Bayesian surprise as a tool for monitoring sensor networks

Wesley Brooks

USGS Wisconsin Water Science Center
and University of Wisconsin-Madison Department of Statistics
Madison, WI

wrbrooks@usgs.gov

May 3, 2012

science for a changing world

Table of Contents

Table of Contents

Intro

- Real-time monitoring: big data

Intro

- Real-time monitoring: big data
- Lots of new instruments are going in
  - That hardware needs to be maintained

Intro

- Real-time monitoring: big data
- Lots of new instruments are going in
  - That hardware needs to be maintained
- Each instrument is producing more data
  - Let's use that data to tell us when there's been a change that needs attention

Presenting Bayesian Surprise

- Automated
- Data-driven
- Detects unusual events in real-time data.

Presenting Bayesian Surprise

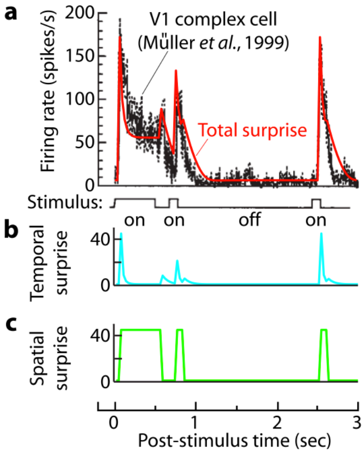
- Automated
- Data-driven
- Detects unusual events in real-time data.
- Basic idea: learn a model for the historical data and compare it to the newest incoming data.

Table of Contents

Table of Contents

Surprise theory

- Original idea of "Bayesian surprise" (2004):
  - Laurent Itti - University of Southern California Neuroscientist
  - Pierre Baldi - University of California-Irvine Computer Scientist
- Used to mimic human response to video images:



Surprise theory

- Adaptation to sensors:
  - Owen Langman's M.S. thesis - UW Limnology, 2009
- Uses identical surprise model (Gamma-Poisson) as Itti and Baldi

Surprise theory

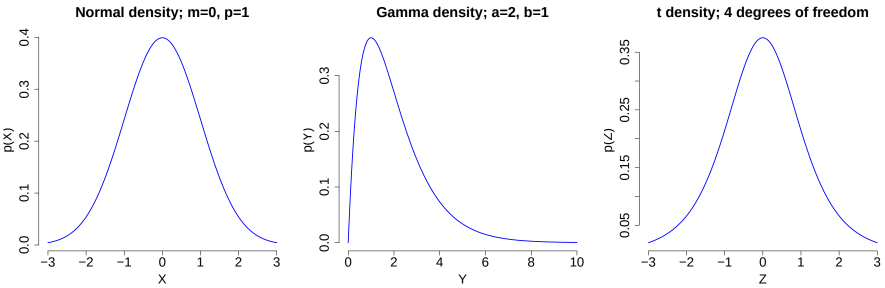
- Problems with original theory:
  - Ad-hoc "memory" parameter must be tuned manually
  - Cannot track mean and variance simultaneously
  - Technically only applicable to discrete data (e.g. counts)

Table of Contents

Bayesian statistics

Bayesian statistics views a probability distribution as representing our degree of belief. This idea can be applied both to our data and to the underlying data-generating model.

- Examples of the three distributions used in this work:
  - $X \sim \text{Normal}(\mu = 0, \tau = 1)$
  - $Y \sim \text{Gamma}(\alpha = 2, \beta = 1)$
  - $Z \sim \text{t}_{\nu=4}(\mu = 0, \sigma^2 = 1)$

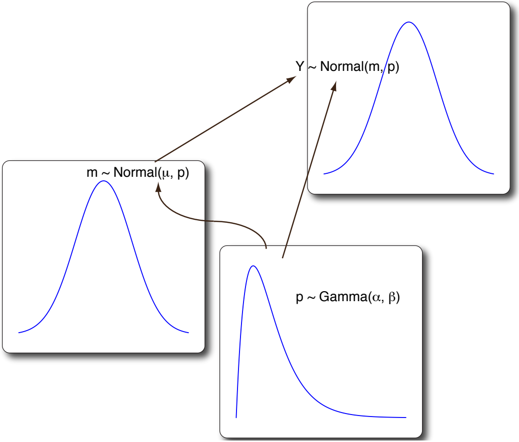


Hierarchical models

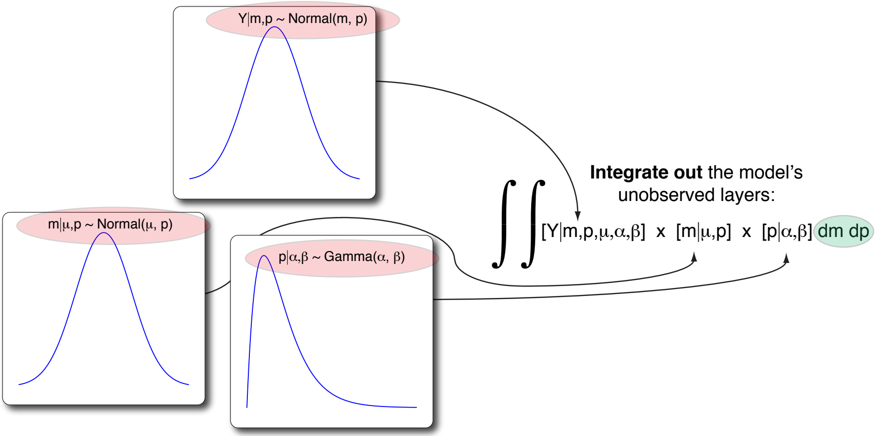
- A hierarchical model has more than one random element
- Randomness at one level feeds into the next

Hierarchical models

- A hierarchical model has more than one random element
- Randomness at one level feeds into the next



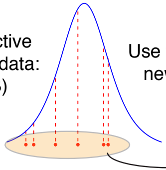
Hierarchical models



Hierarchical models

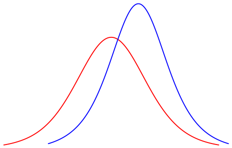
$$\int\int [Y|m,p,\mu,\alpha,\beta] \;\; \text{x} \;\; [m|\mu,p] \;\; \text{x} \;\; [p|\alpha,\beta]\; dm\; dp \;\; = \;\; [Y|\mu,\alpha,\beta]$$

Result is the predictive distribution for new data: $Y|\mu,\alpha,\beta \sim t(\mu,\alpha,\beta)$

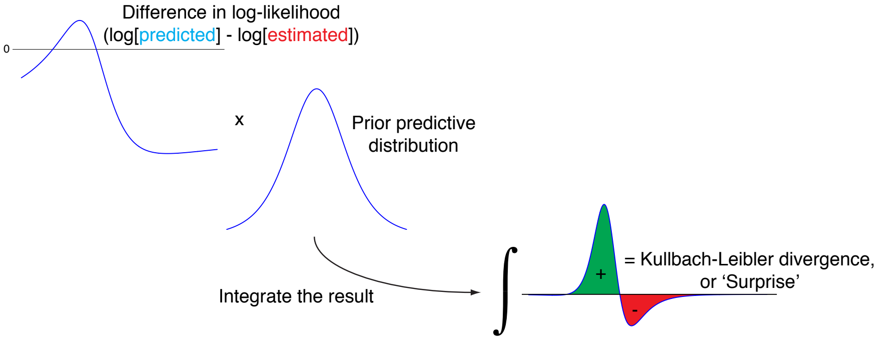


Use the data to estimate new values of μ, α, β

Compare the estimated distribiution to the predictive distribution (Two different t-distributions):



Calculate the surprise



Iterate the process

Use moving windows to iterate the process as new data comes in:

1.5	1.7	...	2.3	2.2	1.9	1.8	2.3	2.2	1.5	1.9	1.8	2.0	1.1	1.4

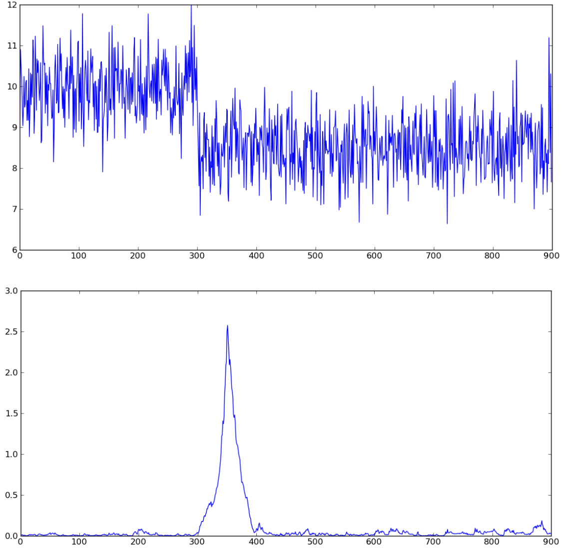
Prior window   Data window

Table of Contents

Table of Contents

Simulated surprise

Surprise generated by a sudden change in mean:



Simulated surprise

Surprise generated by a sudden change in variance:

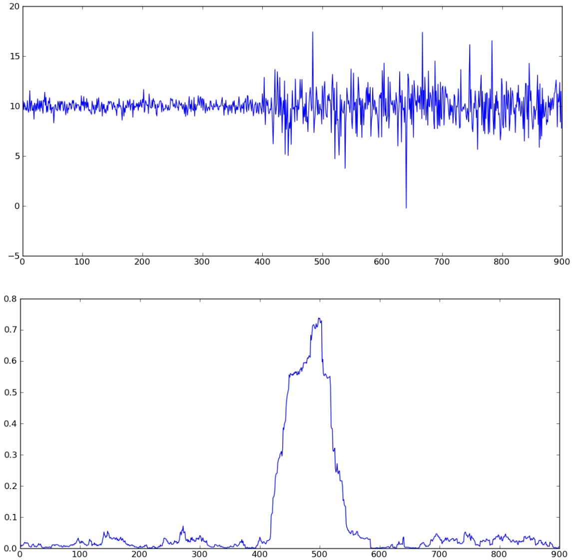
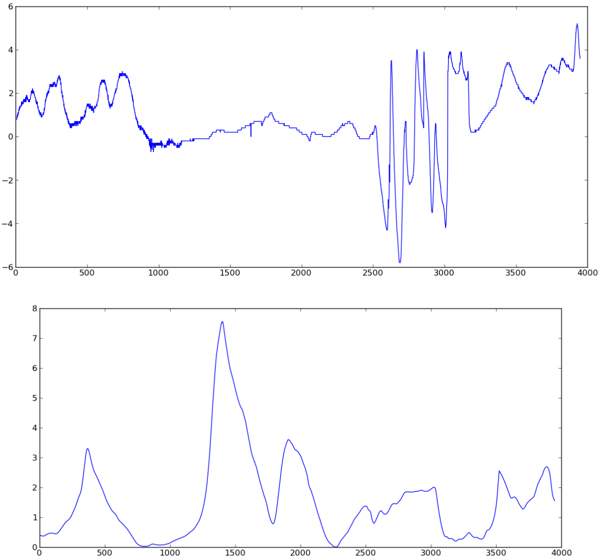


Table of Contents

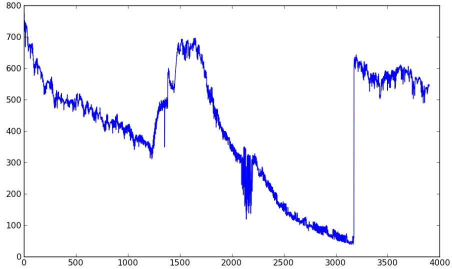
Field data

Pheasant Branch (Middleton, WI) water temp (Dec 2011 - Jan 2012):



Field data

Trout Lake LTER site (northern WI) CDOM (Nov. 2009):



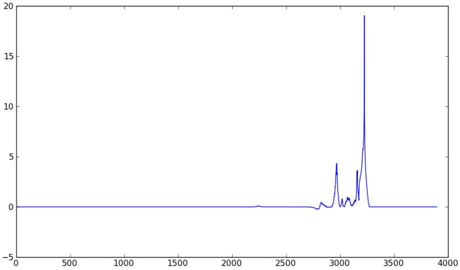
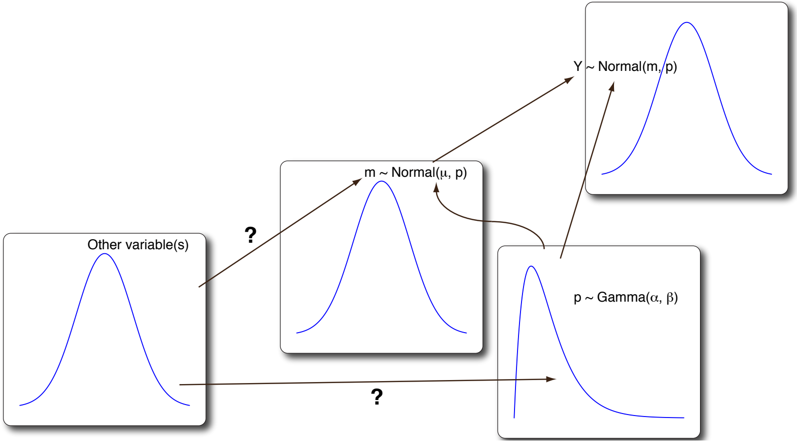


Table of Contents

Future directions

Add dependence on other variables (regression):



Conclusion

- Surprise is a data-driven tool that can help to quickly detect problems with real-time sensors and therefore improve the up-time of a monitoring effort.