# A non-linear method for channel selection

Alison Fowler and Peter Jan van Leeuwen
University of Reading, UK

ESA living planet symposium
Friday 13th September 2013

# Introduction

- Satellite data needs to be thinned so that it can be efficiently stored, transmitted and assimilated.
- Current methods rely on the assumption that the observation operator (the mapping between observation and state space) can be linearised accurately.
  - however satellite data is often a highly non-linear function of the atmospheric state

# Introduction

- Satellite data needs to be thinned so that it can be efficiently stored, transmitted and assimilated.
- Current methods rely on the assumption that the observation operator (the mapping between observation and state space) can be linearised accurately.
  - however satellite data is often a highly non-linear function of the atmospheric state
- Within this talk I'll discuss a new method which does not rely on the linearisation of the observation operator.

## Mutual information

- The first step is to choose an objective measure to quantify the information in the observed data.
- Mutual information measures the reduction in the uncertainty of the state when an observation is made:

$$MI = \int p(\mathbf{y}) \int p(\mathbf{x}|\mathbf{y}) \ln \frac{p(\mathbf{x}|\mathbf{y})}{p(\mathbf{x})} d\mathbf{x} d\mathbf{y}. \tag{1}$$

  - where $\mathbf{x}$ is the state vector, $\mathbf{y}$ is the observation vector.

# Linear approximation

- In many studies of impact for satellite data a linear approximation of *MI* has been used which is consistent with the variational data assimilation scheme (e.g. Rabier et al. 2002)).
- In this case, an analytical expression for mutual information can be given when the prior and likelihood are both Gaussian:

$$MI = \ln |\mathbf{B}^{-1}\mathbf{P_a}|, \tag{2}$$

  - where $\mathbf{B}$ is the error covariance matrix of the prior estimate of the state and $\mathbf{P_a}$ is the error covariance matrix of the analysis estimate of the state.

# Linear approximation

- In many studies of impact for satellite data a linear approximation of $MI$ has been used which is consistent with the variational data assimilation scheme (e.g. Rabier et al. 2002)).
- In this case, an analytical expression for mutual information can be given when the prior and likelihood are both Gaussian:

$$MI = \ln |\mathbf{B}^{-1}\mathbf{P_a}|, \tag{2}$$

  - where $\mathbf{B}$ is the error covariance matrix of the prior estimate of the state and $\mathbf{P_a}$ is the error covariance matrix of the analysis estimate of the state.
- Research questions:
  - How does the linear approximation affect the estimate of information in satellite data?
  - What impact could this have on channel selection?

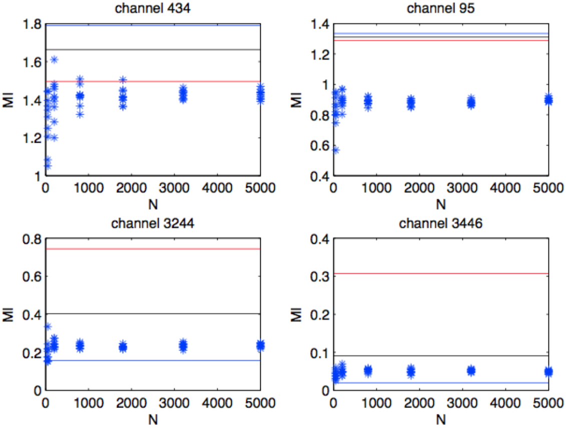
# A non-linear estimate

- It is not possible to give an analytical expression for $MI$ when the observation operator is non-linear.
- But can approximate $MI$ using a sampling method which accounts for the non-linear observation operator.
- This involves taking $M$ samples from the likelihood, $p(\mathbf{y}|\mathbf{x})$, and $N$ samples from the prior, $p(\mathbf{x})$.

$$MI = \sum_{i=1}^{M} p(\mathbf{y}_i) \left( \sum_{j=1}^{N} w_{i,j} \ln(N w_{i,j}) \right), \tag{3}$$

  - where $w_{i,j} = \frac{p(\mathbf{y}_i|\mathbf{x}_j)}{p(\mathbf{y}_i)}$ represents the posterior distribution.

# Comparison of linear (solid lines) and non-linear (stars) estimate of mutual information



# Linearisation error

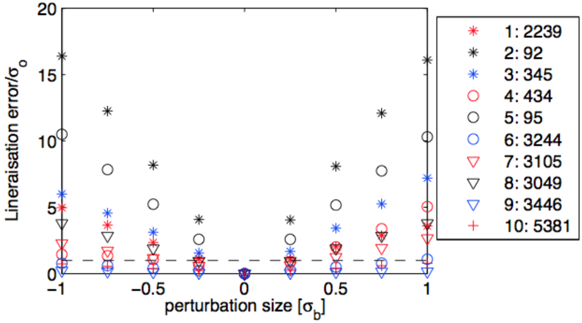



Figure : Linearisation error$=h(x+\delta x)-h(h)-\mathbf{H}\delta x$.

# Channel selection algorithm

1. Remove channels which are known to be poorly modelled by the radiative transfer model (RTTOV) from the channels available for selection,
2. Calculate *MI* for the remaining channels.
3. Select the channel with the greatest *MI*.
4. Update the prior given the information from this channel choice.
5. Repeat steps 2-4 until the desired number of channels have been selected.

# Initial results for IASI data simulated by RTTOV

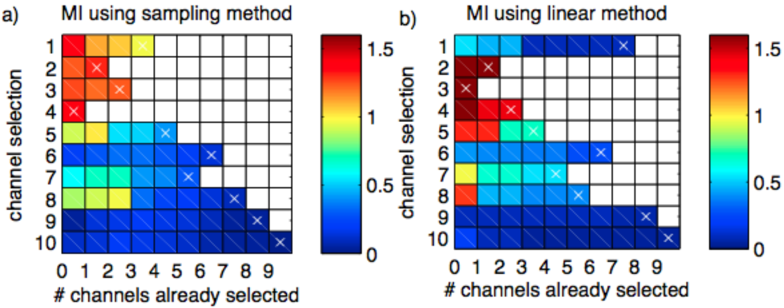



Figure : Channel selection for ten channels shown to be important by Collard 2007, $M = 100$, $N = 5000$.

# The effective sample size

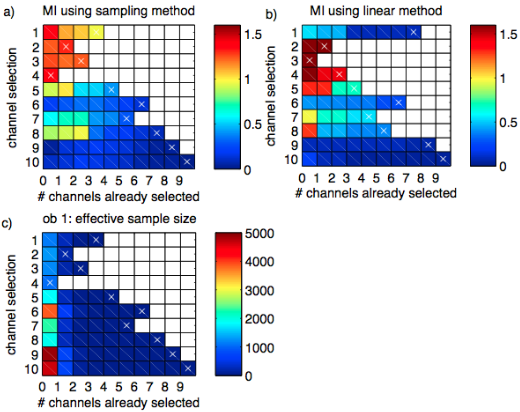


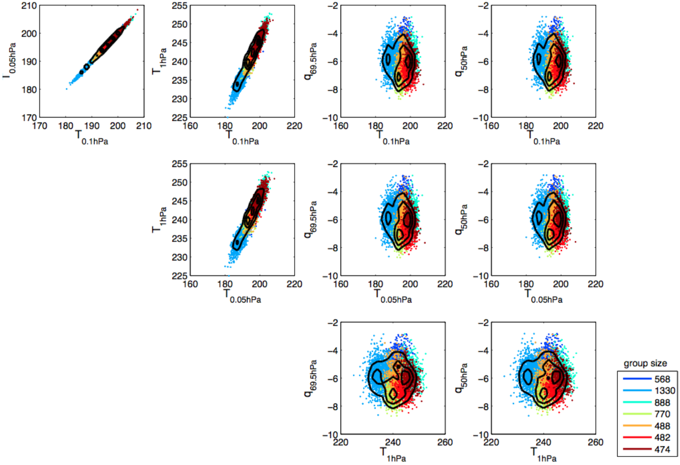

Figure : effective sample size$=(\sum_i^N 1/w_i^2)^{-1}$

# Gaussian mixture resampling

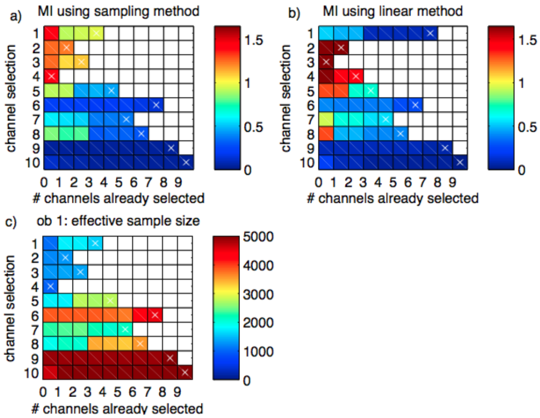
- To keep the effective sample size large we have decided to resample from posterior distribution after each channel selection has been made.
- Each member of the new sample has equal weight and so the effective sample size is restored to $N$.
- To resample from the posterior we need to make some assumptions about its distribution.
  - However we wish to maintain any non-Gaussian structure caused by the non-linear observation operator.
  - This can be done by fitting a Gaussian mixture to the weighted sample with the number of components, $G$ dependent on the exact structure.

$$p(\mathbf{x}|\mathbf{y}) \approx \sum_{i=1}^{G} a_i N(\mu_i, \Sigma_i) \quad (4)$$

# Gaussian mixture resampling



# Gaussian mixture resampling



## Summary

- Satellite observations are a non-linear function of the state variables of interest. As such a linear approximation to their information content may be misleading.
- A sampling approximation to *MI* has been demonstrated which is free from assumptions about linearity. This has shown that for some channels the linear approximation to *MI* is indeed poor.
- Although this estimate is free from assumptions about the linearity it does suffer from under sampling when the region of uncertainty is small.
- Resampling from the posterior distribution after each channel is selected can alleviate this problem.

# References


- Rabier et al., 2002: Channel selection methods for IASI radiances. **Q. J. R. Met. Soc.**, *128*, 1011-1027.
- Collard, 2007: Selection of IASI channels for use in NWP. **Q. J. R. Met. Soc.**, *133*, 1977-1991.

# References


- Rabier et al., 2002: Channel selection methods for IASI radiances. **Q. J. R. Met. Soc.**, *128*, 1011-1027.
- Collard, 2007: Selection of IASI channels for use in NWP. **Q. J. R. Met. Soc.**, *133*, 1977-1991.


*Thank you for listening, any questions?*

# Weighting functions

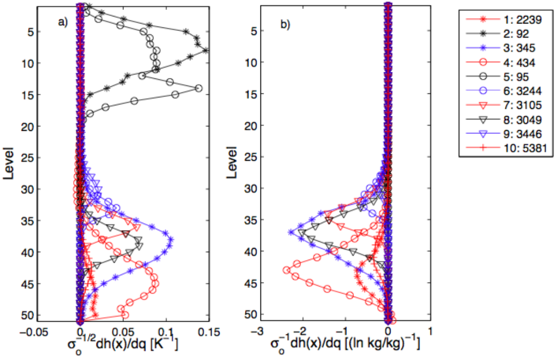



Figure 4: Weighting functions normalised by the observation error standard deviation, for the ten channels used in the channel selection. a) Sensitivity to changes in temperature. b) Sensitivity to changes in humidity.