tudor
PUBLIC RESEARCH CENTRE HENRI TUDOR

University of South Australia

Flinders UNIVERSITY ADELAIDE • AUSTRALIA

tudor PUBLIC RESEARCH CENTRE HENRI TUDOR

JOHNS HOPKINS UNIVERSITY

Predicting the future? Comparative Forecasting and a test for Persistence in the El Nino Southern Oscillation … and a few new ideas

Dan Zachary

May 27, 2014

tudor
PUBLIC RESEARCH CENTRE HENRI TUDOR

University of South Australia

Flinders University Adelaide • Australia

tudor
PUBLIC RESEARCH CENTRE HENRI TUDOR

Johns Hopkins University

Brief Introduction

The Bayesian Binary Tree (BBT) model

A path expectation model

Forecasting

Conclusions

Belinda Chiera (University of South Australia)

Jerzy Filar (Flinders University)

Adrian H. Gordon (University of South Australia) – deceased

Daniel Zachary (Johns Hopkins, Tudor)

The Motivation

- ***El Nino Southern Oscillation (ENSO):***

"well known source of interannual climatic variability with far-reaching flow-on effects"

- **Economic impact:** agriculture, commercial fishing, construction & tourism.

- **Environmental impact:** drought and flooding.

- **Human impact:** loss of life and livelihood.

ENSO forecasting now an important part of policy making.

Tradeoff? complex models at the cost of manageability?
Does this really translate into forecast quality?

Alternative? a simple, probabilistic (Bayesian) forecast that exploits *persistence* in ENSO.

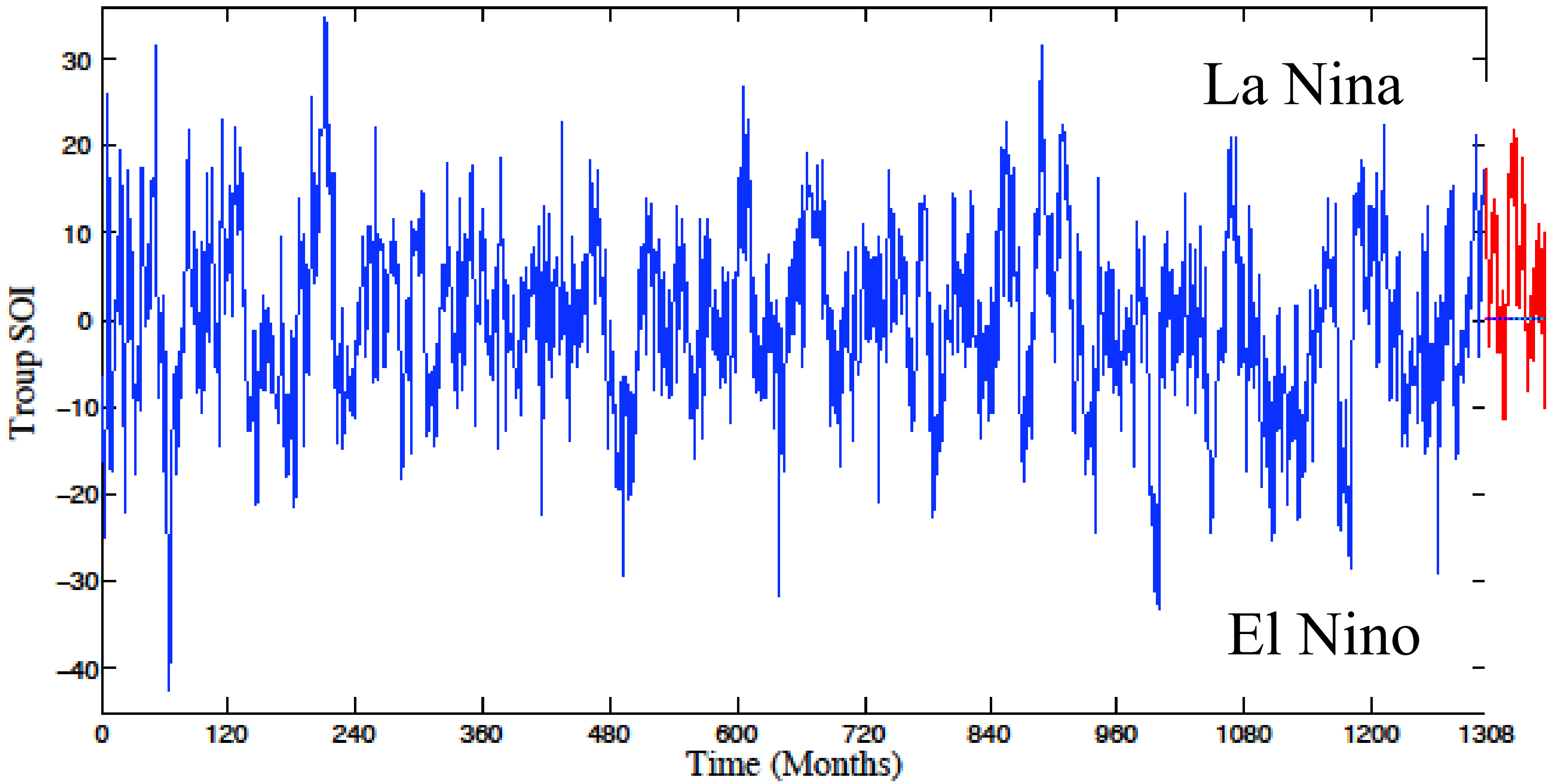
SOI: The Southern Oscillation Index

$$SOI = 10 \cdot \frac{P_d - \overline{P_d}}{\sigma_{P_d}}$$

P_d : difference between average Tahiti MSLP & Darwin MSLP

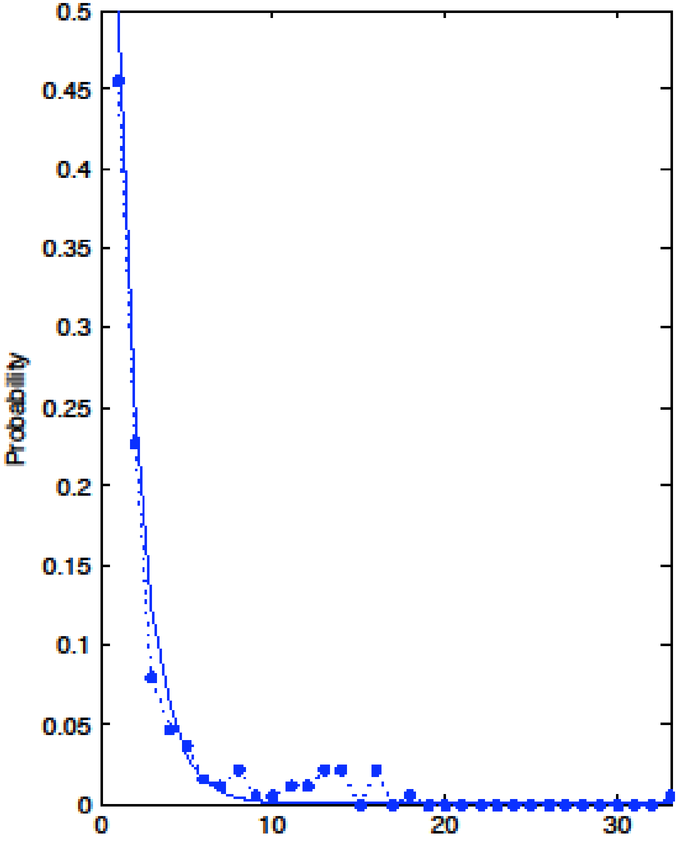
$\overline{P_d}$: long-term average of P_d.

σ_{P_d} : long-term standard deviation of P_d.



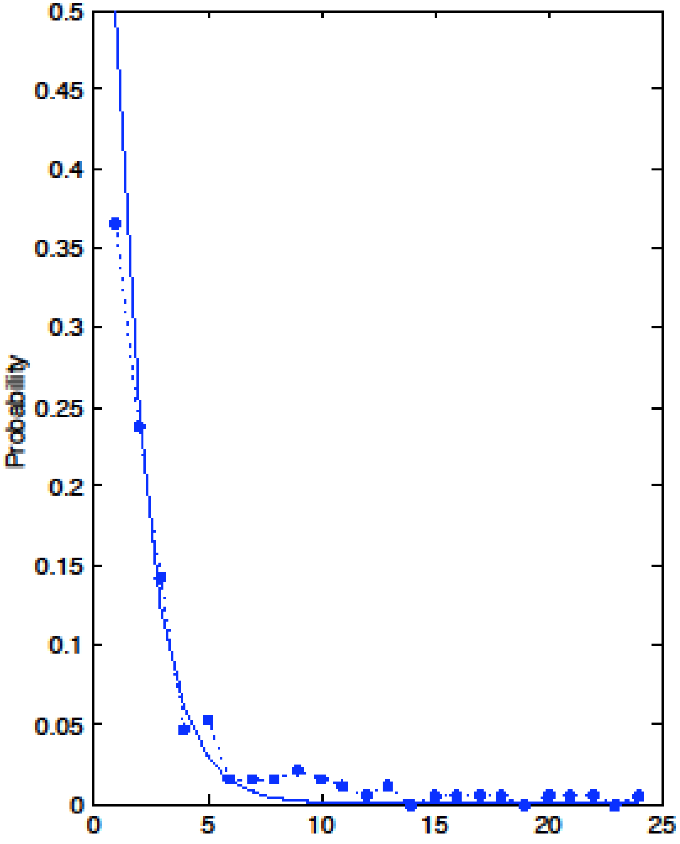
SOI Probability distribution vs a Binomial distribution (p = 0.5)

Negative Valued Events



Months

Positive Valued Events



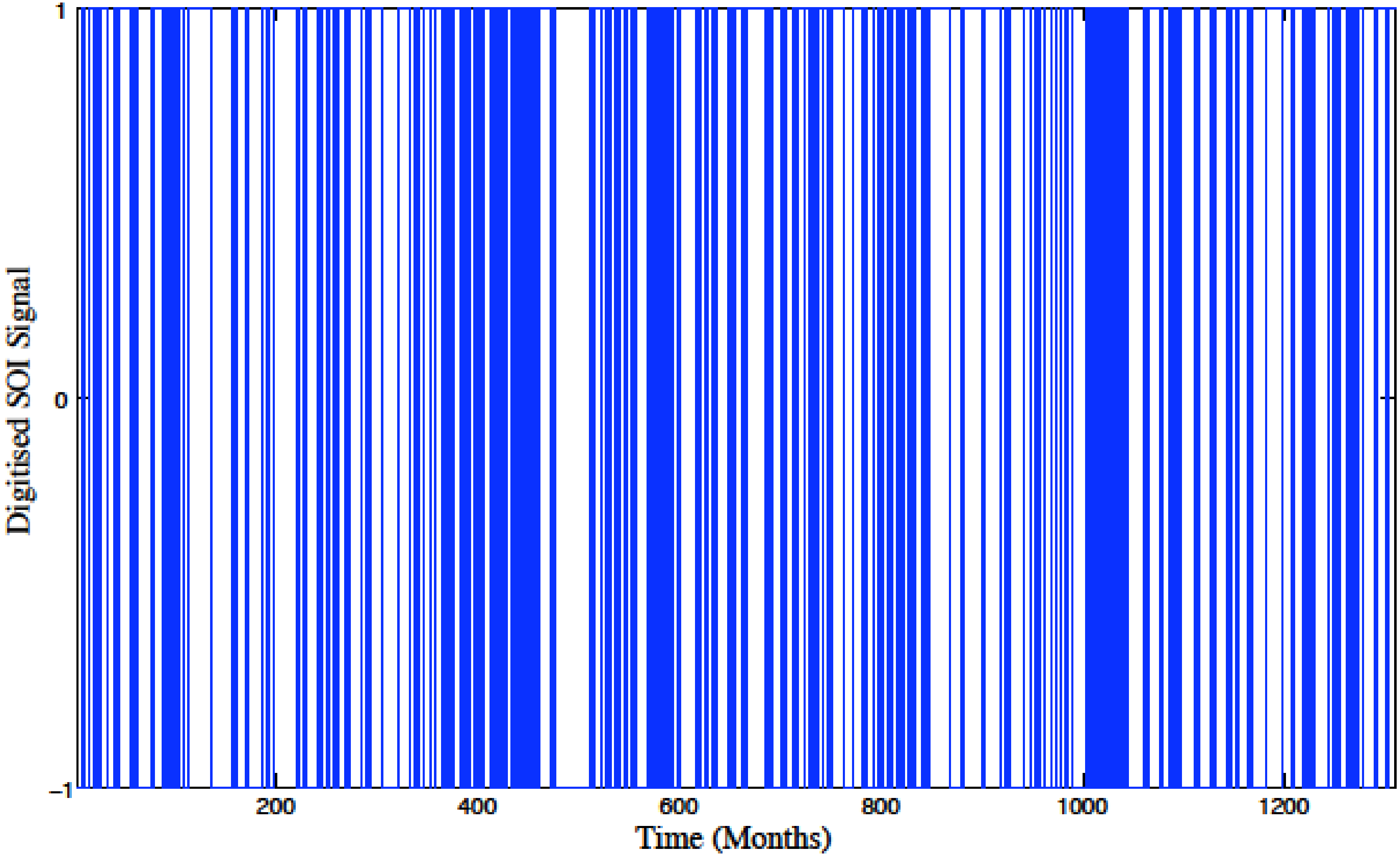
Months

SOI : Digitised

Digitised SOI = +1 if SOI ≥ 0: Up event

-1 if SOI < 0: Down event

*Blue/White bands signal **persistence** in the series*



Up Events:

$\mathcal{U}_1$ = *Number of Up episodes of length 1*

$\mathcal{U}_2$ = *Number of Up episodes of length 2*

...

$\mathcal{U}_{\mathbf{u}}$ = *Number of Up episodes of length* ***u***

$$\mathcal{U}_T = \sum_{i=1}^{\mathbf{u}} \mathcal{U}_i$$

Down Events:

$\mathcal{D}_1$ = Number of Down episodes of length 1

$\mathcal{D}_2$ = Number of Down episodes of length 2

...

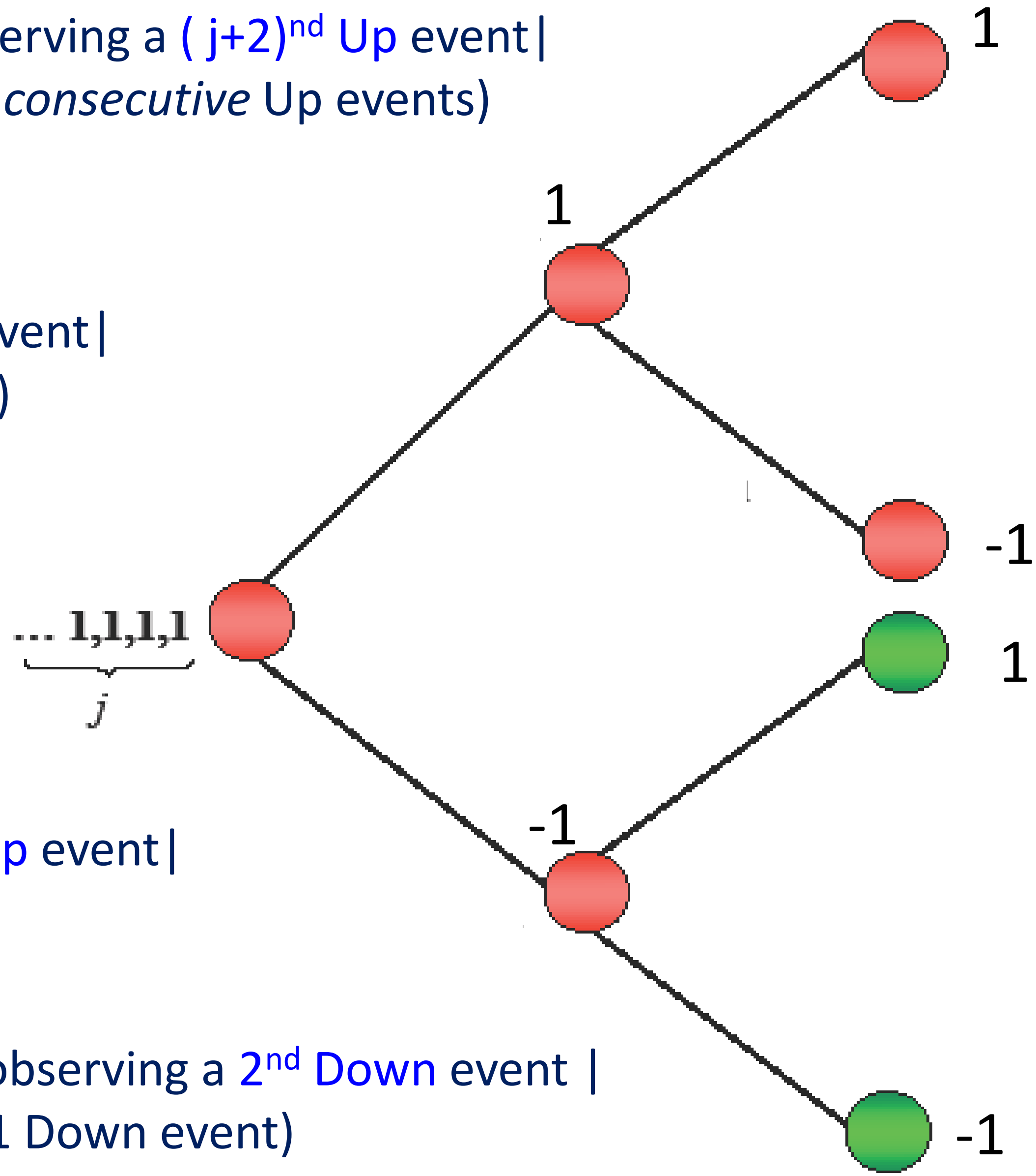
$\mathcal{D}_{\mathbf{d}}$ = Number of Down episodes of length **d**

$$\mathcal{D}_T = \sum_{i=1}^{\mathbf{d}} \mathcal{D}_i$$

The BBT model

P(observing a (j+2)nd Up event|
j+1 *consecutive* Up events)

P(observing a *(j+1)st Up* event|
j *consecutive* Up events)



P(not observing a (j+1)st Up event|
j consecutive Up events)

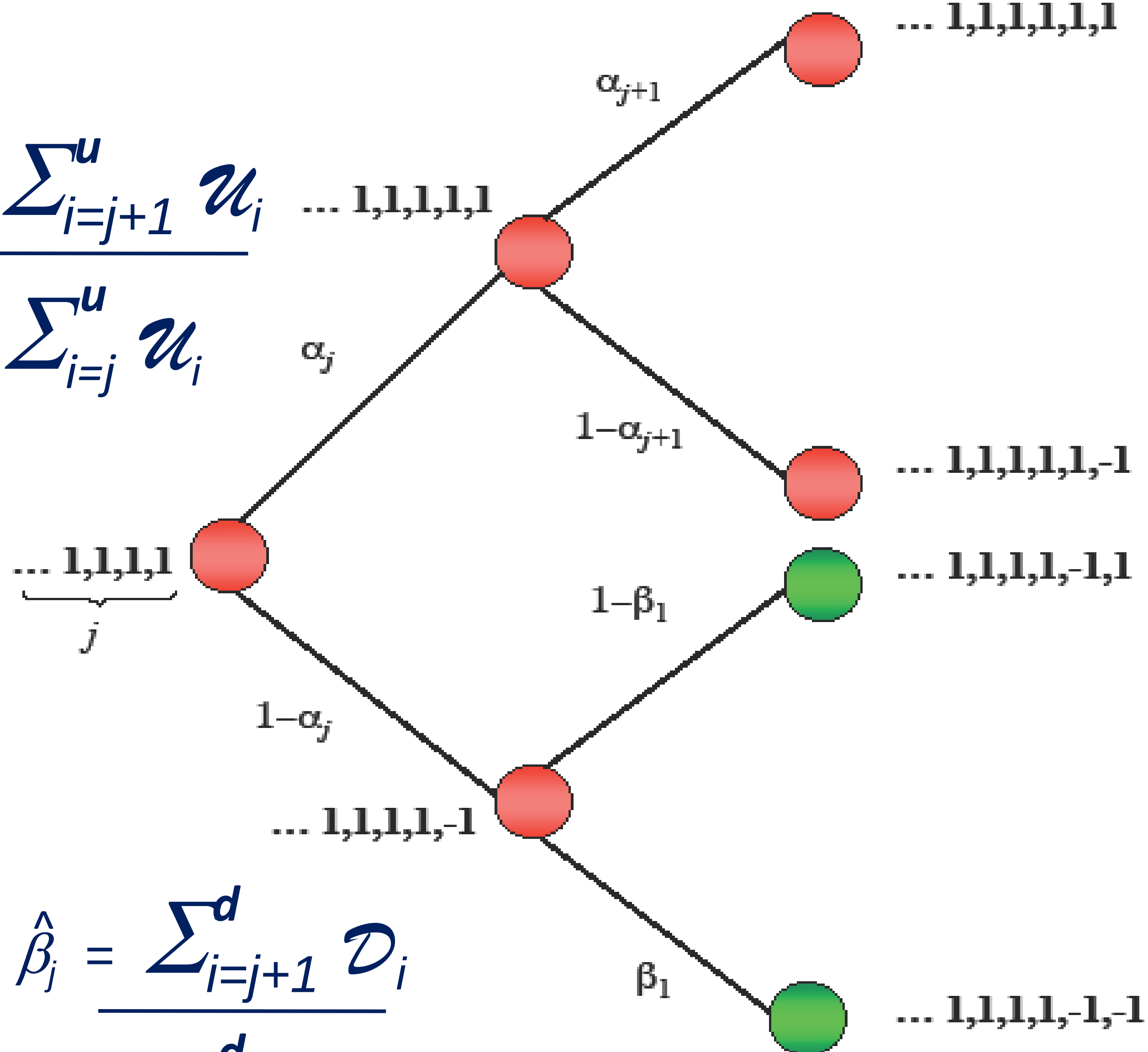
P(observing a 2nd Down event |
1 Down event)

tudor
PUBLIC RESEARCH CENTRE HENRI TUDOR

The Bayesian Binary Tree Model (BBT)

The BBT model

$$\hat{\alpha}_j = \frac{\sum_{i=j+1}^{u} \mathcal{U}_i}{\sum_{i=j}^{u} \mathcal{U}_i}$$



So how do we use a probability model of +1 and -1 events to reconstruct and forecast the SOI?

$$\hat{\beta}_j = \frac{\sum_{i=j+1}^{d} \mathcal{D}_i}{\sum_{i=j}^{d} \mathcal{D}_i}$$

The BBT model

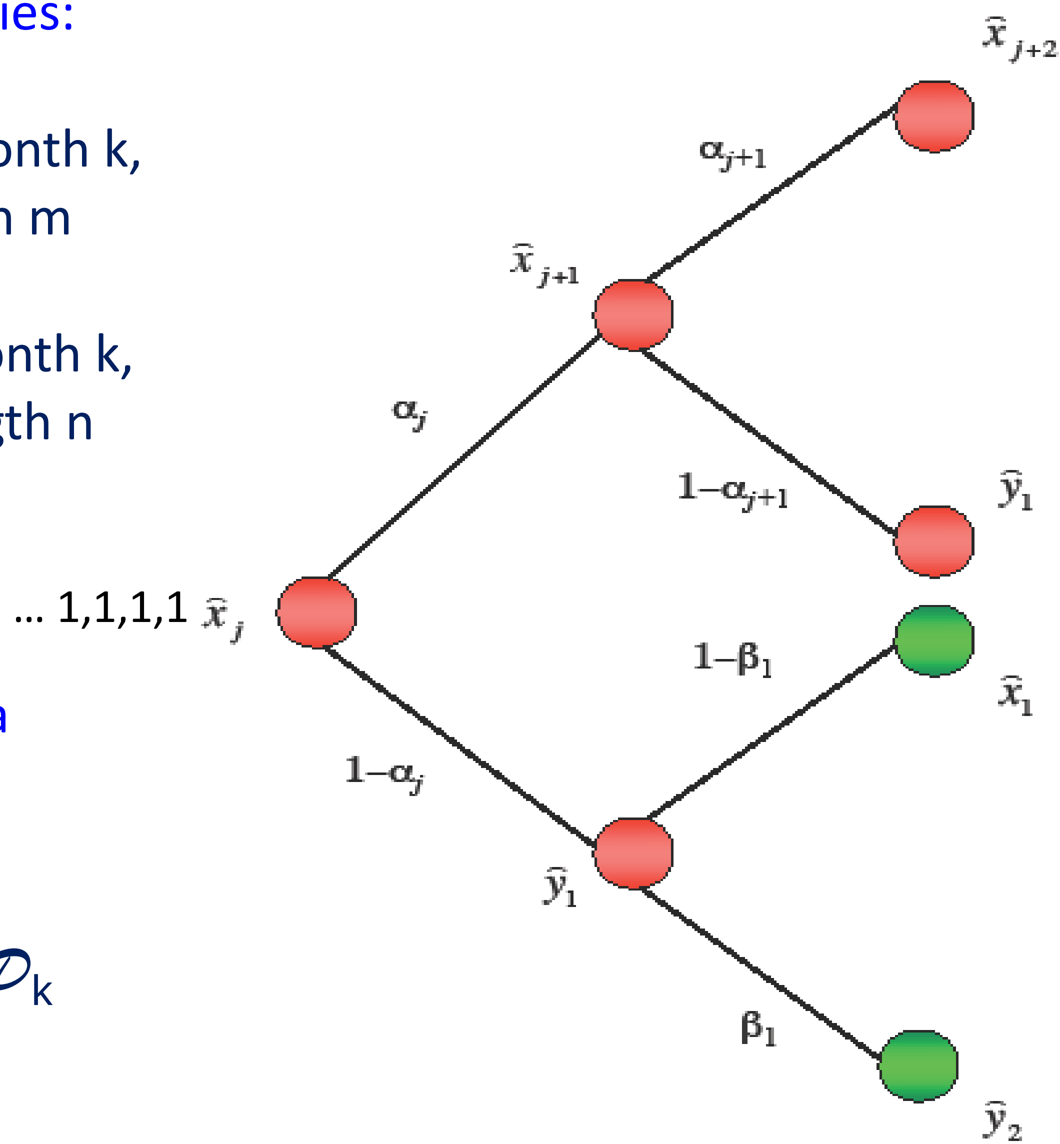
- Use the Up/Down episode histories:

  $\overline{x}_{m,k}$ = average *SOI value* for month k,
  in an *Up* cluster of length m

  $\overline{y}_{n,k}$ = average *SOI value* for month k,
  in a *Down* cluster of length n

- Take into account we may enter a cluster part-way through:

$$\mathcal{U}_{j.} = \sum_{k>j} \mathcal{U}_k \quad \text{and} \quad \mathcal{D}_{j.} = \sum_{k>j} \mathcal{D}_k$$



Forecasting – Part 1/3

- Use the Up/Down episode histories:

  $\bar{x}_{m,k}$ = average *SOI value* for month k, in an *Up* cluster of length m

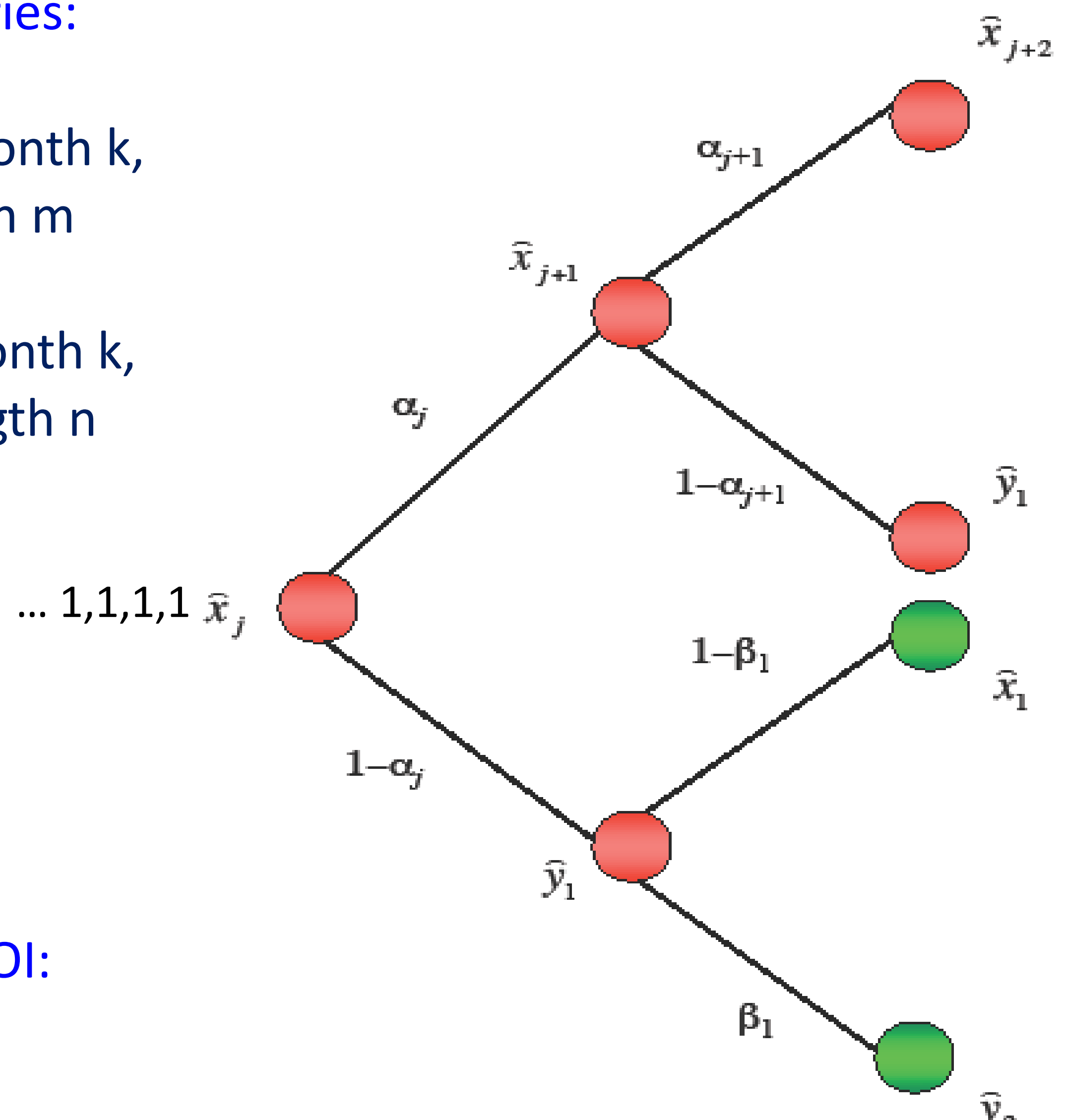
  $\bar{y}_{n,k}$ = average *SOI value* for month k, in a *Down* cluster of length n

- Estimate the positively valued SOI:

$$\hat{x}_{j+1} = \sum_{k=1}^{\mathbf{u}-j} \frac{\mathcal{U}_{j+k}}{\mathcal{U}_{j\cdot}} \bar{x}_{j+k,j+1}$$

- Estimate the negatively valued SOI:

$$\hat{y}_{j+1} = \sum_{k=1}^{\mathbf{d}-j} \frac{\mathcal{D}_{j+k}}{\mathcal{D}_{j\cdot}} \bar{y}_{j+k,j+1}$$



Forecasting – Part 2/3

- Construct Forecast Estimators:

$$\hat{z}(t+k) = E[SOI(t+k)]$$

$$\hat{z}_{\mathcal{U}}(t+k) = E[SOI(t+k) \mid \mathcal{U}]$$

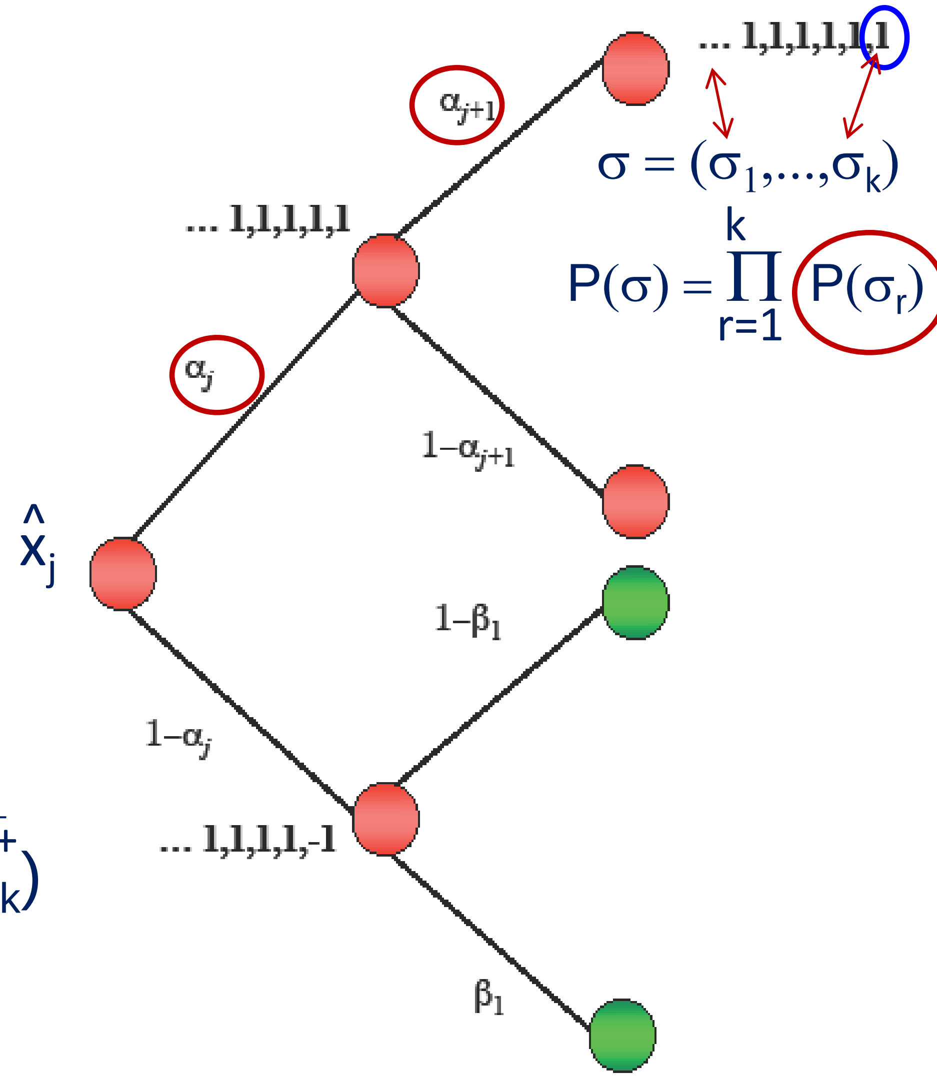
$$\hat{z}_{\mathcal{D}}(t+k) = E[SOI(t+k) \mid \mathcal{D}]$$

- Conditionalise on "Up":

$$A_k^+ = \{\sigma \mid \sigma_k = +1\}$$

$$P(A_k^+) = \sum_{\sigma \in A_k^+} P(\sigma) \quad \text{and} \quad \gamma^+(\sigma) = \frac{P(\sigma)}{P(A_k^+)}$$

$$\boxed{z_{\mathcal{U}}(t+k) = \sum_{\sigma \in A_k^+} \gamma^+(\sigma)\, \hat{x}_{t+k}(\sigma)}$$



$$\sigma = (\sigma_1, \ldots, \sigma_k)$$

$$P(\sigma) = \prod_{r=1}^{k} P(\sigma_r)$$

Forecasting – Part 3/3

- Construct Forecast Estimators:

$$\hat{z}(t+k) = E[SOI(t+k)]$$

$$\hat{z}_{\mathcal{U}}(t+k) = E[SOI(t+k) \mid \mathcal{U}]$$

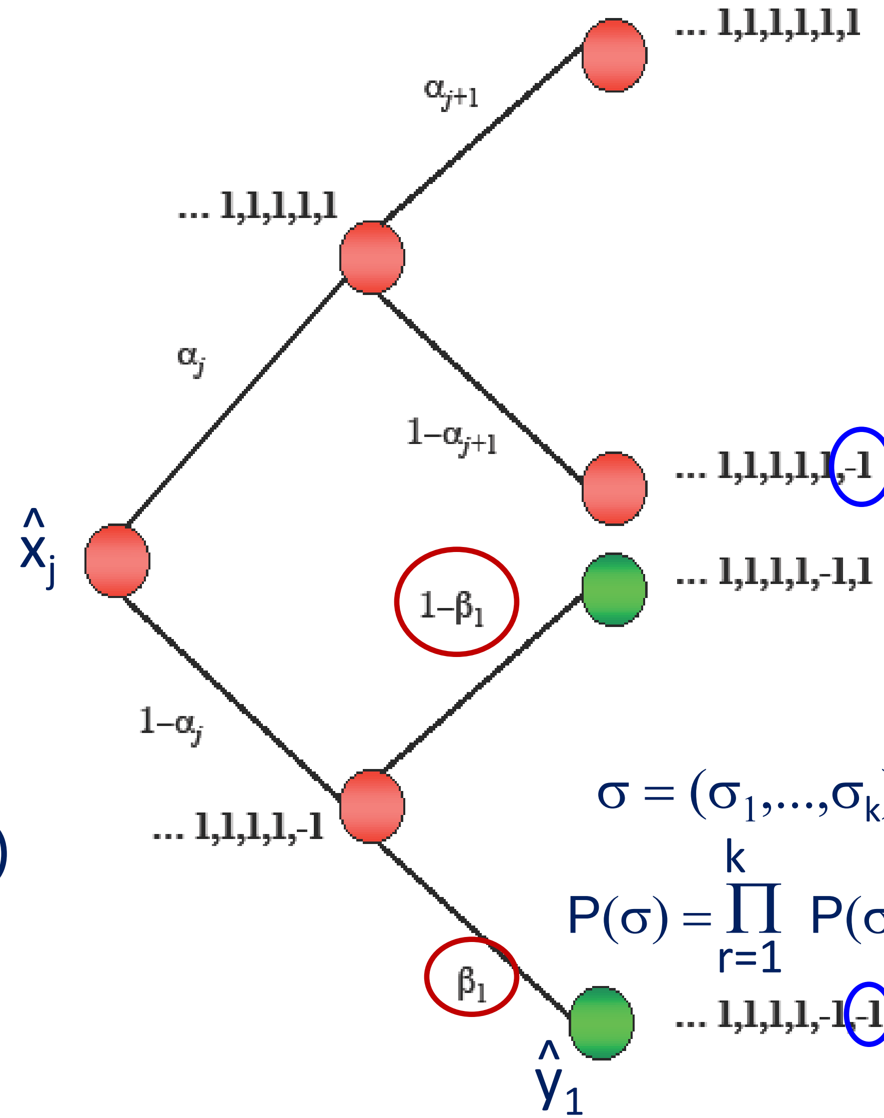
$$\hat{z}_{\mathcal{D}}(t+k) = E[SOI(t+k) \mid \mathcal{D}]$$

- Conditionalise on "Down":

$$A_k^- = \{\sigma \mid \sigma_k = -1\}$$

$$P(A_k^-) = \sum_{\sigma \in A_k^-} P(\sigma) \quad \text{and} \quad \gamma^-(\sigma) = \frac{P(\sigma)}{P(A_k^-)}$$

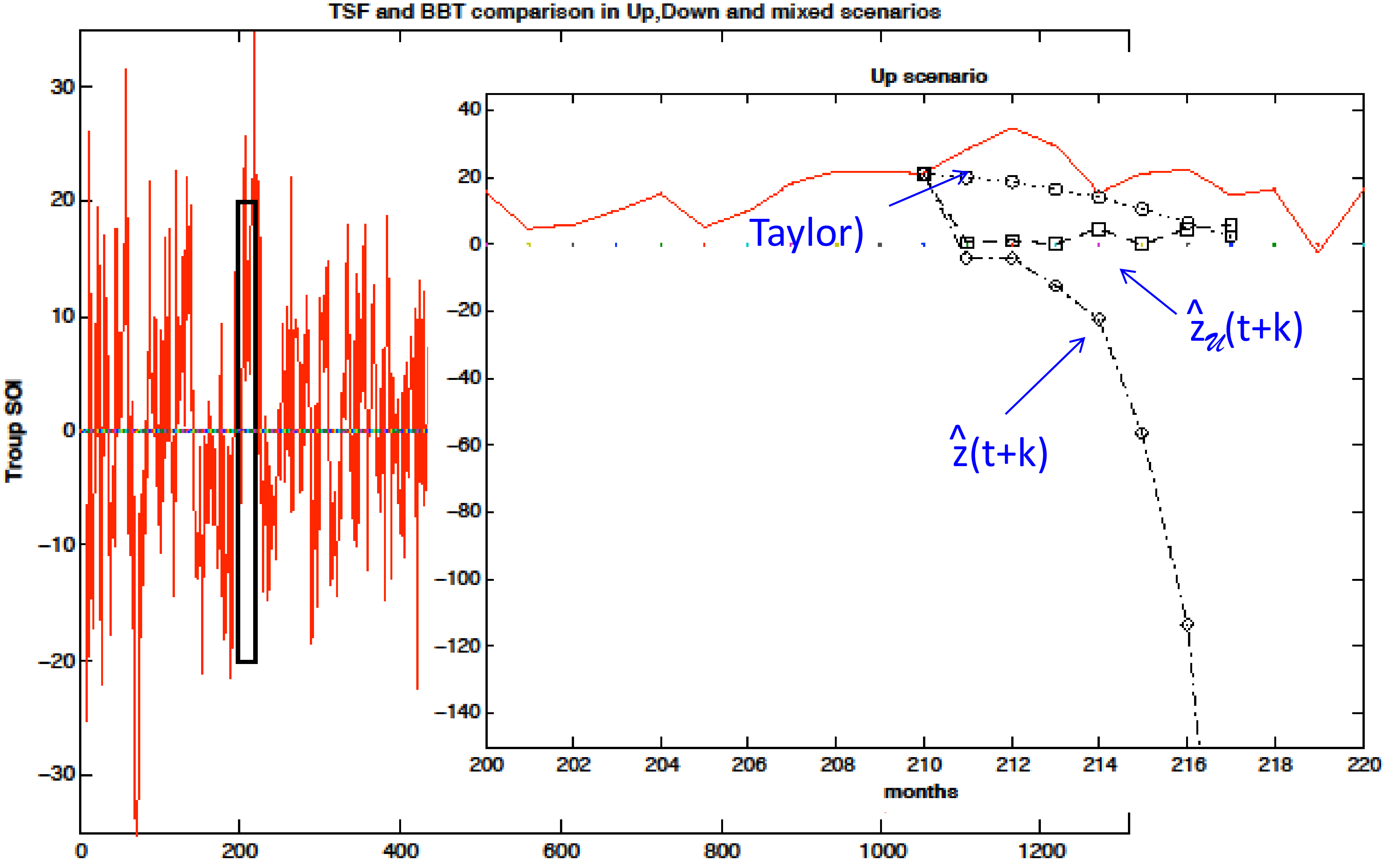
$$\boxed{z_{\mathcal{D}}(t+k) = \sum_{\sigma \in A_k^-} \gamma^-(\sigma)\, \hat{y}_{t+k}(\sigma)}$$



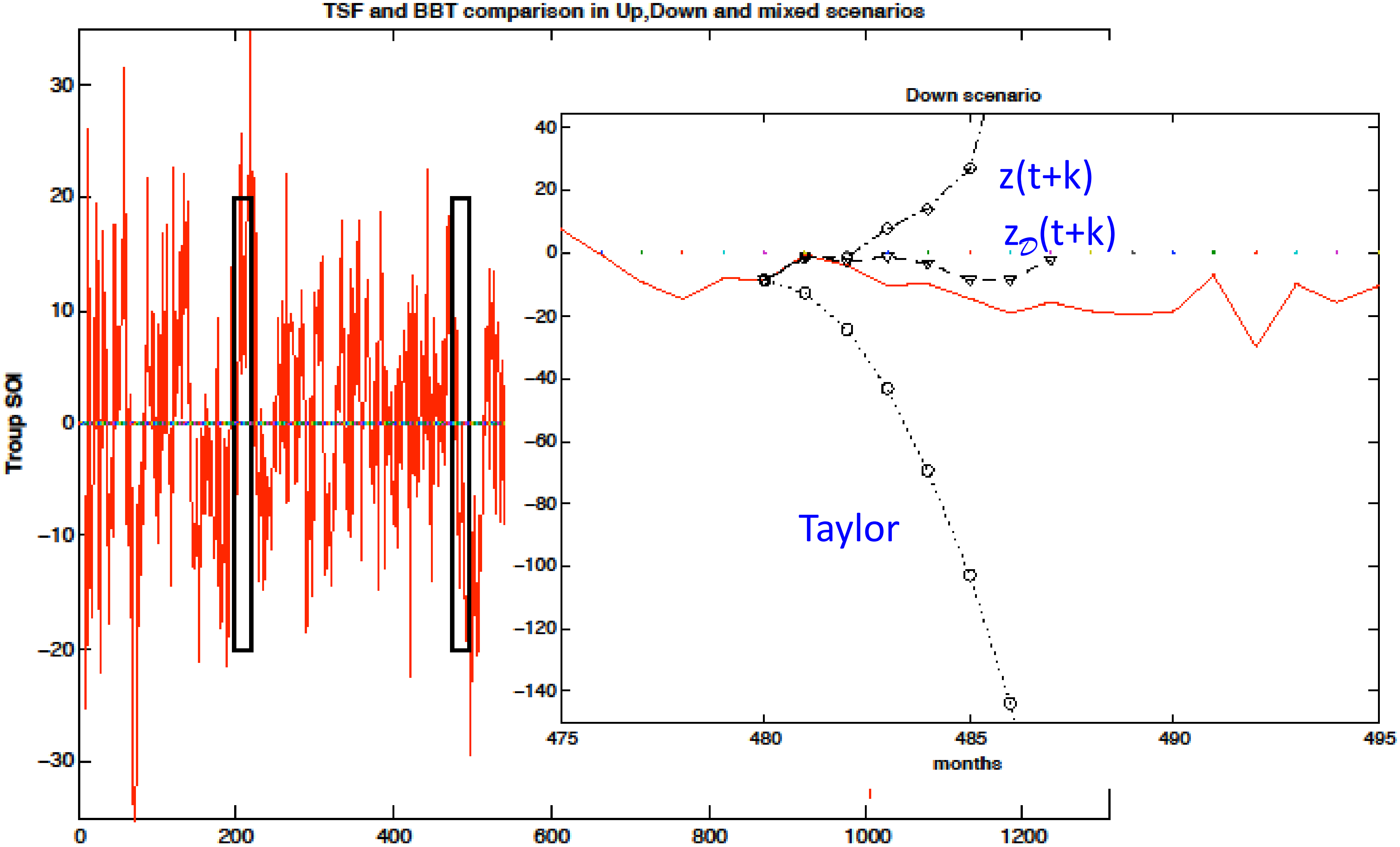
$$\sigma = (\sigma_1, \ldots, \sigma_k)$$

$$P(\sigma) = \prod_{r=1}^{k} P(\sigma_r)$$

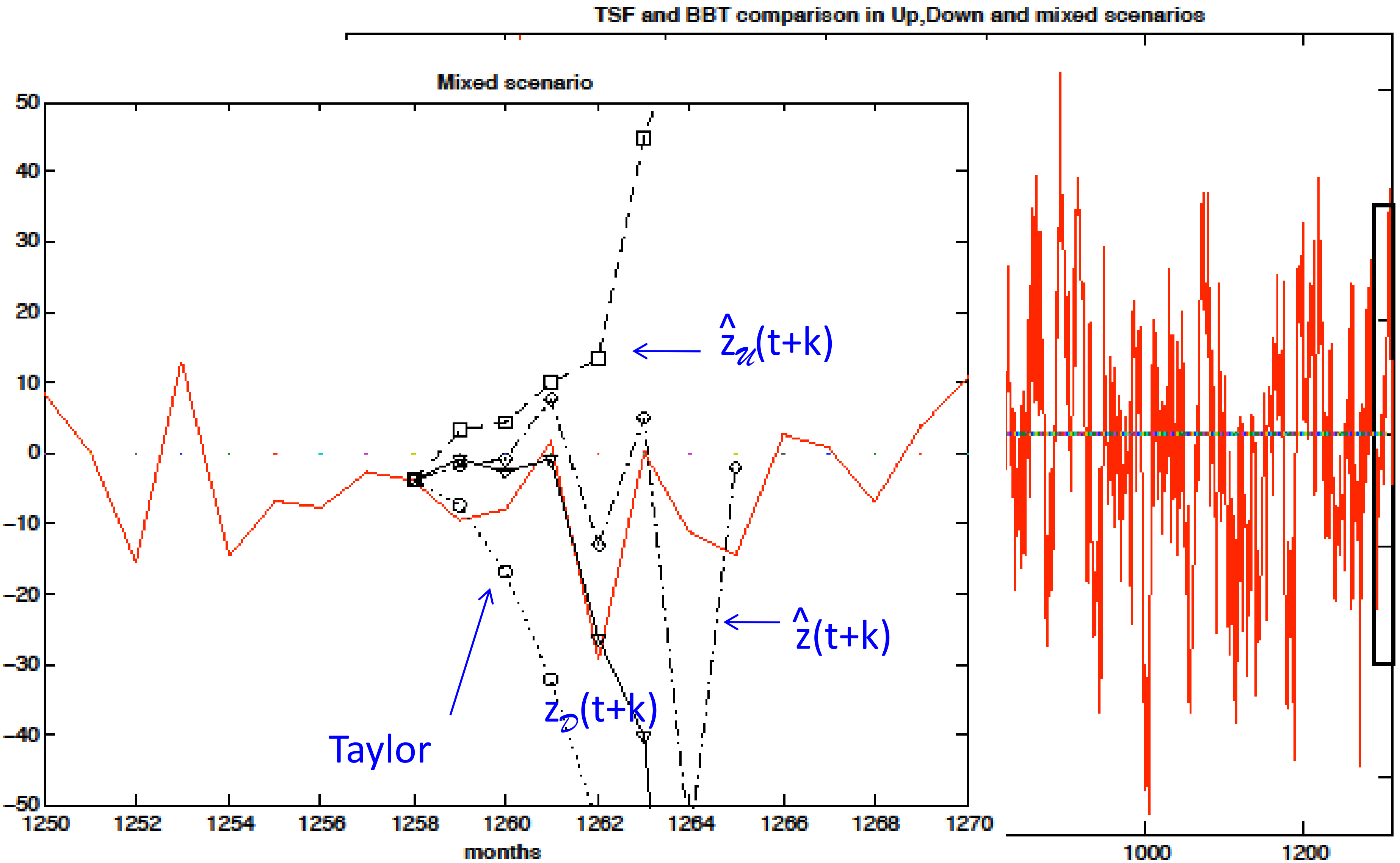
BBT model representation – up scenario



Model representation – down scenario



Model representation – mixed scenario (2010)

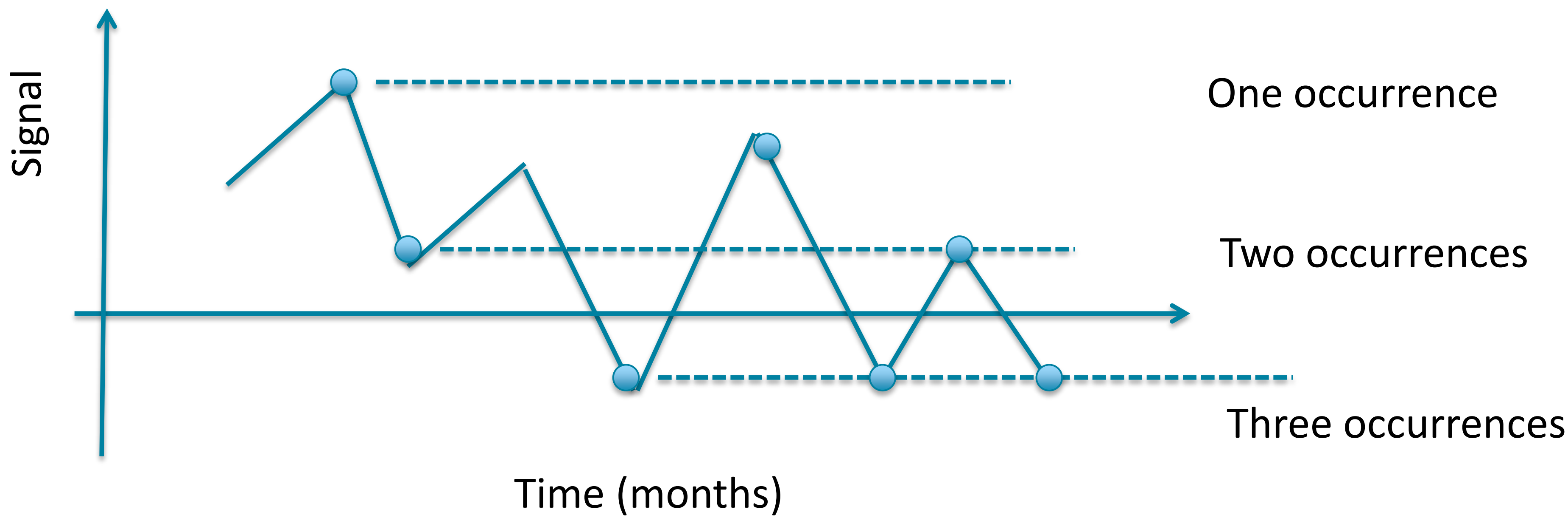


tudor
PUBLIC RESEARCH CENTRE HENRI TUDOR

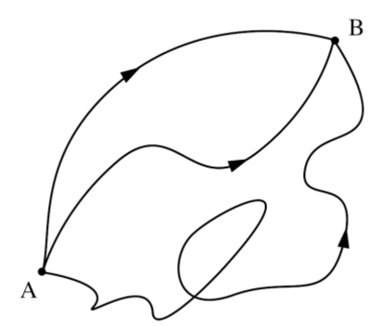
A path expectation model

Building on the expectation of each level

Find the ensemble of expectation values for each 'level' of the SOI

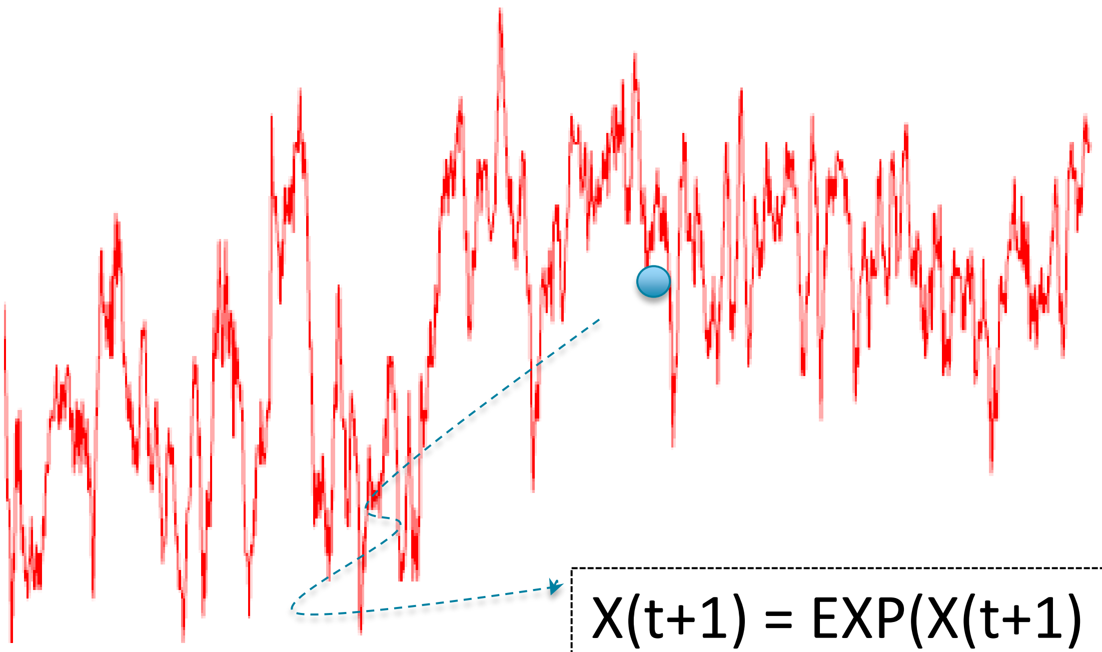


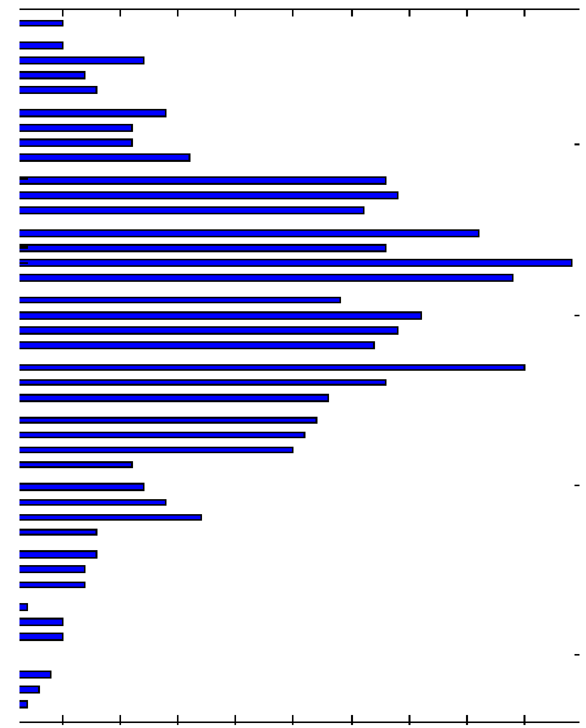
Inspired by the concept of most probable path of particles developed in quantum theory – finding the most probable path from A to B.



An alternative strategy

Find the ensemble of expectation values for each ‘level’ of the SOI





X(t+1) = EXP(X(t+1)

X(t+2) = EXP(X(t+2)

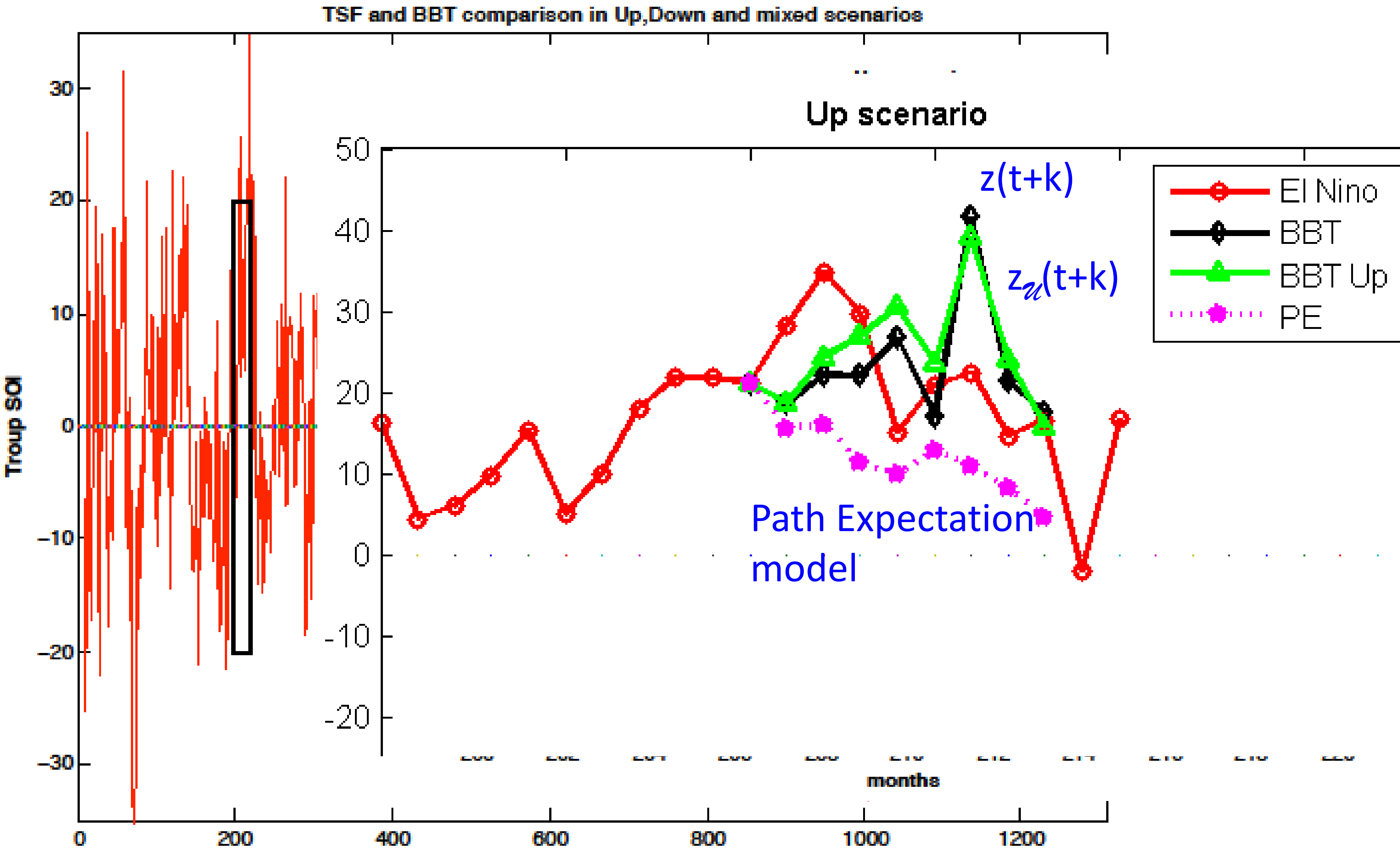
X(t+2) = EXP(X(t+3)

.

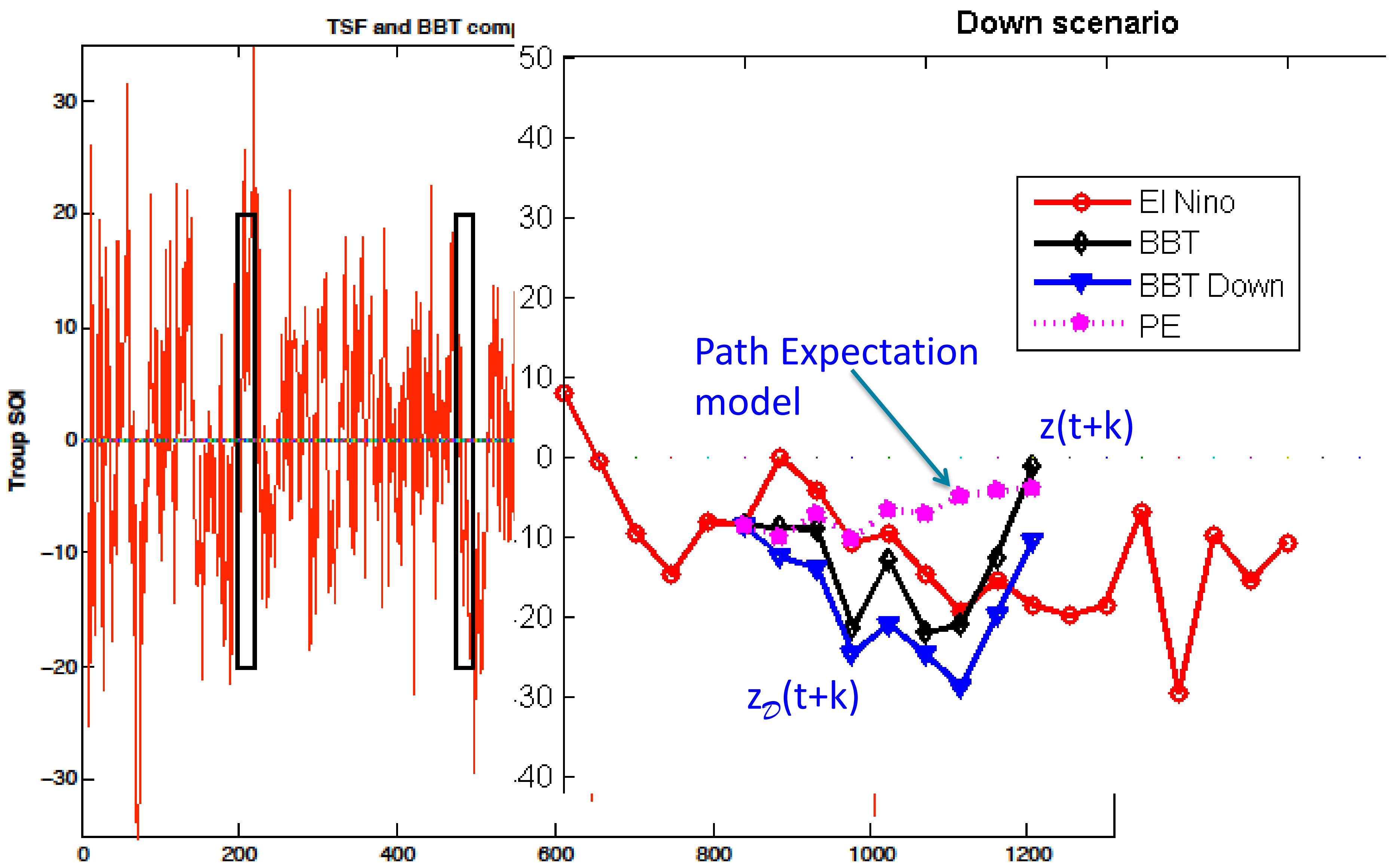
.

.

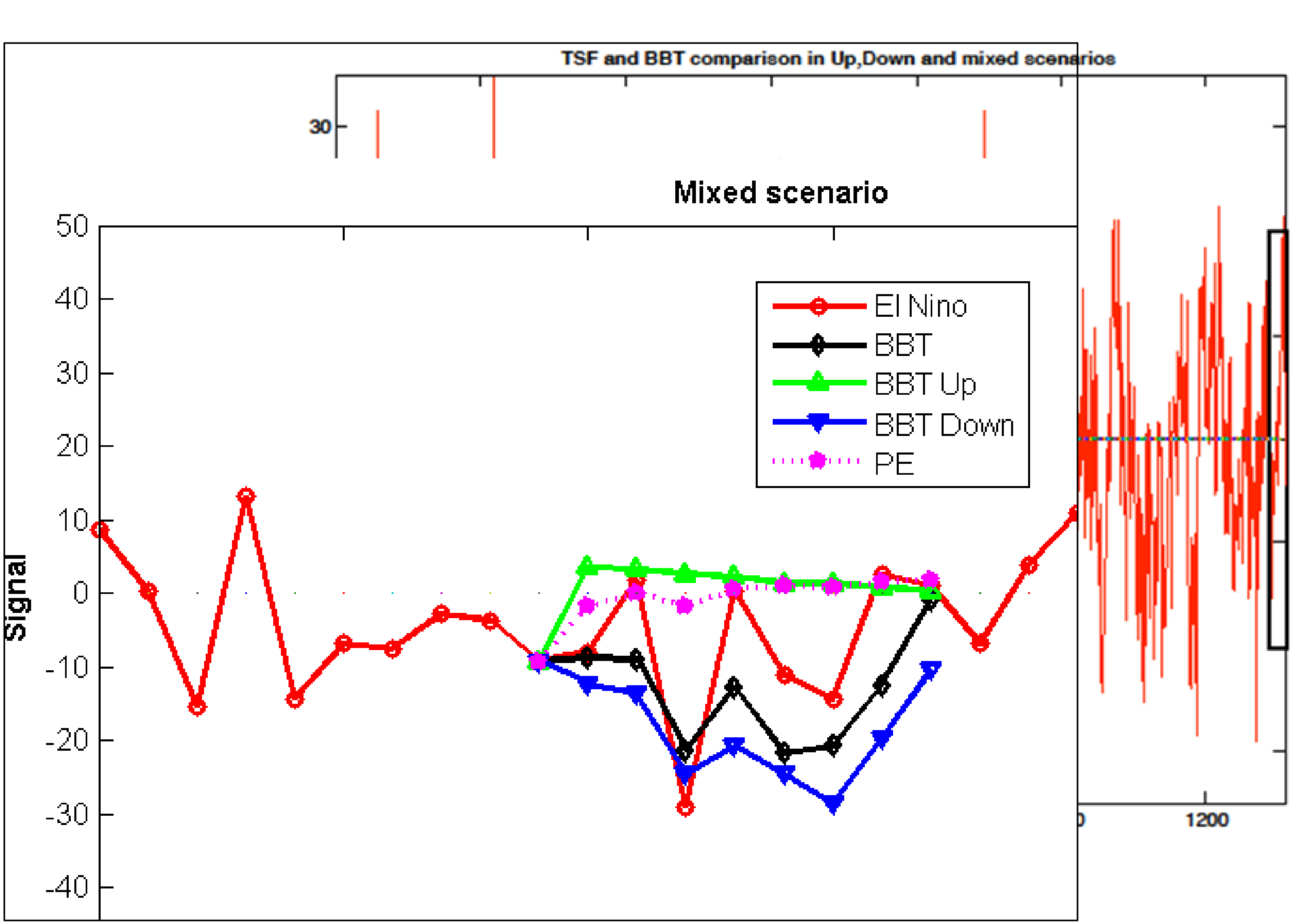
PE representation and comparisons – up scenario



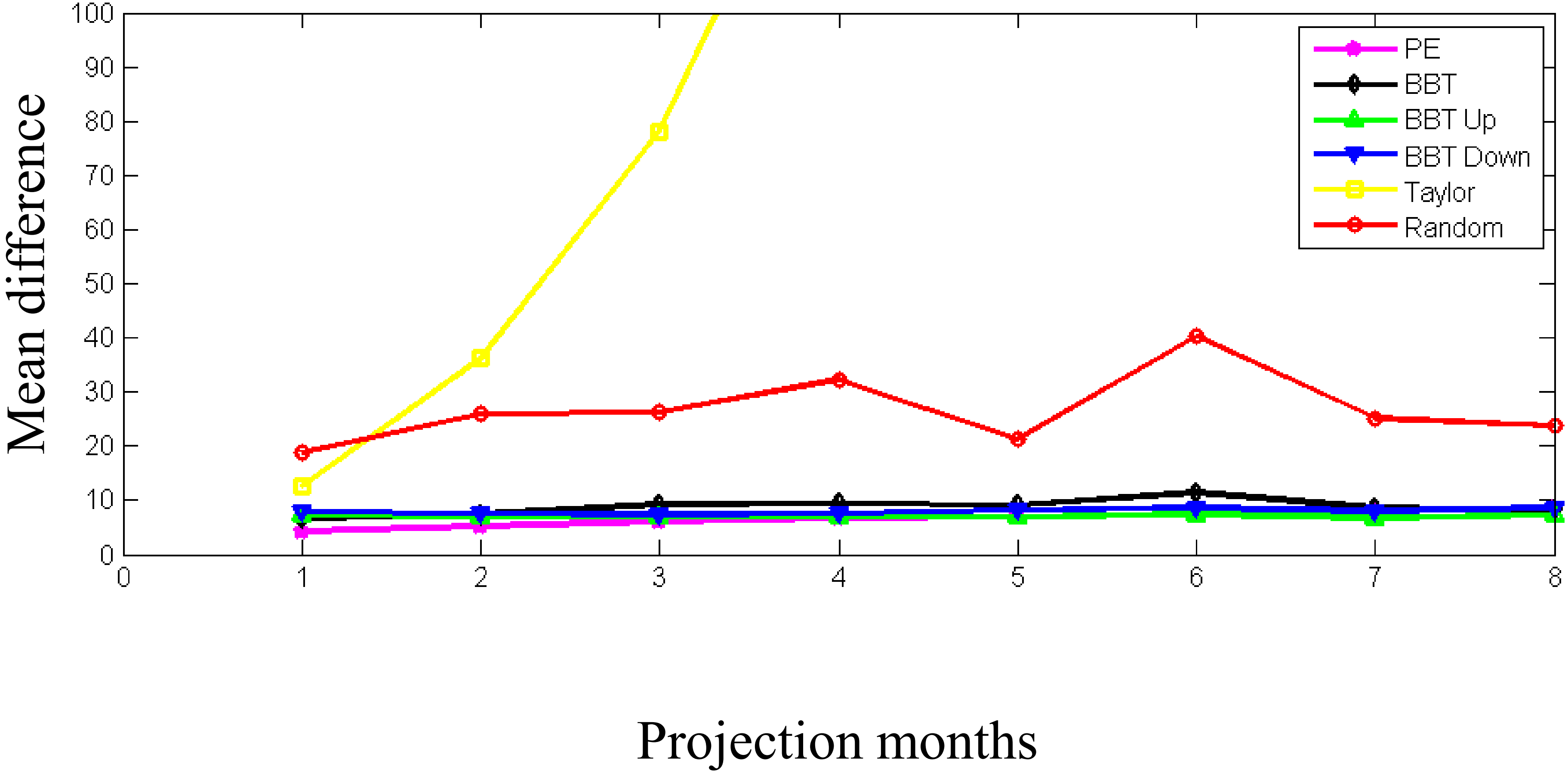
PE representation and comparisons – down scenario



PE representation and comparisons – mixed scenario



Mean differences - overview



tudor
PUBLIC RESEARCH CENTRE HENRI TUDOR

Forecasting

SOI Forecast – compared with IRI/CPC SST forecast

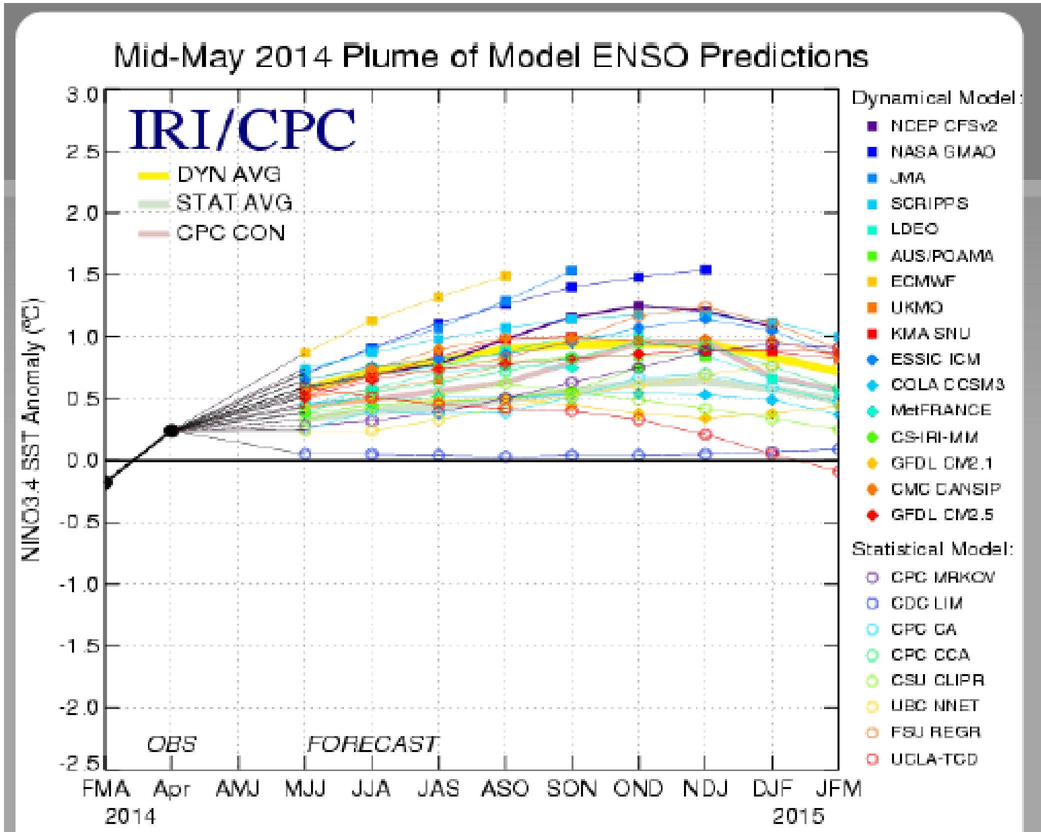
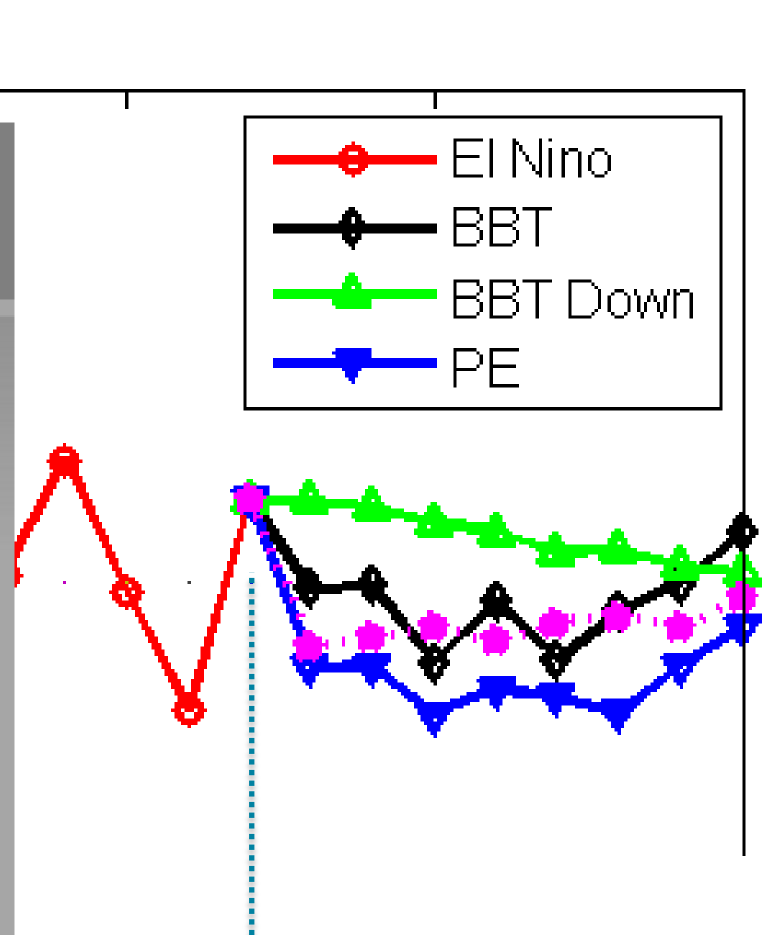


Figure provided by the International Research Institute (IRI) for Climate and Society (updated 13 May 2014).



April 2014

Conclusions

Summary

- The BBT models outperforms the TSF, Random model, and better adapts to the changing trend of the SOI.

- WORTH NOTING: BBT (Up/Down) typically do better forecast in periods of sustained Up/Down. BBT (unconditional) does better in mixed periods.

- Methods do reasonably good forecasts for several months beyond the forecast point, although there is generally considered a a breakdown in predictability power after this elapse of time.

- SHORT-TERM AIM: achieved. Simple schemes particularly useful in economic SOI-based forecasting (Austria - S. America research fund).

- LONG-TERM AIM: To develop a generalized the statistical method (e.g.Path Expectation model) and improve the quality of the forecasting tool & for a more realistic (e.g. tri-state) ENSO process, as described by the SOI.

tudor
PUBLIC RESEARCH CENTRE HENRI TUDOR

Thank you

Zachary D.S. On the sustainability of an activity, Nature: Scientific Reports, 2014, accepted.

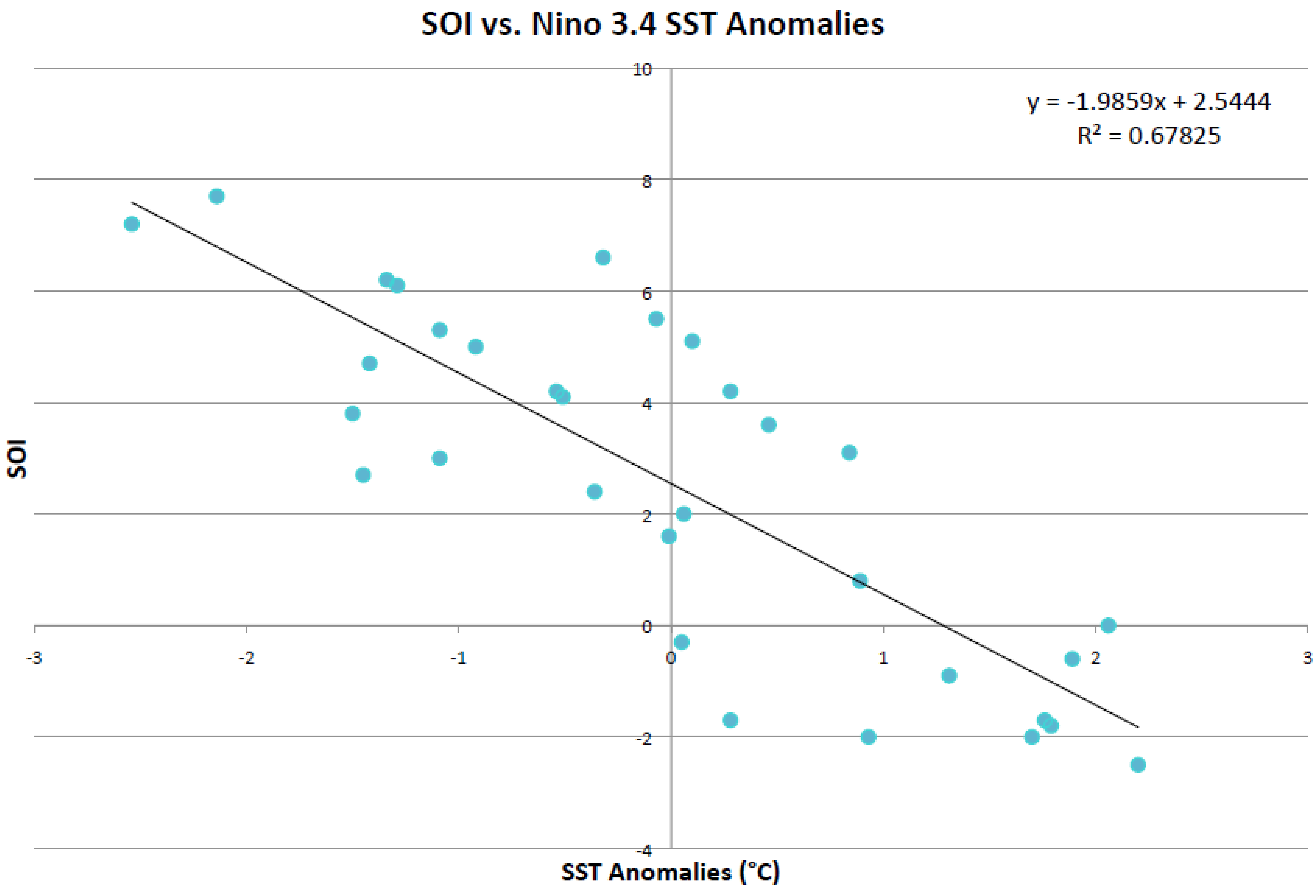
Zachary D.S. and Dobsen, S. Does urban space evolve deterministically or entropically? An exploration of urban development models for Sheffield, UK in relation to decentralized energy policies, Energy Policy (In review), 2013

Reis, L.A., Drouet L., Zachary D.S., Peters B., Melas D., Leopold U., Implementation of a full air quality model in an integrated assessment, Environmental Modeling and Software, in press, 2013

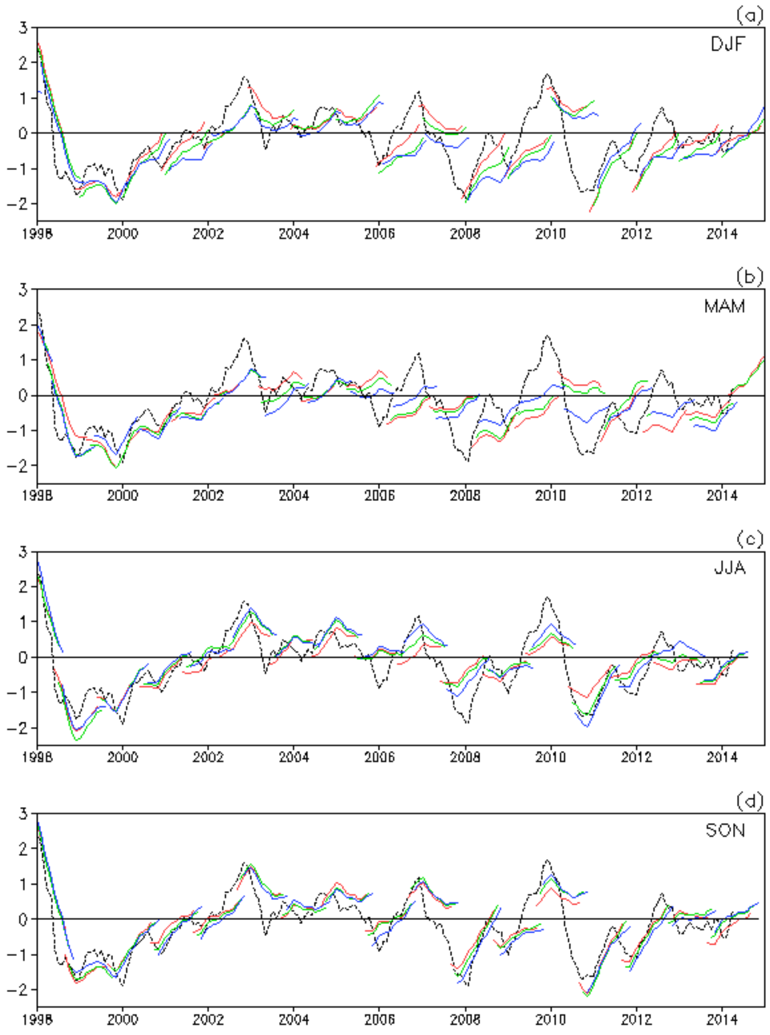
Reis L.A., Melas D., Peters D., Zachary D.S., Developing a fast photochemical Calculator for an integrated assessment model, International Journal of Environmental Pollution, Vol. 50, Nos. 1/2/3/4, 2012

Zachary D.S., Drouet L., Leopold U., Reis L.A. Trade-offs Between Energy Cost and Health Impact in a Regional Coupled Energy-Air Quality Model, Environmental Research Letters, 6(201), doi:10.1088/1748-9326/6/2/024021, 2011.

A new strategy

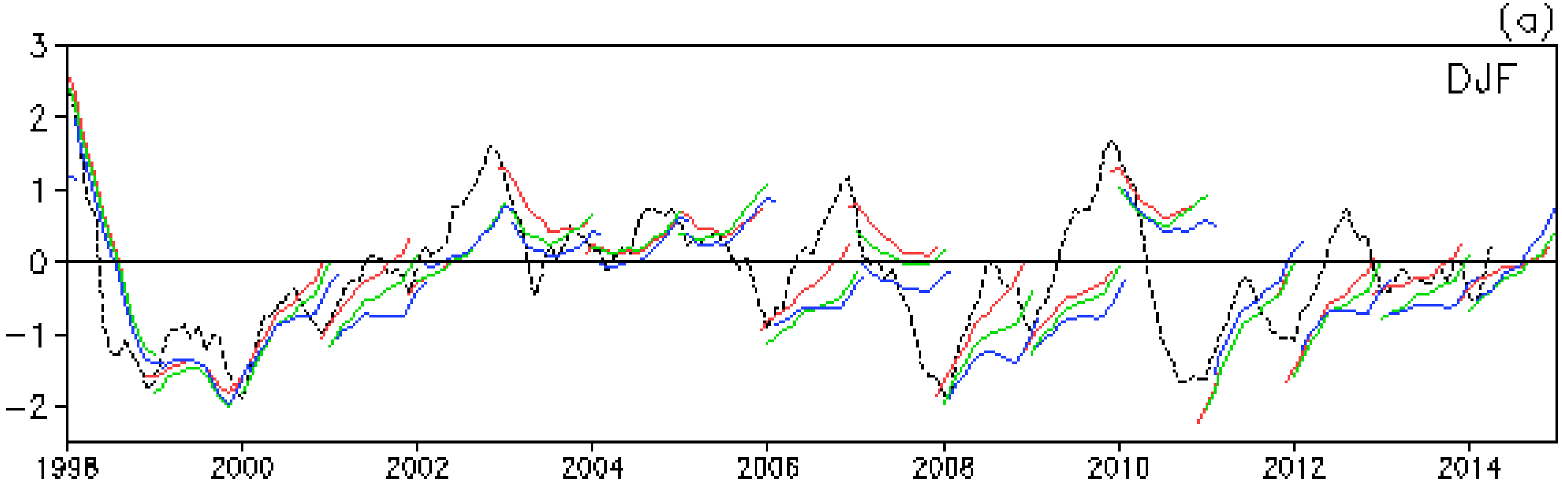


A comparison with NCEP/CPC Markov model



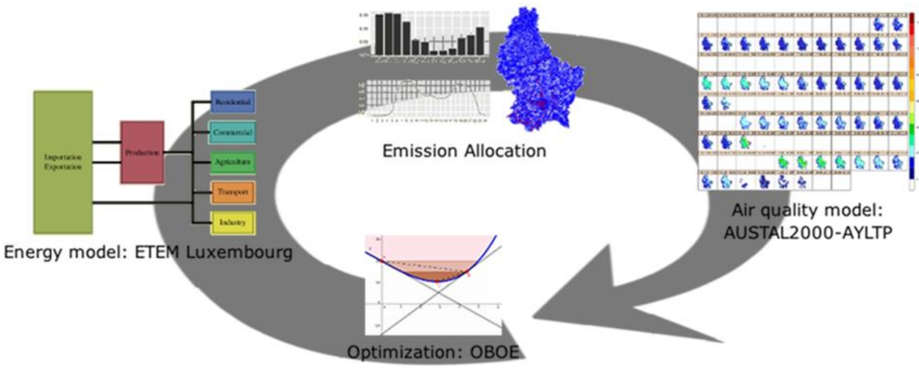
Time evolution of observed and predicted SST anomalies in the Nino 3.4 region (up to 12 lead months) by the NCEP/CPC Markov model (Xue et al. 2000, *J. Climate*, **13**, 849-871).

A comparison with NCEP/CPC Markov model

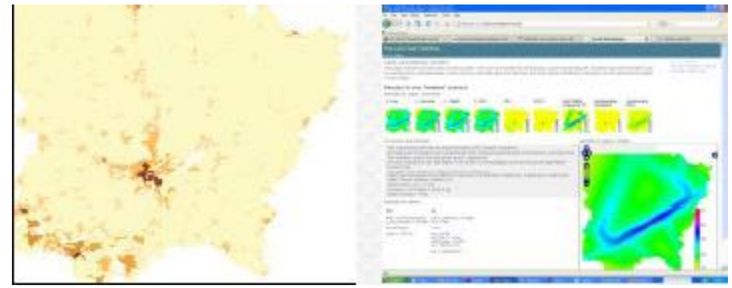


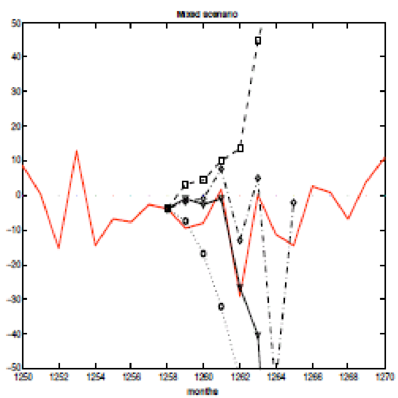
TUDOR Modeling Group – Research snapshot

Integrated assessment: Energy-AQ

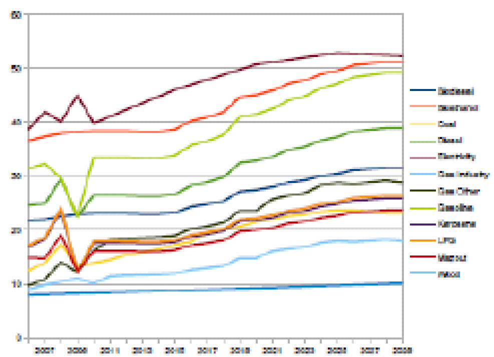


Integrated assessment: Impact



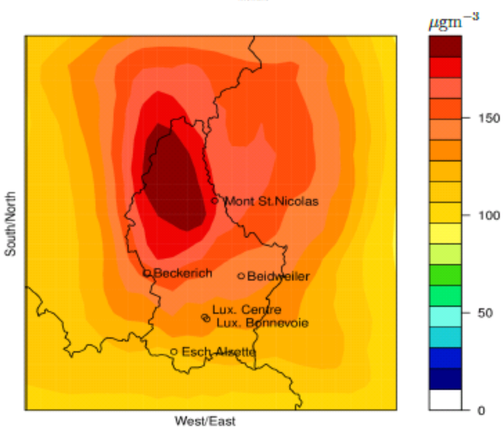


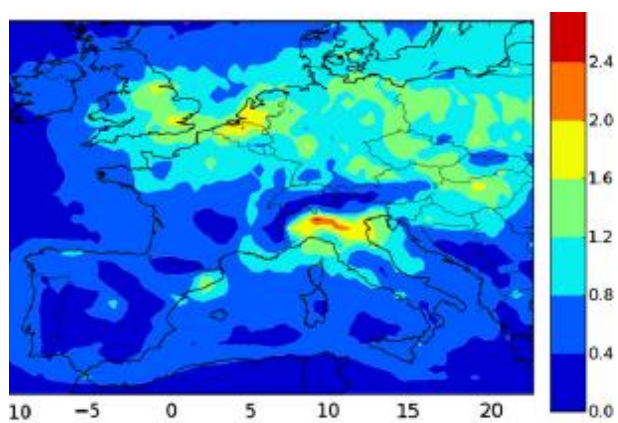
Bayesian risk (atmospheric forecasting)



Energy projections

AQ





Geospatial analysis