

Invariance d'échelle dans le trafic Internet: aspects méthodologiques



Steve UHLIG

Département d'Ingénierie Informatique.
Université catholique de Louvain

Agenda

- **Invariance d'échelle**
- **Invariance d'échelle de 2^{ième} ordre**
- **Invariance d'échelle d'ordre supérieur**
- **Conclusions**

Invariance d'échelle

Invariance d'échelle

- **Définition**: une statistique se comporte comme une loi puissance du temps
- Si les incréments du processus $X(t)$ sont remplacés par les coefficients d'ondelette $d(j,k)$, l'invariance d'échelle peut être étudiée à travers le comportement de $E/d(j,k)^q$
- La loi puissance peut être linéaire ou non-linéaire:
 - auto-similarité : $E/d(j,k)^q \propto \exp(qH \ln(2^j))$
 - multi-fractal : $E/d(j,k)^q \propto \exp(H(q) \ln(2^j))$

Modèles d'invariance d'échelle ≠ modèles physiques

- L'invariance d'échelle est une *propriété statistique* du signal.
- Les modèles d'invariance d'échelle ne doivent pas nécessairement avoir le moindre rapport avec la *physique* du signal.
- L'invariance d'échelle est *un formalisme* permettant de décrire un signal.

Invariance d'échelle de 2^{ième} ordre

Invariance d'échelle de 2^{ième} ordre

- Identification de l'invariance d'échelle de 2^{ième} ordre à l'aide du *logscale diagram* (LD)

[VA99] :

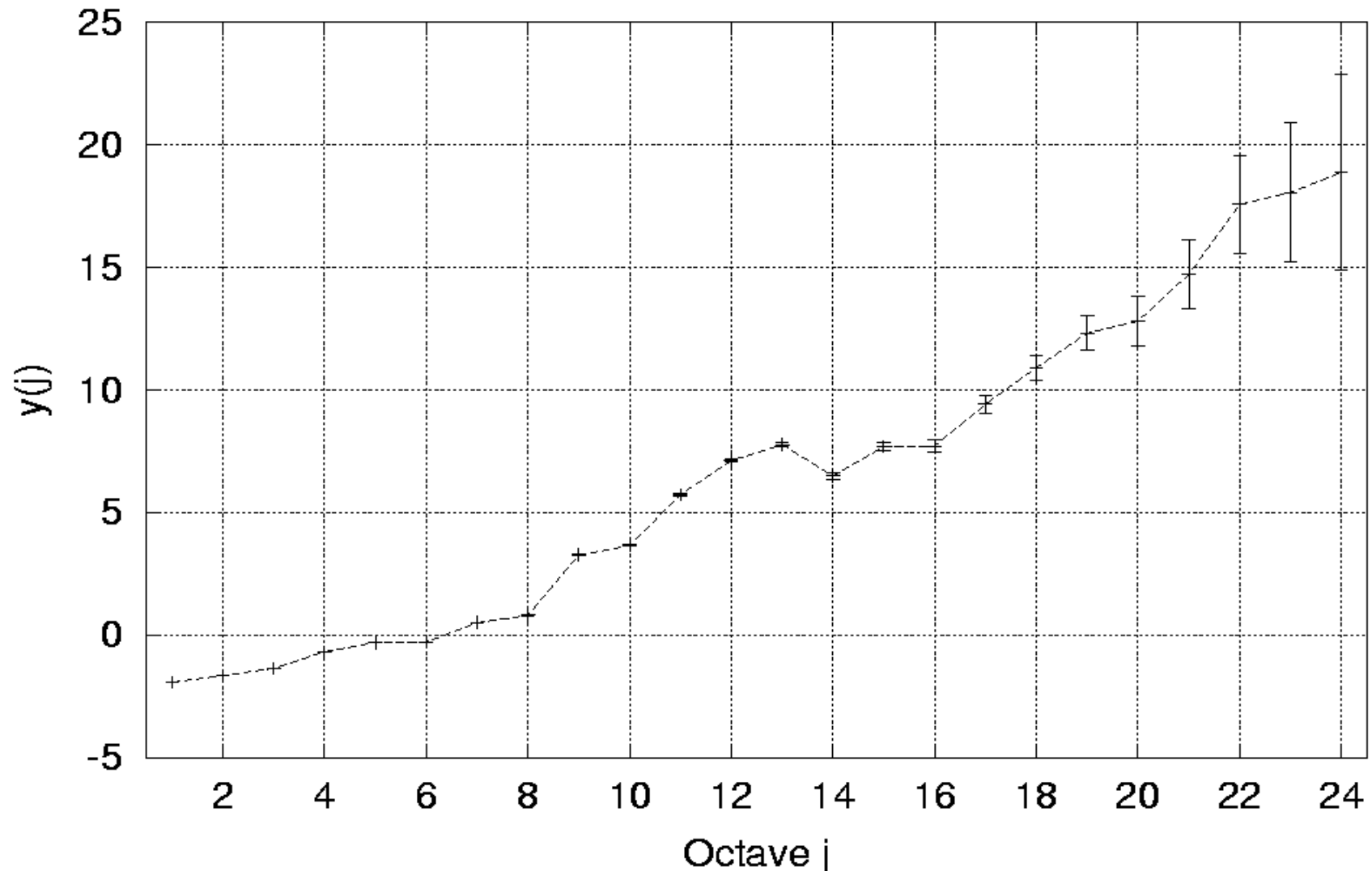
– afficher $\log_2(\mu_j)$ en fonction de j où

$$\mu_j = \sum_k |d(j,k)|^2$$

- L'identification de l'invariance d'échelle nécessite l'alignement de $\log_2(\mu_j)$ ainsi que des intervalles de confiance !

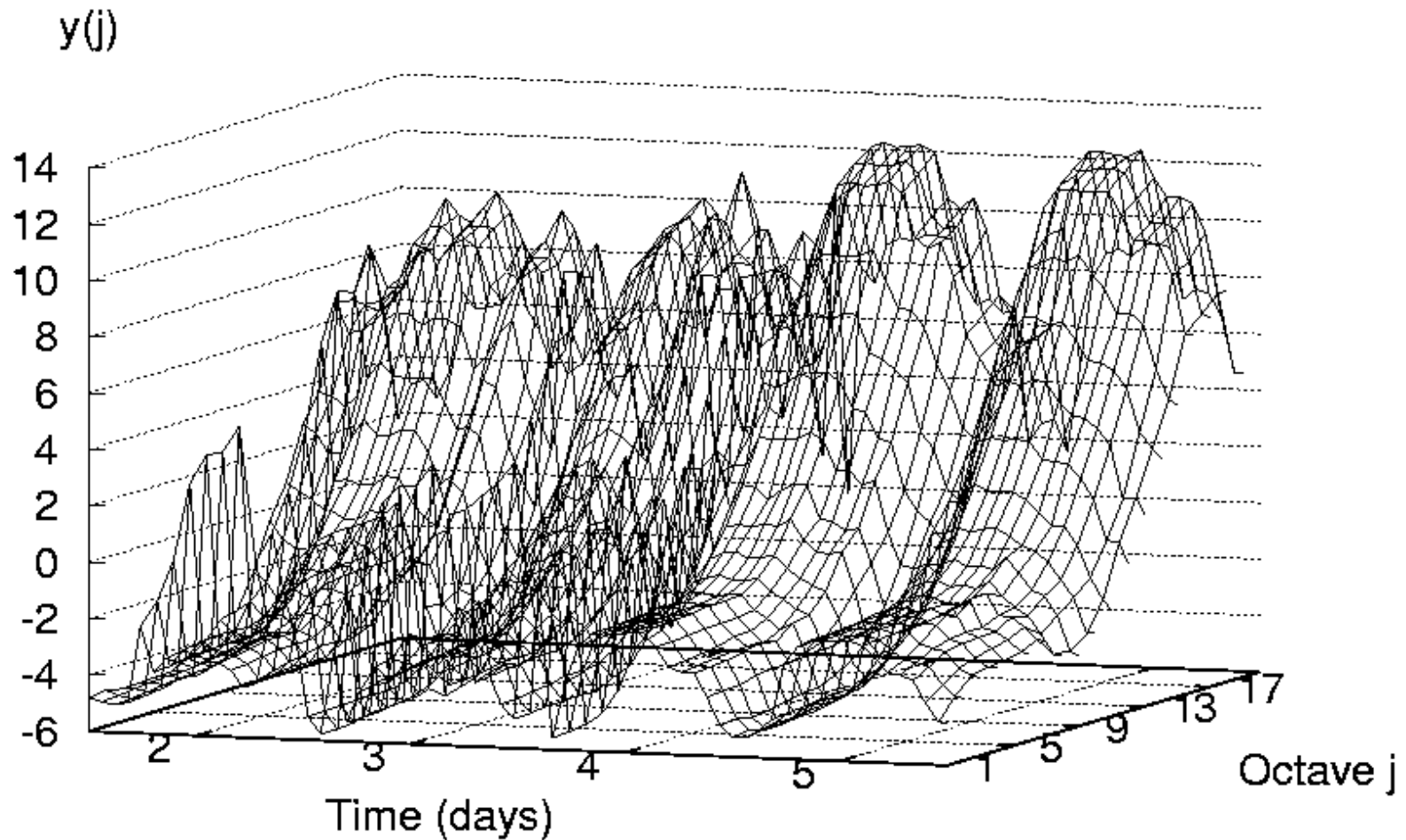
Faire confiance au LD

Logscale diagram for Auckland incoming flow arrivals



Ne faites pas confiance au LD !

3D-LD for Auckland incoming flow arrivals



Invariance d'échelle d'ordre supérieur

Composants dans l'invariance d'échelle

- **Deux régimes d'invariance d'échelle ont été identifiés dans le trafic Internet:**
 - **multi-fractal aux échelles de temps fines [FGW98]**
 - **Dépendance à longue portée aux grandes échelles de temps [WPRT02]**
- **La cause physique du comportement multi-fractal pourrait être la cascade de TCP [FGW98], celle de la dépendance à longue portée les queues lourdes couplées au comportement ON/OFF [TWS97].**

Indentification des composants d'invariance d'échelle

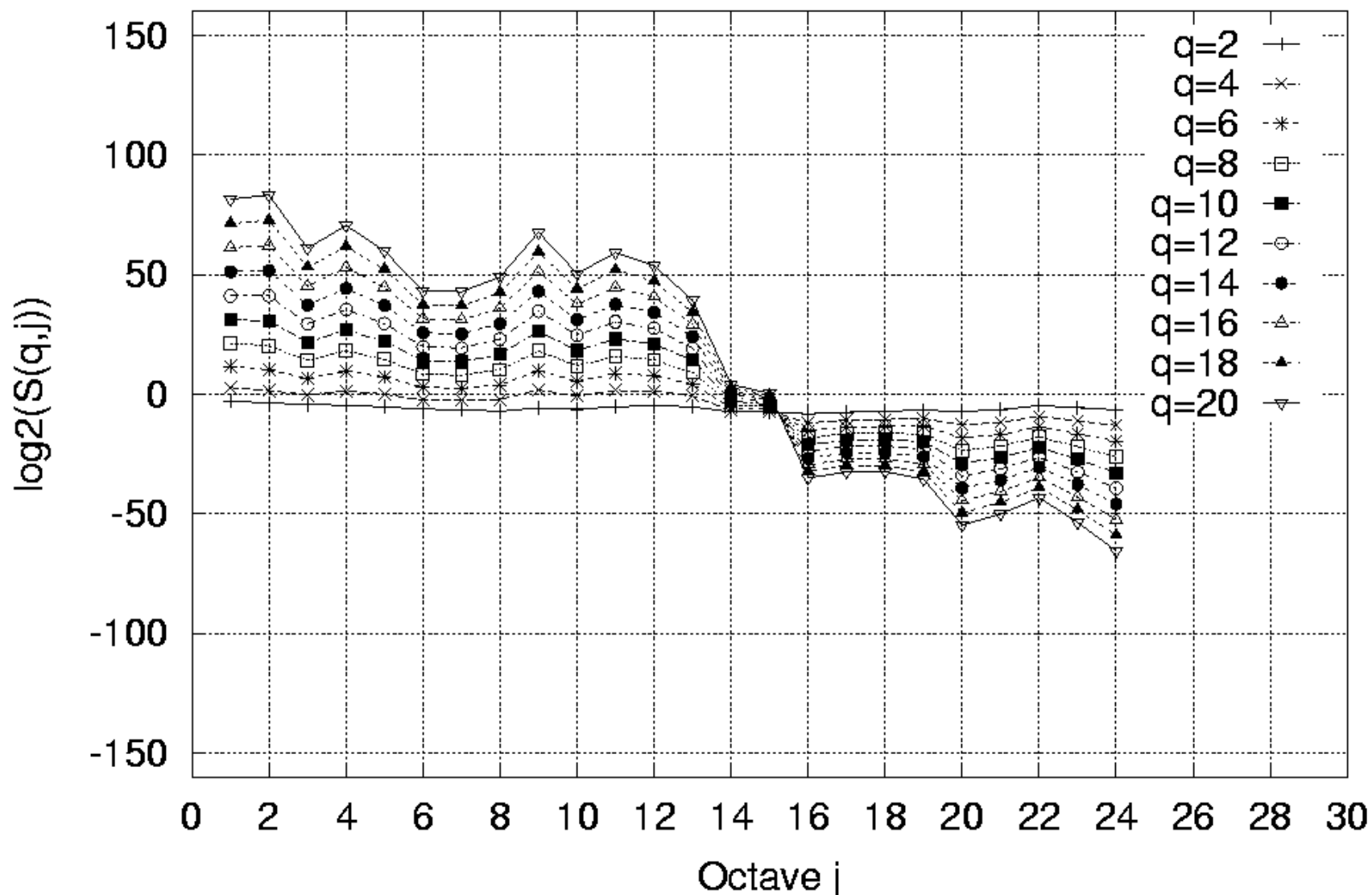
- **La littérature en réseau utilise les estimateurs LD et MD [AFTV00] pour identifier l'invariance d'échelle.**
- **Peu d'études ont essayé de vérifier que l'invariance d'échelle est une propriété invariante du signal, pas un biais provenant d'un sous-ensemble du signal.**
- **Est-il possible de faire mieux que la littérature en utilisant l'information contenue dans $d(j,k)$?**

Invariance d'échelle d'ordre supérieur

- Il est possible de distinguer entre différents types d'invariance d'échelle en étudiant comment les moments q évoluent en fonction des échelles de temps j
- La fonction de partition $S(q,j) = \sum_k |2^{-j/2} d(j,k)|^q$
- $M(j) = \langle \ln(S(q+1,j)) - \ln(S(q,j)) / \ln(S(2,j)) - \ln(S(1,j)) \rangle_q$
 - pas d'invariance: $M(j) < 1$
 - auto-similarité: $M(j) \sim 1$
 - multi-fractal: $M(j) > 1$

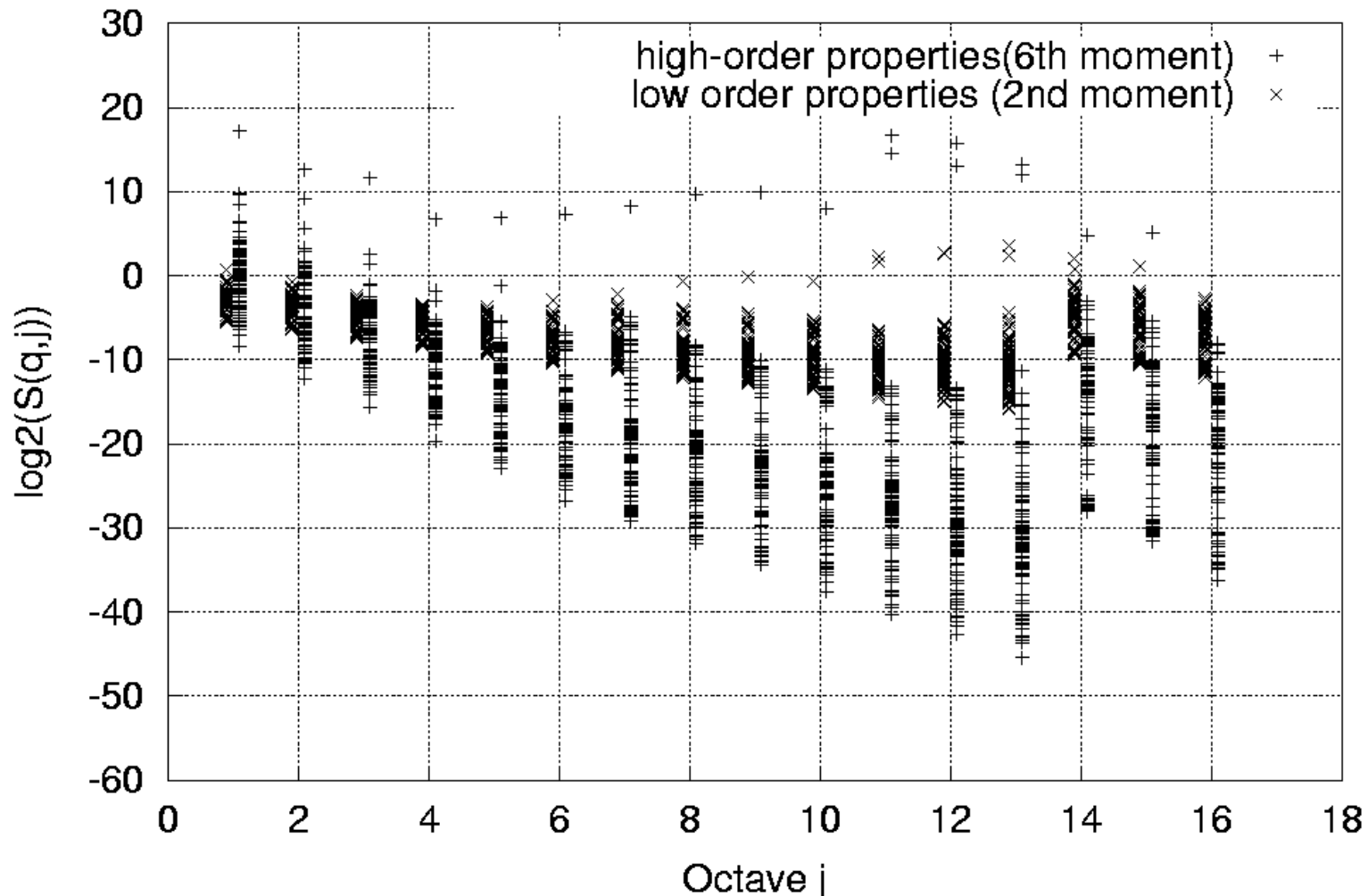
Faire confiance à $S(q,j)$

Partition function for Auckland incoming flow arrivals

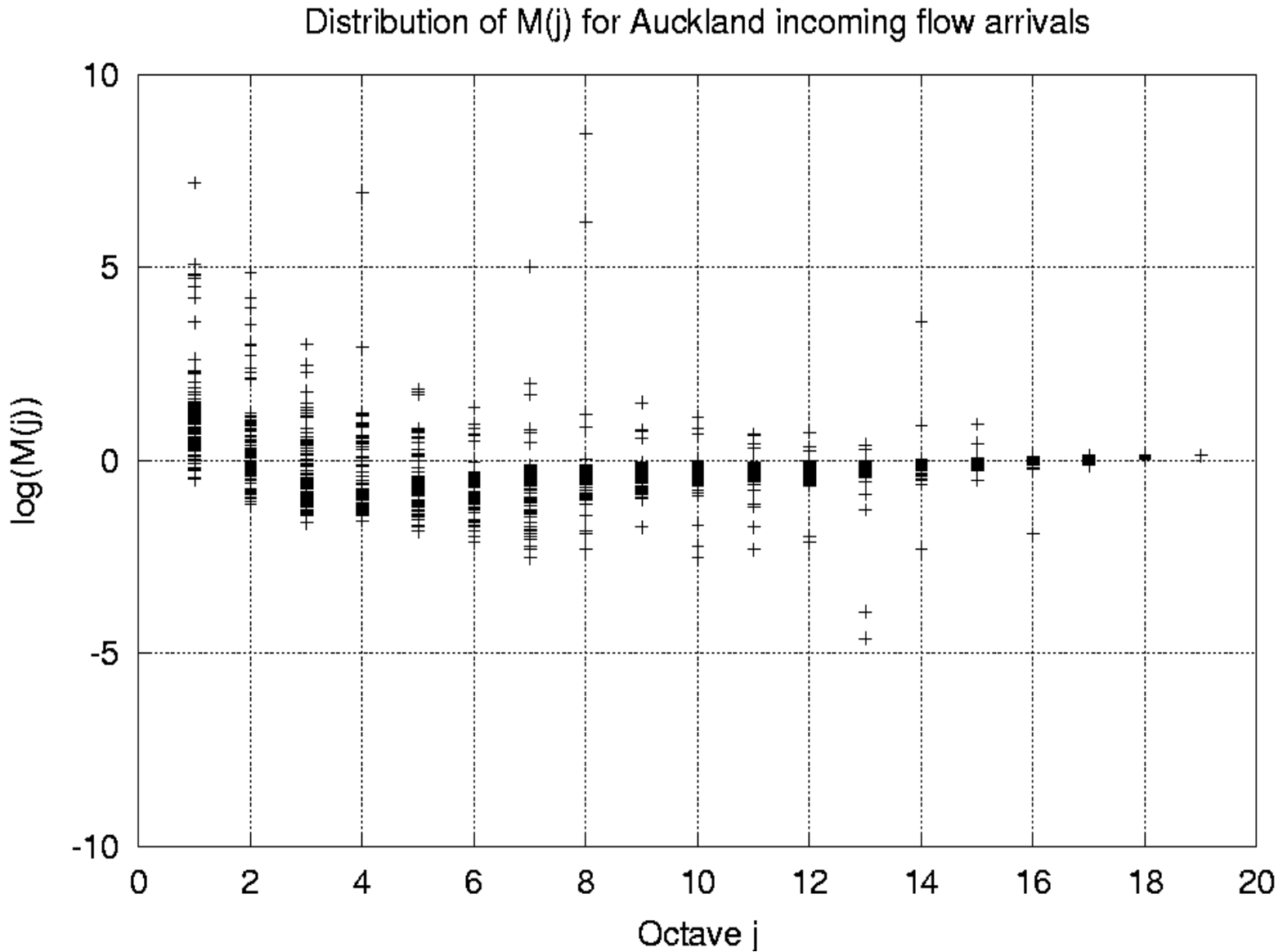


Ne faites pas confiance à $S(q,j)$

Distribution of high order properties for Auckland incoming flow arrivals



Le biais de la non-stationarité



Conclusions

- **Il existe assez d'outils à ce jour pour analyser des signaux complexes contenant de l'invariance d'échelle complexe, encore faut-il les utiliser !**
- **Questions ouvertes en traitement de signal:**
 - **Comment pouvoir se focaliser sur des échelles de temps limitées sans souffrir d'un biais provenant d'autres échelles de temps ?**
 - **Comment vérifier que la non-stationarité ne biaise pas les résultats, ainsi que les hypothèses de modélisation ?**

- **[FGW98] A. Feldmann, W. Willinger and A. Gilbert. *Data networks as cascades: Investigating the Multifractal Nature of Internet WAN Traffic*. Proc. of ACM SIGCOMM, 1998.**
- **[TWS97] M. Taqqu, W. Willinger and R. Sherman. *Proof of a Fundamental Result in Self-Similar Traffic Modeling*. Comp. Comm. Rev., 1997.**
- **[VA00] D. Veitch and P. Abry. *A wavelet based joint estimator for the parameters of LRD*. IEEE Transactions on Information Theory, 45(3), April 1999.**
- **[AFTV00] Abry, Flandrin, Taqqu, Veitch. *Wavelets for the analysis, estimation and synthesis of scaling data*. Book chapter of "Self Similar Network Traffic Analysis and Performance Evaluation", Wiley, 2000.**
- **[WPRT02] W. Willinger, V.Paxson, R. Riedi and M. Taqqu. *Long-range dependence and data network traffic*. Book chapter of "Theory and applications of long-range dependence", Birkhäuser, Boston, 2002.**