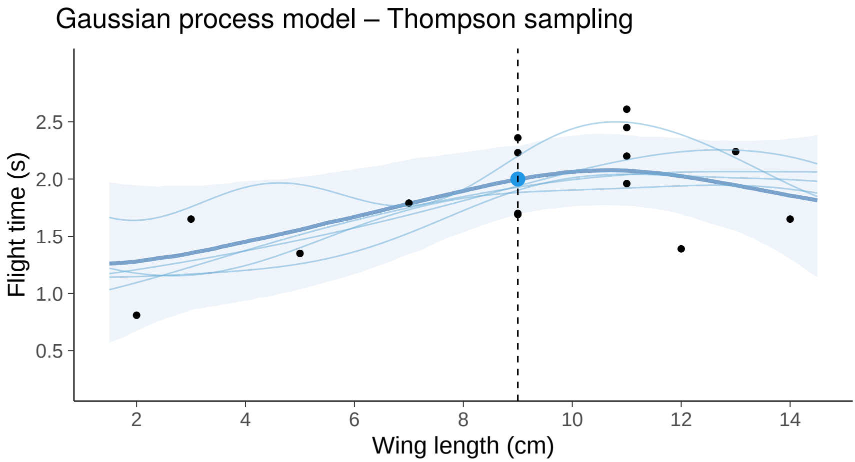
Bayesian optimization of wing length



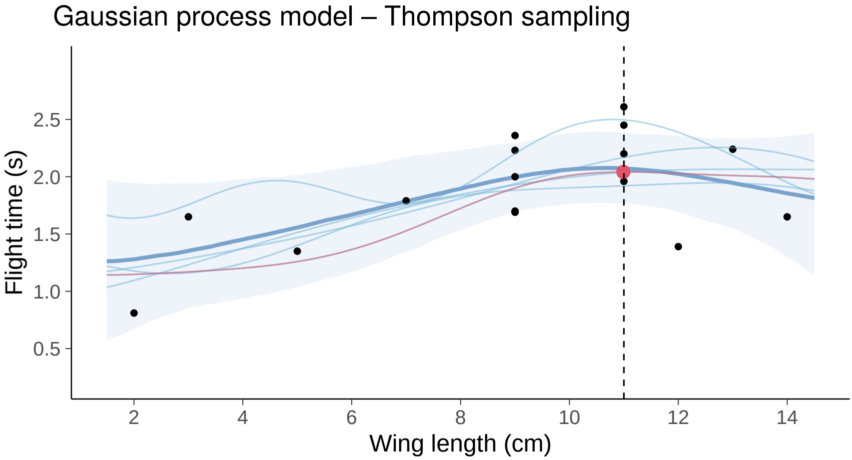
Bayesian optimization of wing length



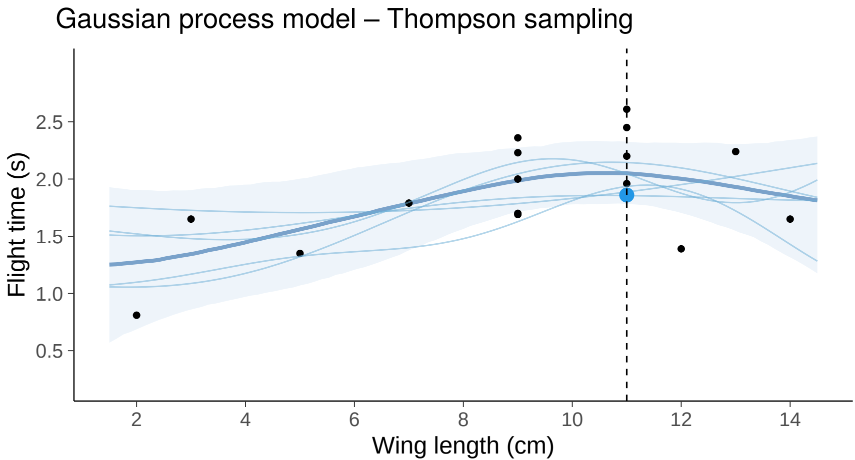
Bayesian optimization of wing length



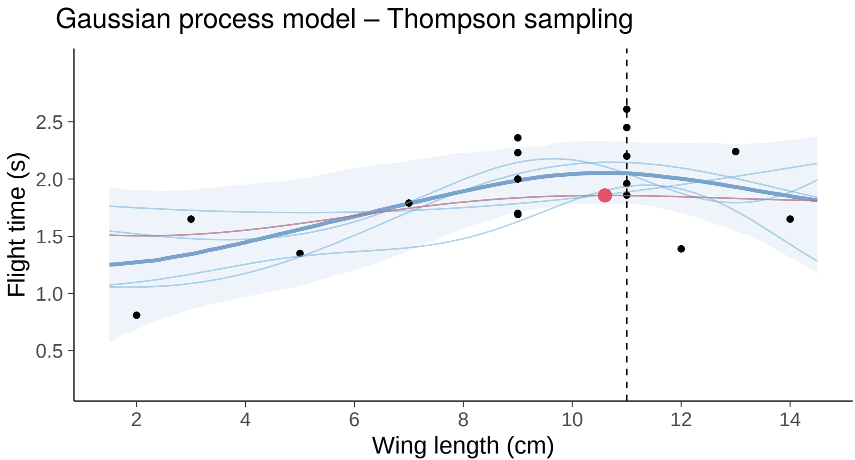
Bayesian optimization of wing length



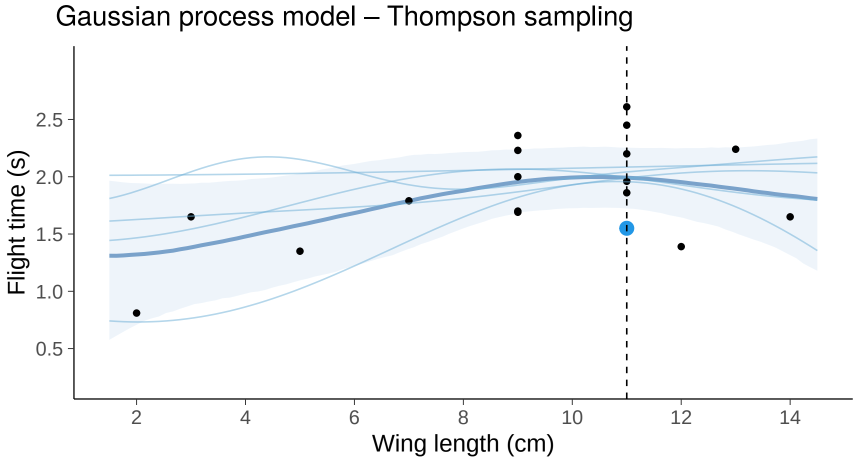
Bayesian optimization of wing length



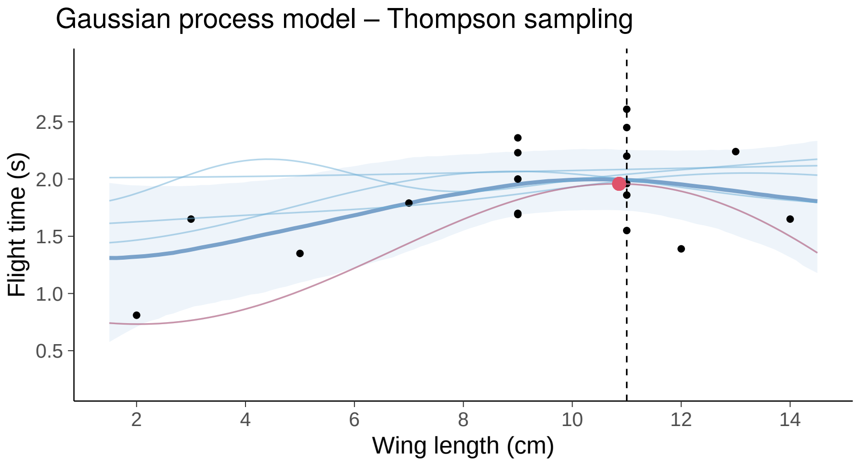
Bayesian optimization of wing length



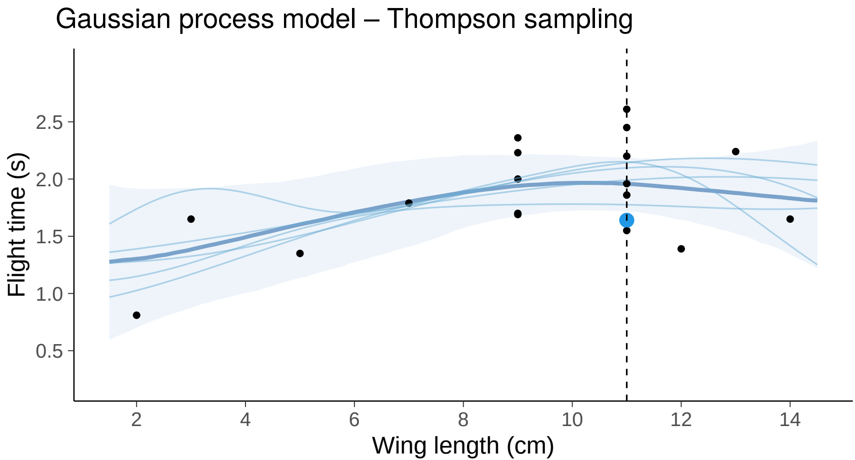
Bayesian optimization of wing length



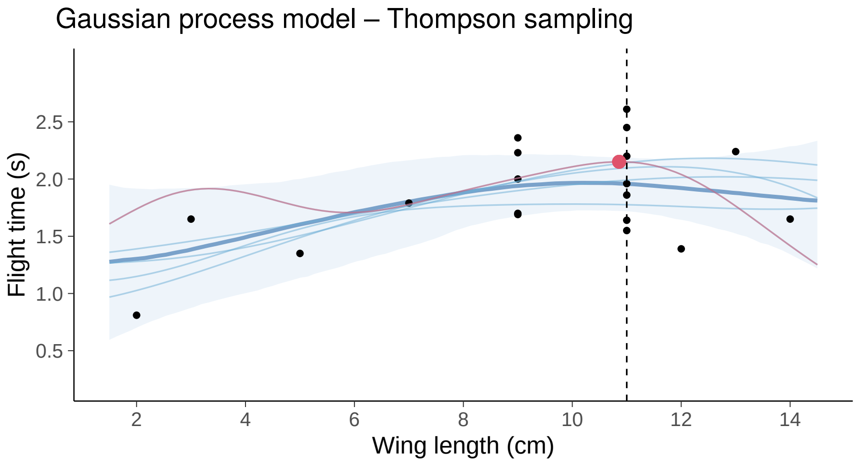
Bayesian optimization of wing length



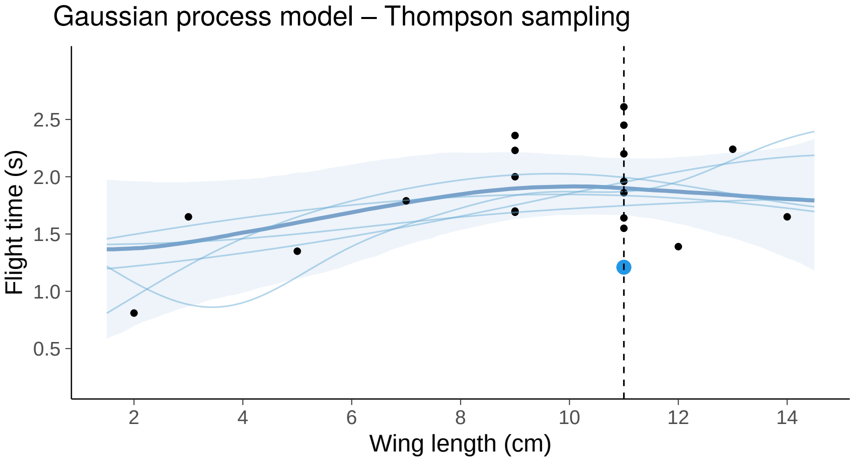
Bayesian optimization of wing length



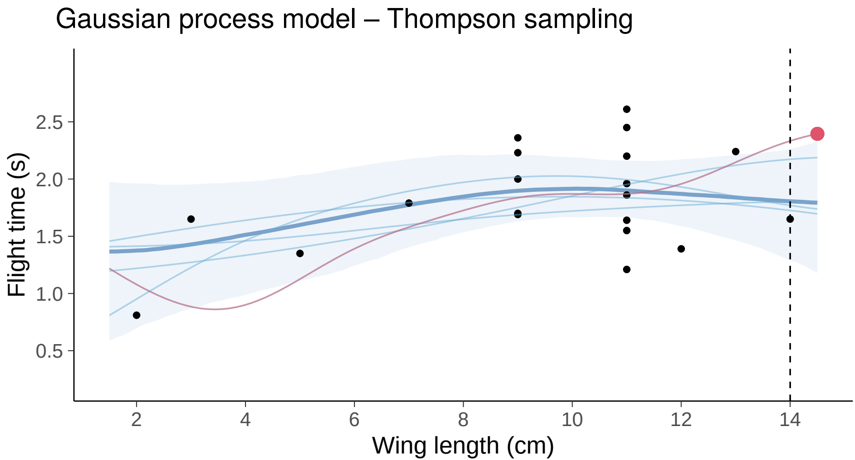
Bayesian optimization of wing length



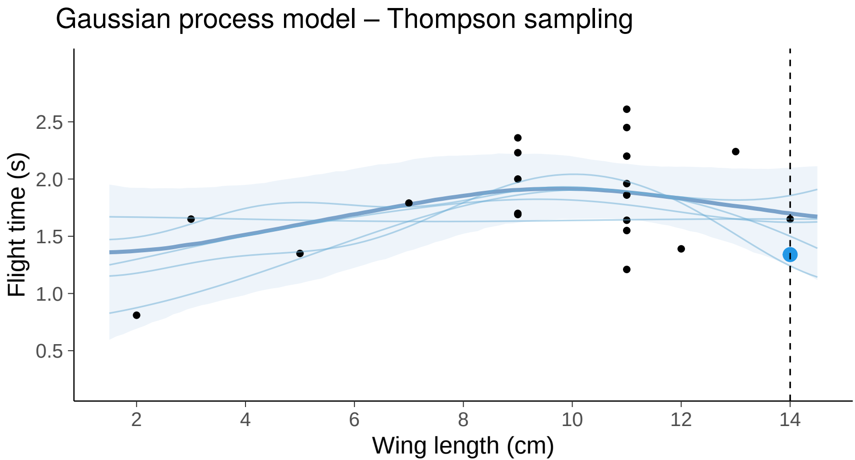
Bayesian optimization of wing length



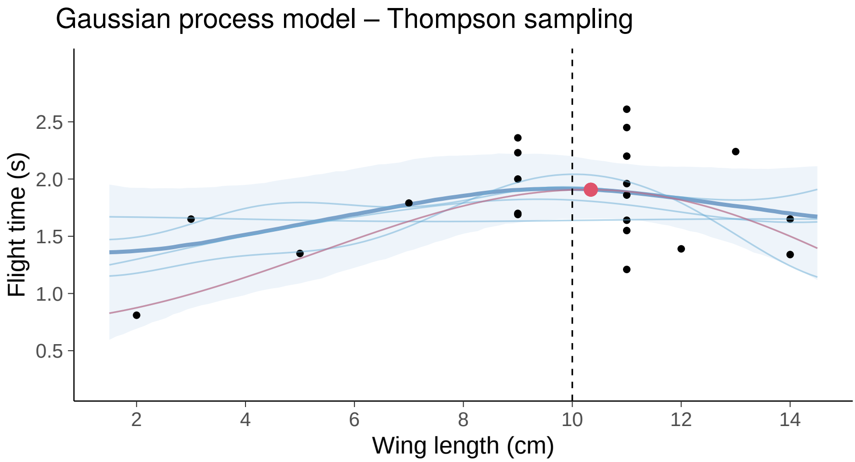
Bayesian optimization of wing length



Bayesian optimization of wing length

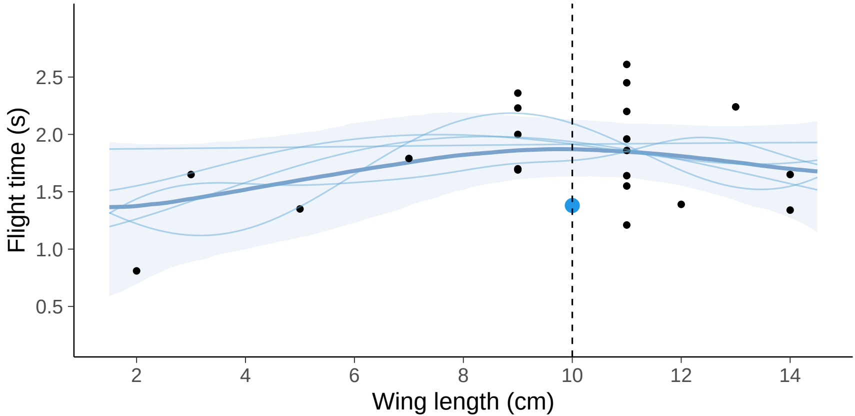


Bayesian optimization of wing length

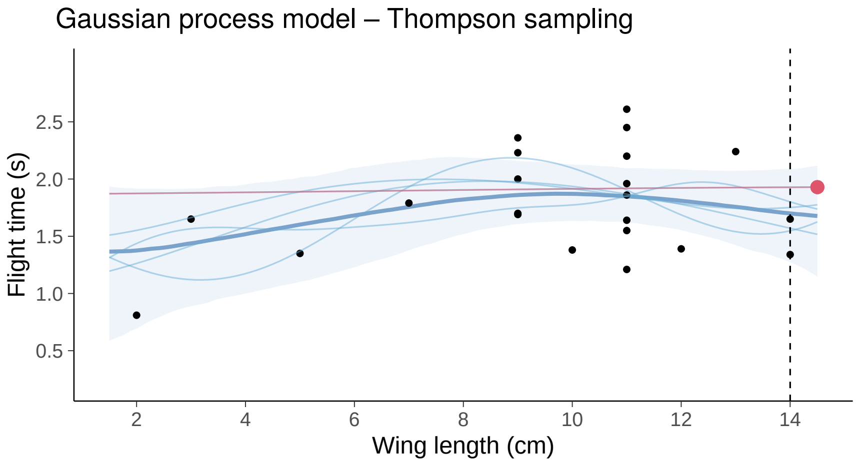


Bayesian optimization of wing length

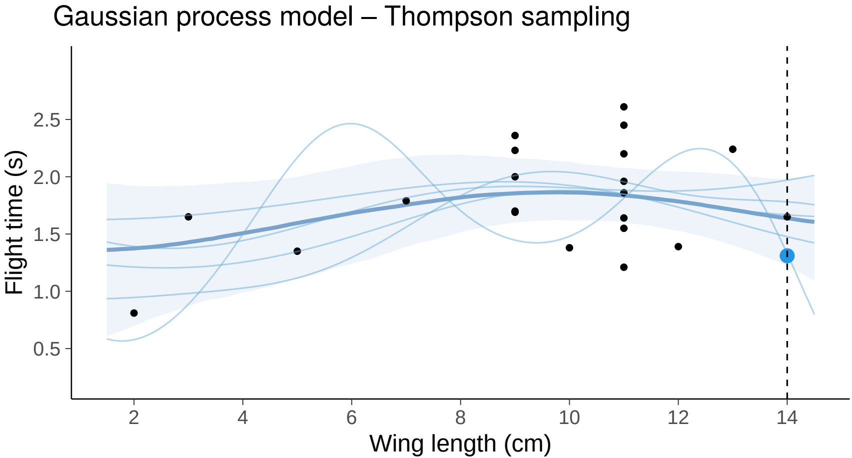
Gaussian process model – Thompson sampling



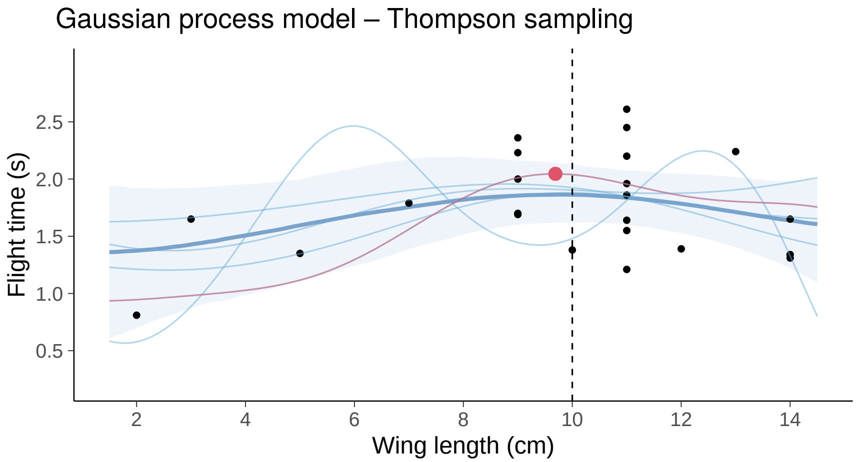
Bayesian optimization of wing length



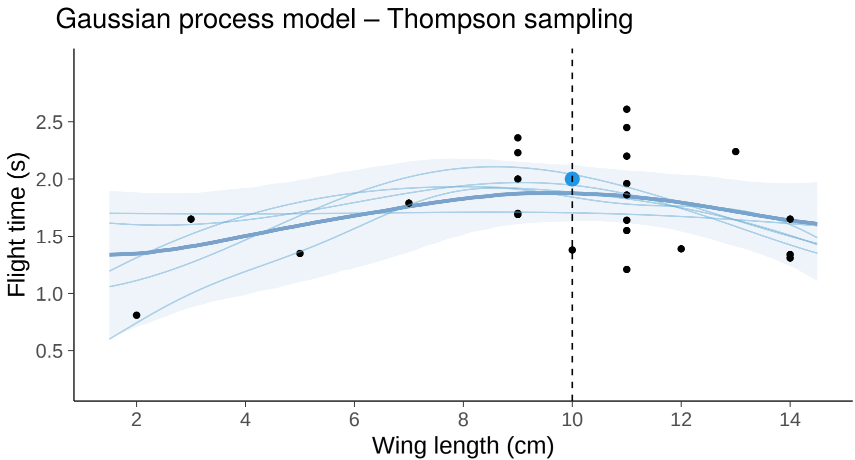
Bayesian optimization of wing length



Bayesian optimization of wing length

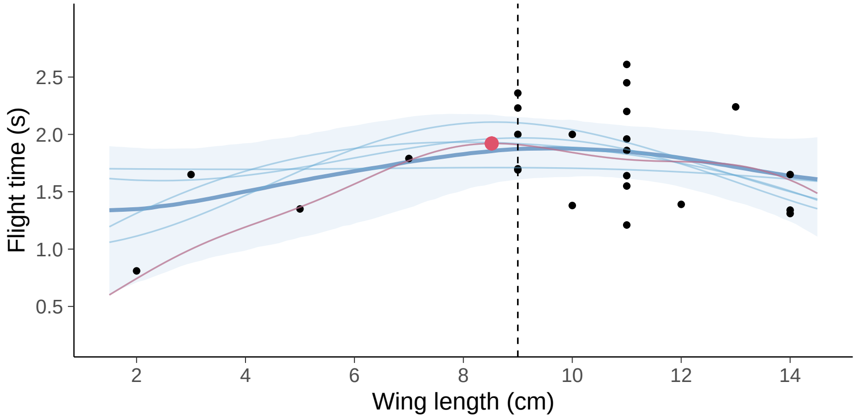


Bayesian optimization of wing length

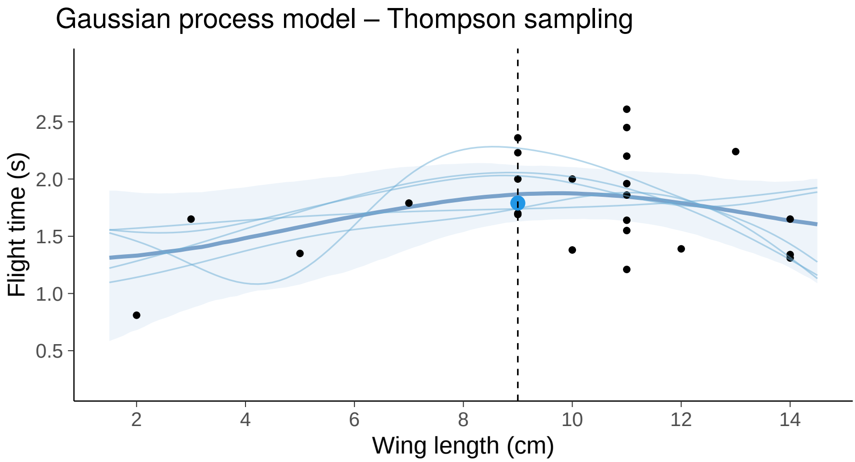


Bayesian optimization of wing length

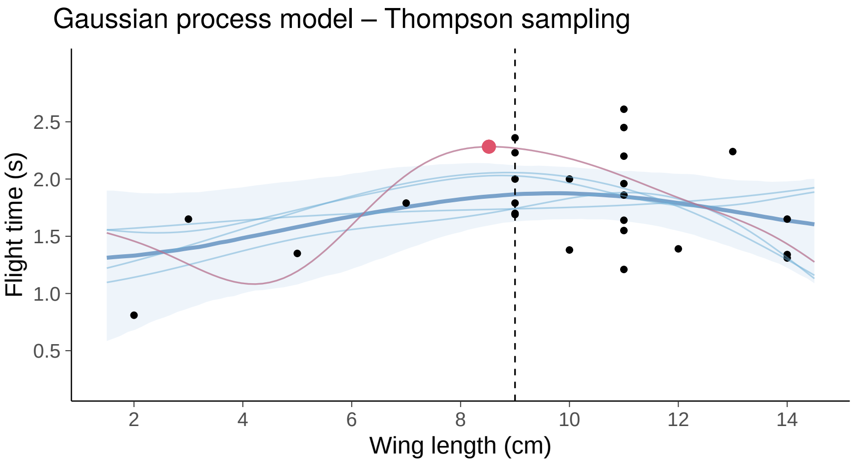
Gaussian process model – Thompson sampling



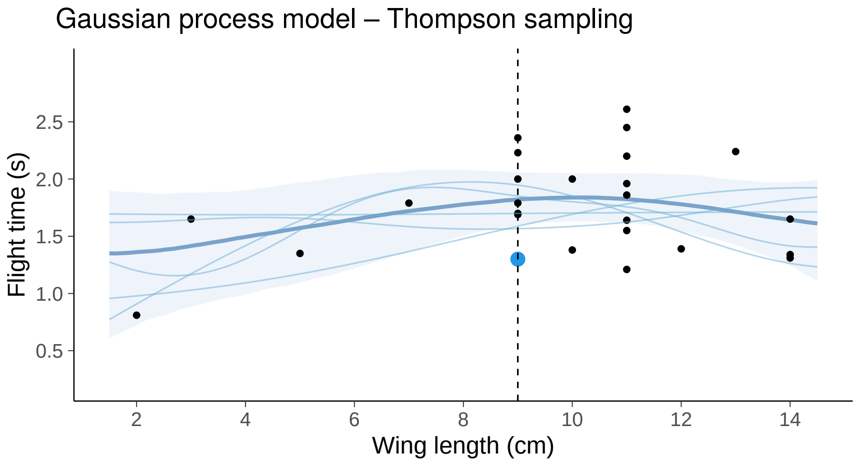
Bayesian optimization of wing length



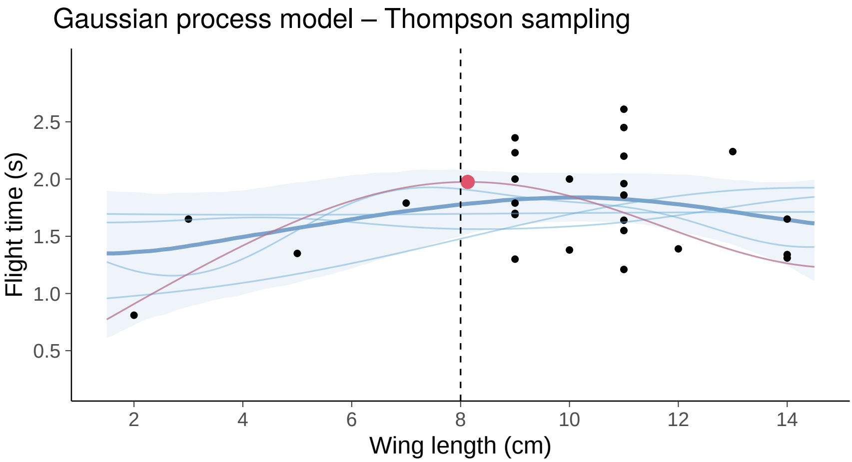
Bayesian optimization of wing length



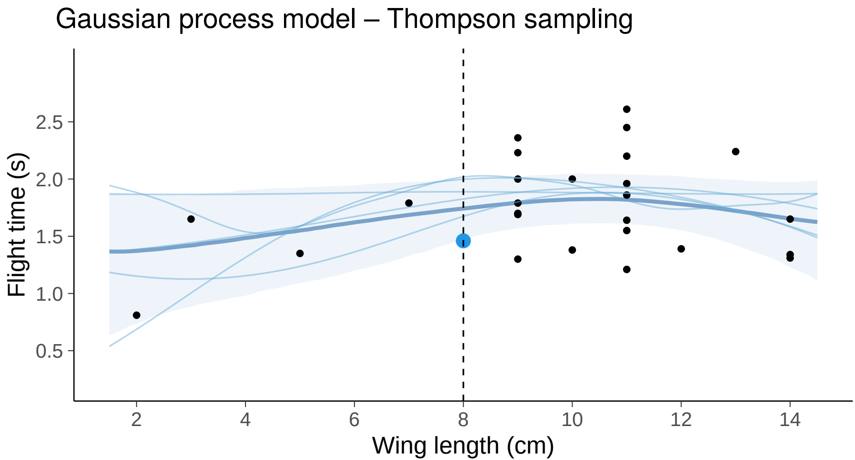
Bayesian optimization of wing length



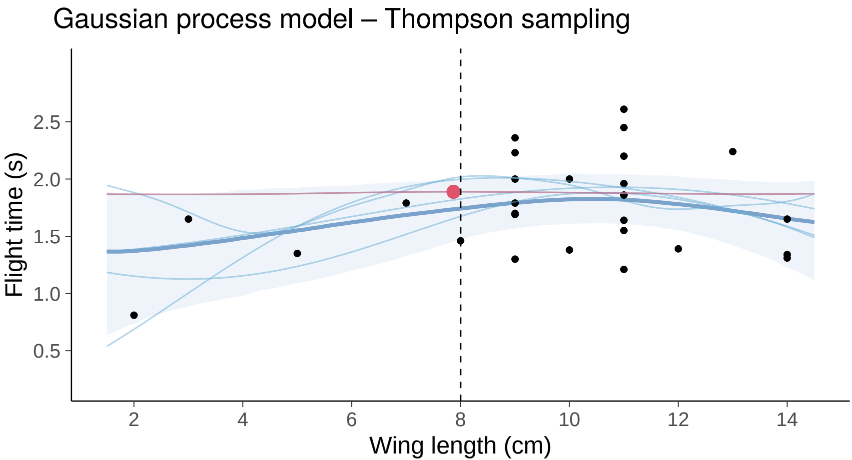
Bayesian optimization of wing length



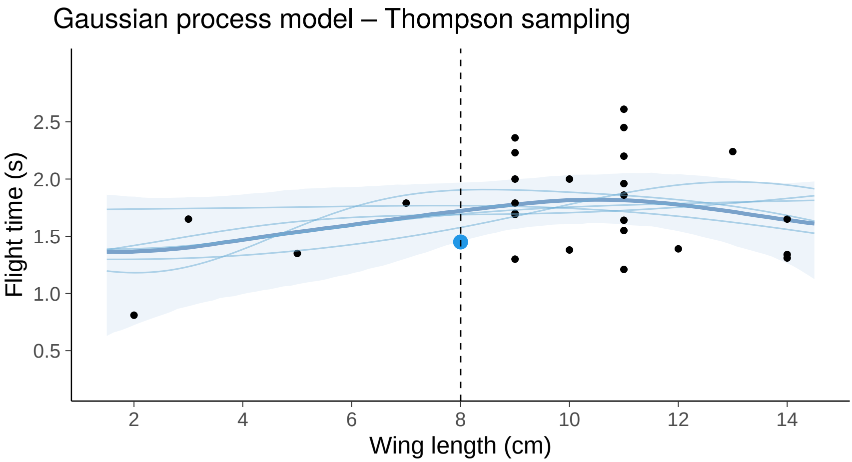
Bayesian optimization of wing length



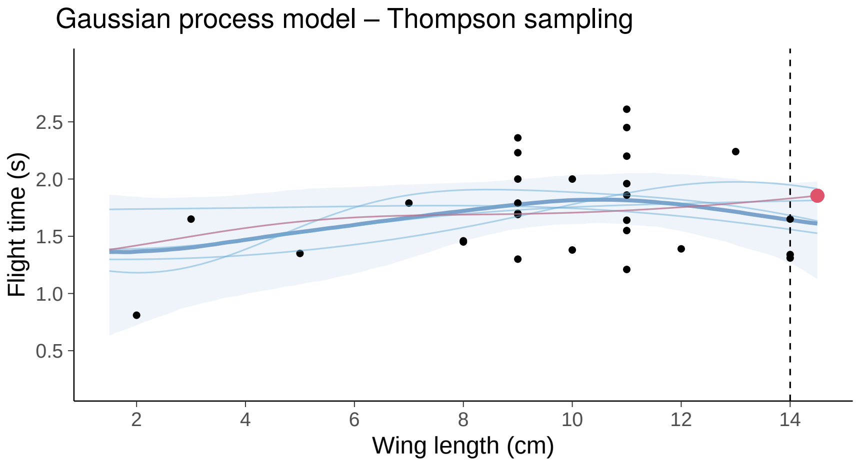
Bayesian optimization of wing length



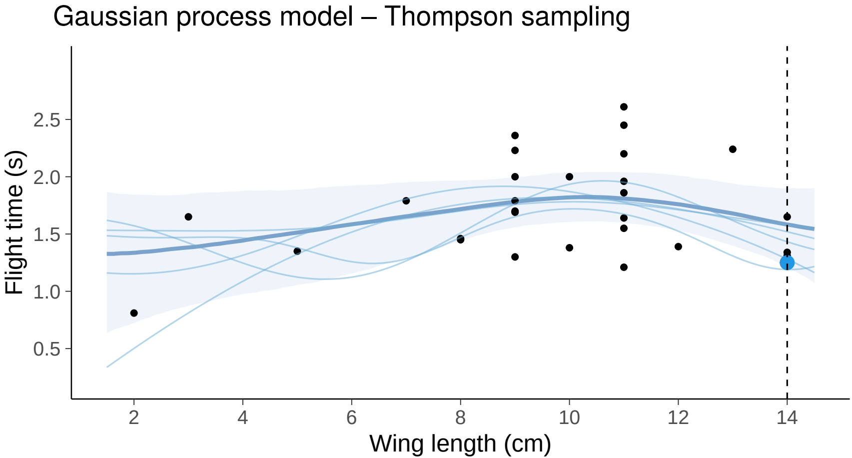
Bayesian optimization of wing length



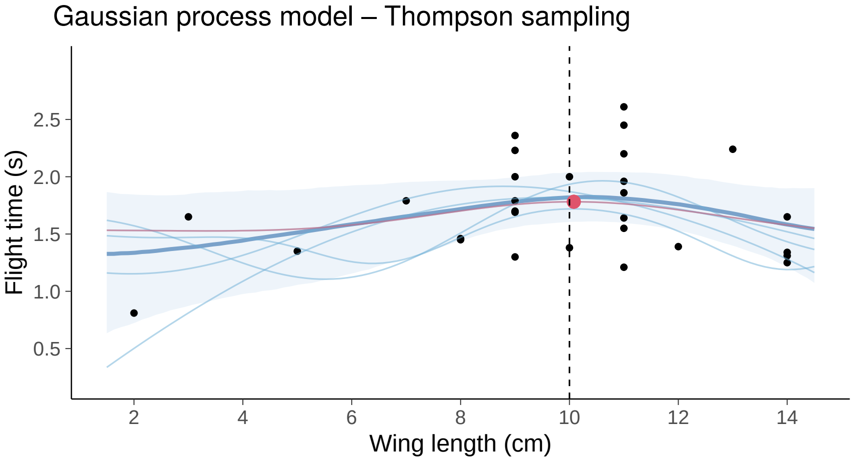
Bayesian optimization of wing length



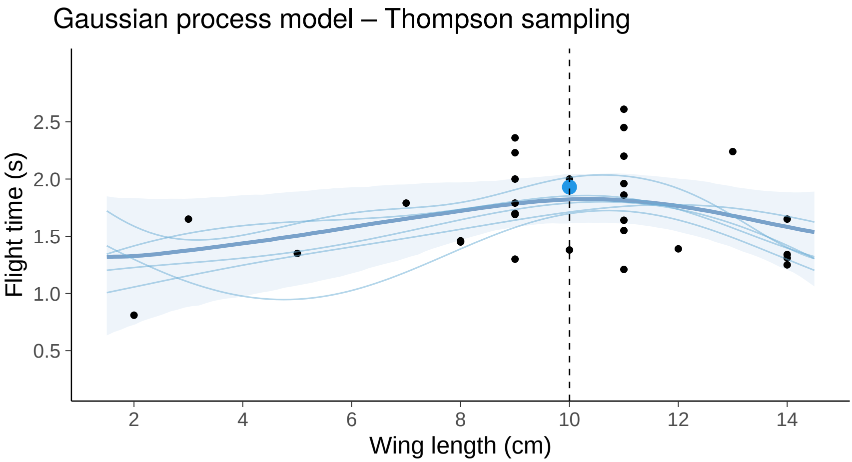
Bayesian optimization of wing length



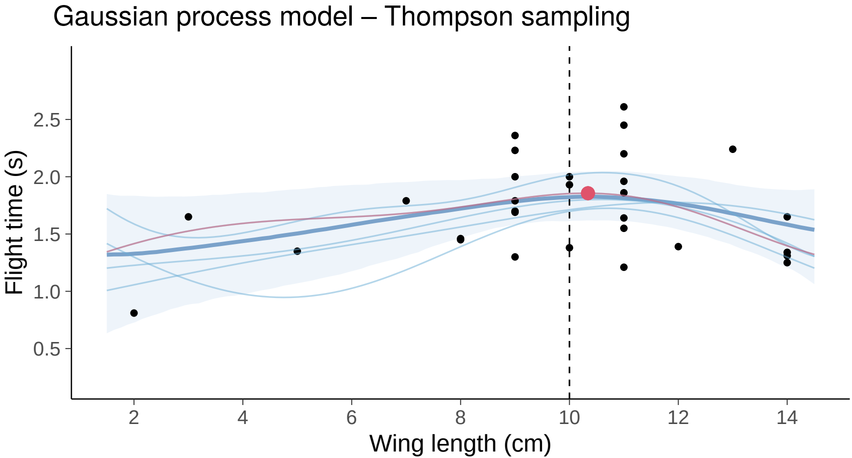
Bayesian optimization of wing length



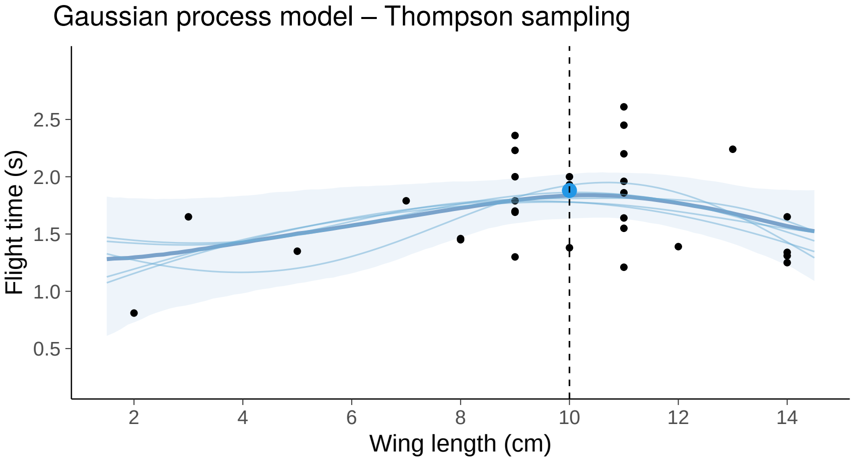
Bayesian optimization of wing length



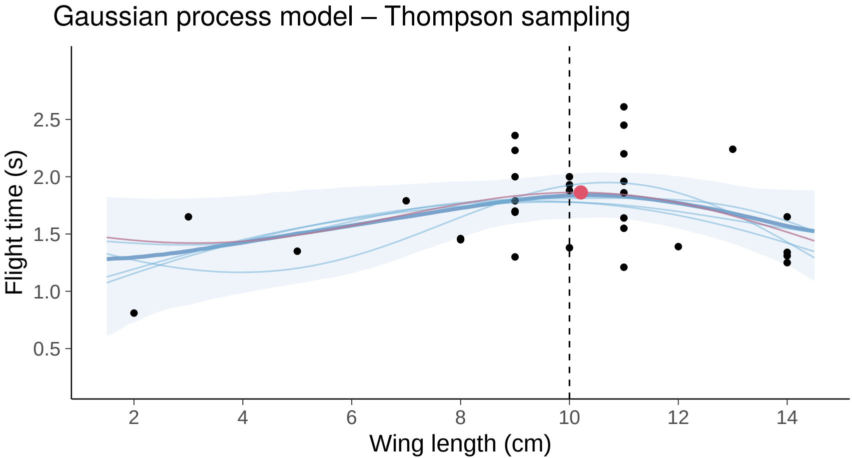
Bayesian optimization of wing length



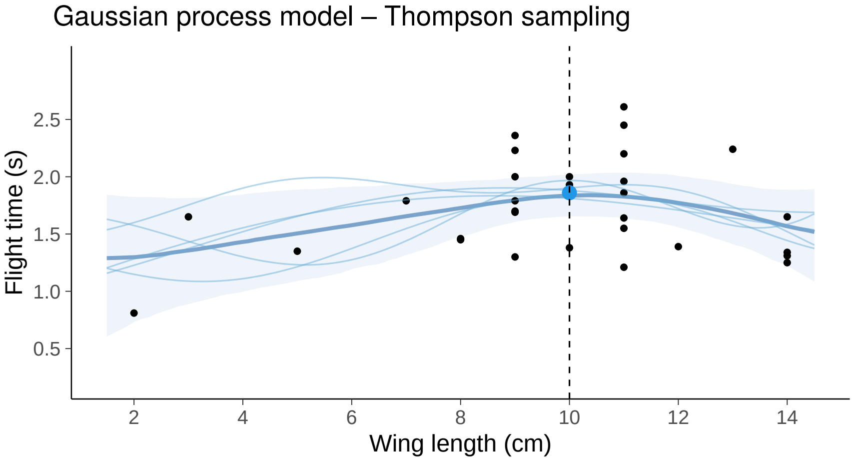
Bayesian optimization of wing length



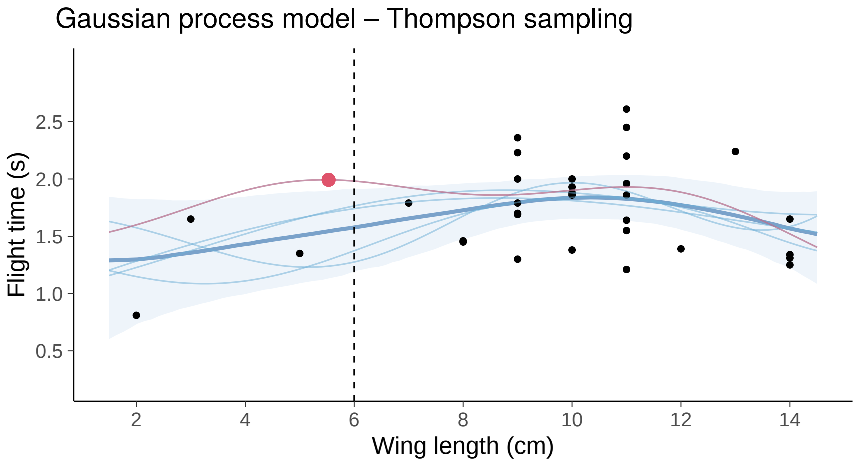
Bayesian optimization of wing length



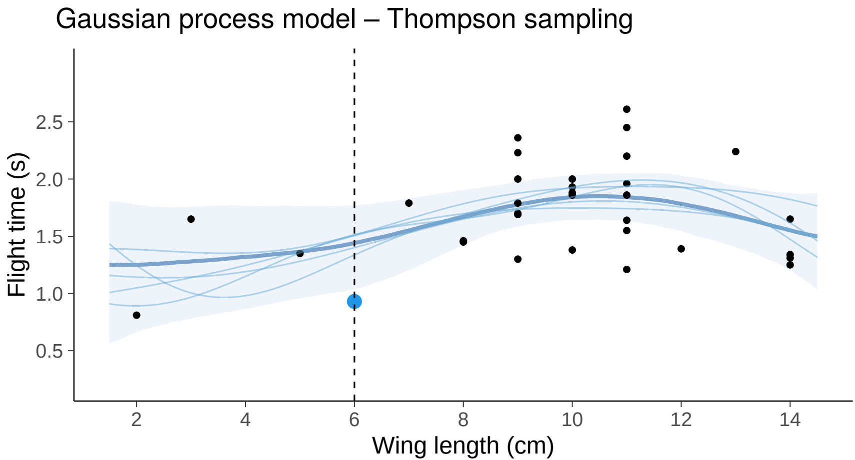
Bayesian optimization of wing length



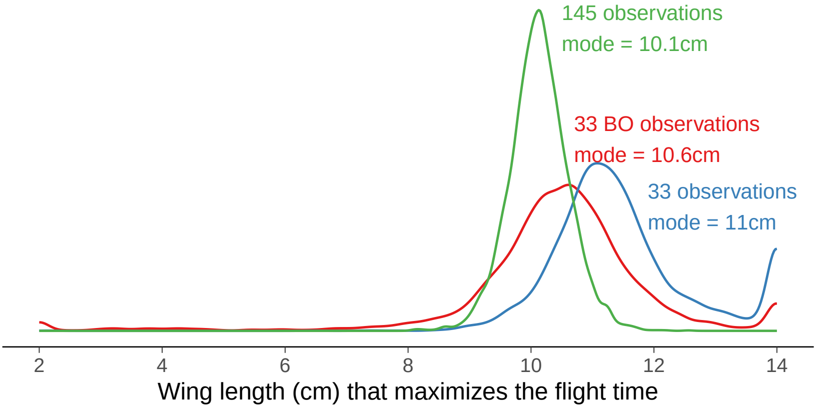
Bayesian optimization of wing length



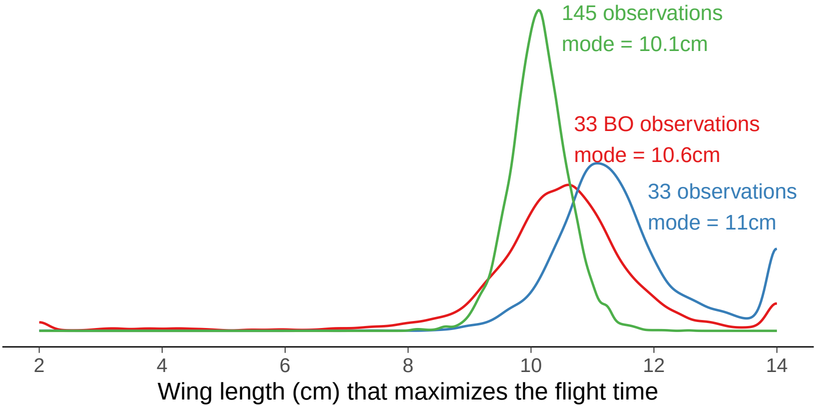
Bayesian optimization of wing length



Bayesian optimization of wing length



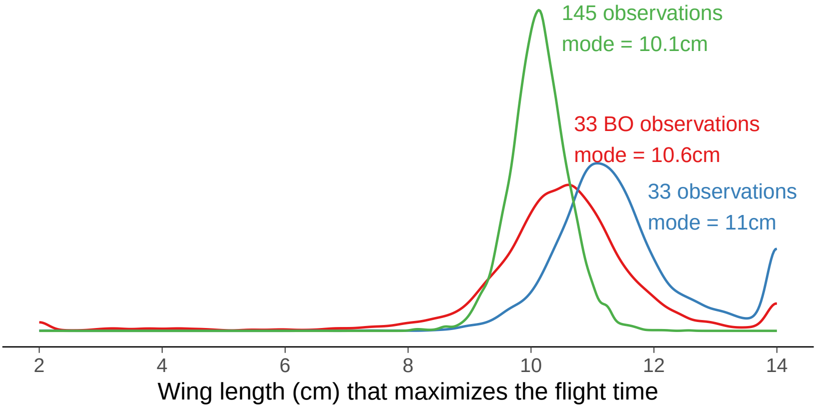
Bayesian optimization of wing length



33 BO obs. post. Wasserstein-1 distance ≈ 0.77

33 first obs. post. Wasserstein-1 distance ≈ 1.36

Bayesian optimization of wing length



33 BO obs. post. Wasserstein-1 distance ≈ 0.77

33 first obs. post. Wasserstein-1 distance ≈ 1.36

We obtain about 50% increase in efficiency

Examples of big Bayesian decision making success stories

- Bayesian optimization of ML algorithms
- Bayesian optimization of new medical molecules
- Bayesian optimization of new materials
- A/B testing
- Customer retention / satisfaction
- Marketing