A Bayesian Classifier for Climate Model Ensemble Selection

Presentation, Helsinki, August 2012

Peter Friis-Hansen & Luca Garré
2012.08.27

MANAGING RISK

The need for “accurate” climate projections

- Is the uncertainty of climate important? What’s the value of climate uncertainty?

The need for “accurate” climate projections

- Is the uncertainty of climate important? What’s the value of climate uncertainty?

- Santa Barbara desalination plant:
  - From web.mit.edu: “*In 1992, construction of the Santa Barbara Desalination Plant was completed after a five-year drought in California. The facility was expected to produce enough water to compensate for severe loss in groundwater reserves during the drought. […] Once the drought ended, however, the pricey facility was placed in an active standby mode […]. The plant has since been decommissioned…*”
  - Maladaptation, Glob. Env. Change, 2010, Vol. 20: “*[The] desalination plant built in Santa Barbara in 1990 […] has never been used due to ample rainfall since its construction. It is still maintained ostensibly as an ‘insurance policy’ (City of Santa Barbara, 2009), but more likely because it is a sunk cost. Due to ample rainfall during the years following completion, cheaper surface water has been available and it has not been necessary to run the plant.*”

The need for “accurate” climate projections

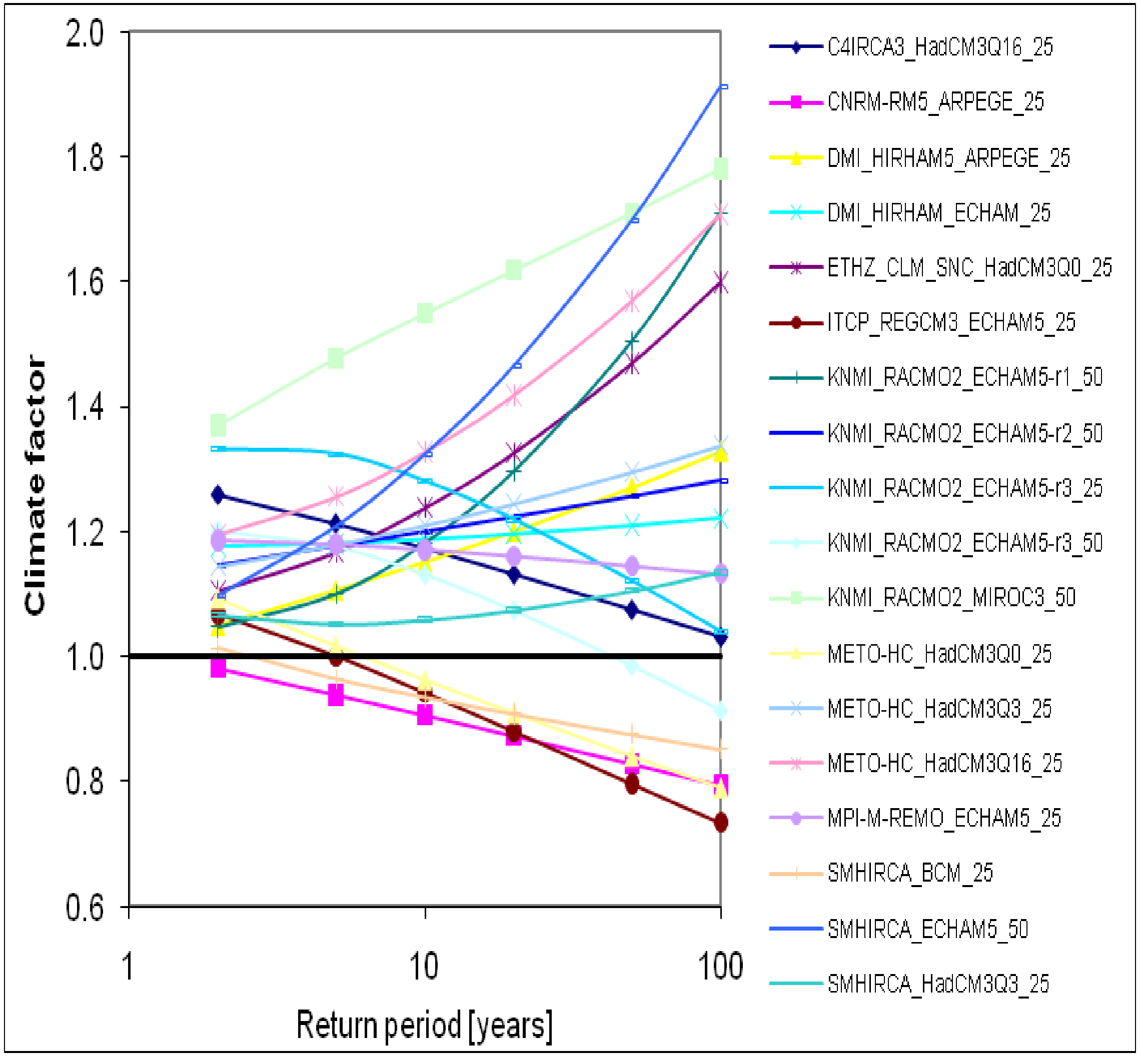
- Is the uncertainty of climate important? What’s the value of climate uncertainty?

- Santa Barbara desalination plant:
  - From web.mit.edu: “*In 1992, construction of the Santa Barbara Desalination Plant was completed after a five-year drought in California. The facility was expected to produce enough water to compensate for severe loss in groundwater reserves during the drought. […] Once the drought ended, however, the pricey facility was placed in an active standby mode […]. The plant has since been decommissioned…*”
  - Maladaptation, Glob. Env. Change, 2010, Vol. 20: “*[The] desalination plant built in Santa Barbara in 1990 […] has never been used due to ample rainfall since its construction. It is still maintained ostensibly as an ‘insurance policy’ (City of Santa Barbara, 2009), but more likely because it is a sunk cost. Due to ample rainfall during the years following completion, cheaper surface water has been available and it has not been necessary to run the plant.*”

- Cost of the plant: 34 million USD

Limitations of climate models - Ensemble variability

- The climate and its modelling is complex
- Current climate models have limitations and if used and interpreted wrongly can introduce gross errors
- The figure shows anticipated precipitation in the city of Århus ➔ need a method to select and weigh different climate models



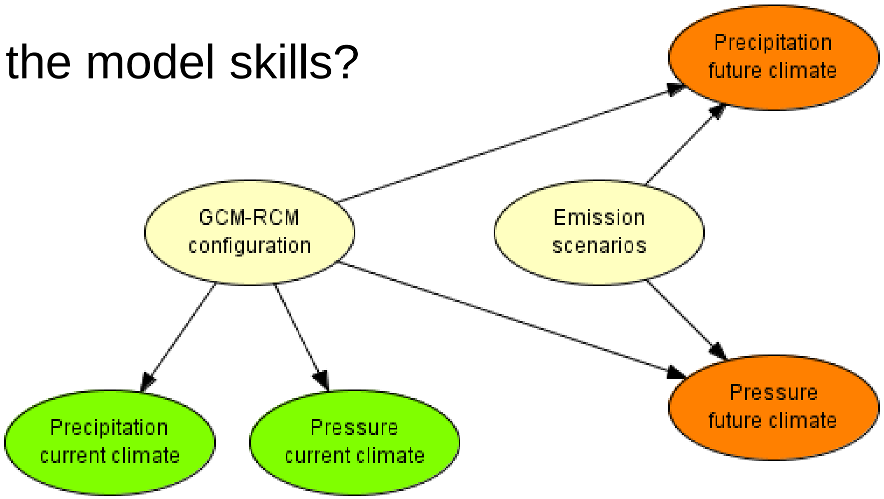
Data courtesy of DHI, Denmark.

Climate change: GCM classifier

- Citation from Tebaldi & Knutti (2007):

  - «*There is evidence that combining multi-model ensembles increases the skill, reliability and consistency of model forecasts*»
  - «*Combining information from several models is reported to be superior to a single-model forecast*»
  - «*Intuitively, it makes perfect sense to trust and hence weigh the better models more than the poor models*»

- How may we identify and quantify the model skills?



Tebaldi, C. and Knutti, R. (2007). ***The use of the multi-model ensemble in probabilistic climate projections*****. Philosophical transactions of the royal society A.**

Bayes rule

Bayes rule:

$$P(A|B) = \frac{1}{P(B)} P(B|A)P(A)$$

Reverend Thomas Bayes 1702-1761

Bayes rule

Reverend Thomas Bayes 1702-1761

Bayes rule:

$$P(A|B) = \frac{1}{P(B)} P(B|A) P(A)$$

Posterior distribution of models

Likelihood function of precipitation

Prior distribution of models

MANAGING RISK

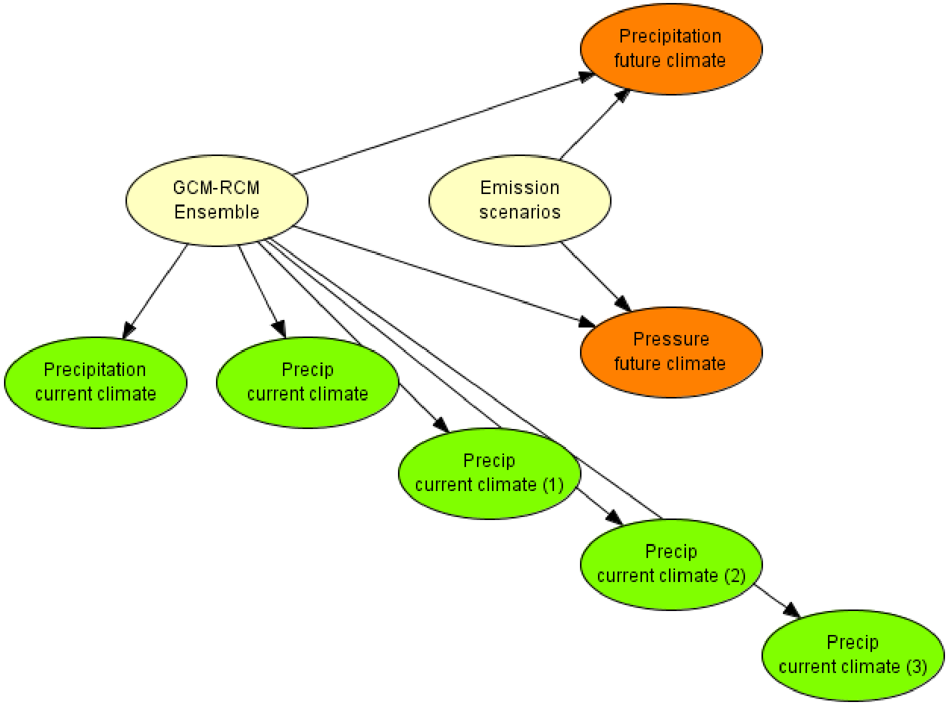
The math in the Bayesian updating…

-
- $P[\boldsymbol{M}|obs_1, obs_2, \dots, obs_i \dots obs_n] \propto$

$$\prod_{models\ m=1..l} \left\{ \int_{all\ x} f_m(x) \cdot f_{obs_i(x)} \mathrm{d}x \ \cdot P[model_m] \right\}$$

- In which,
- $f_m(x)$ is the distribution of each climate model prediction,
- $f_{obs_i(x)}$ is the observed distribution of the environmental parameter, and
- $P[model_m]$ is the prior probability of the plausibility of the climate model
- A small amount (3%) of colored noise (1/x) is added to the model-predicted distributions to avoid singularites in the likelihood function

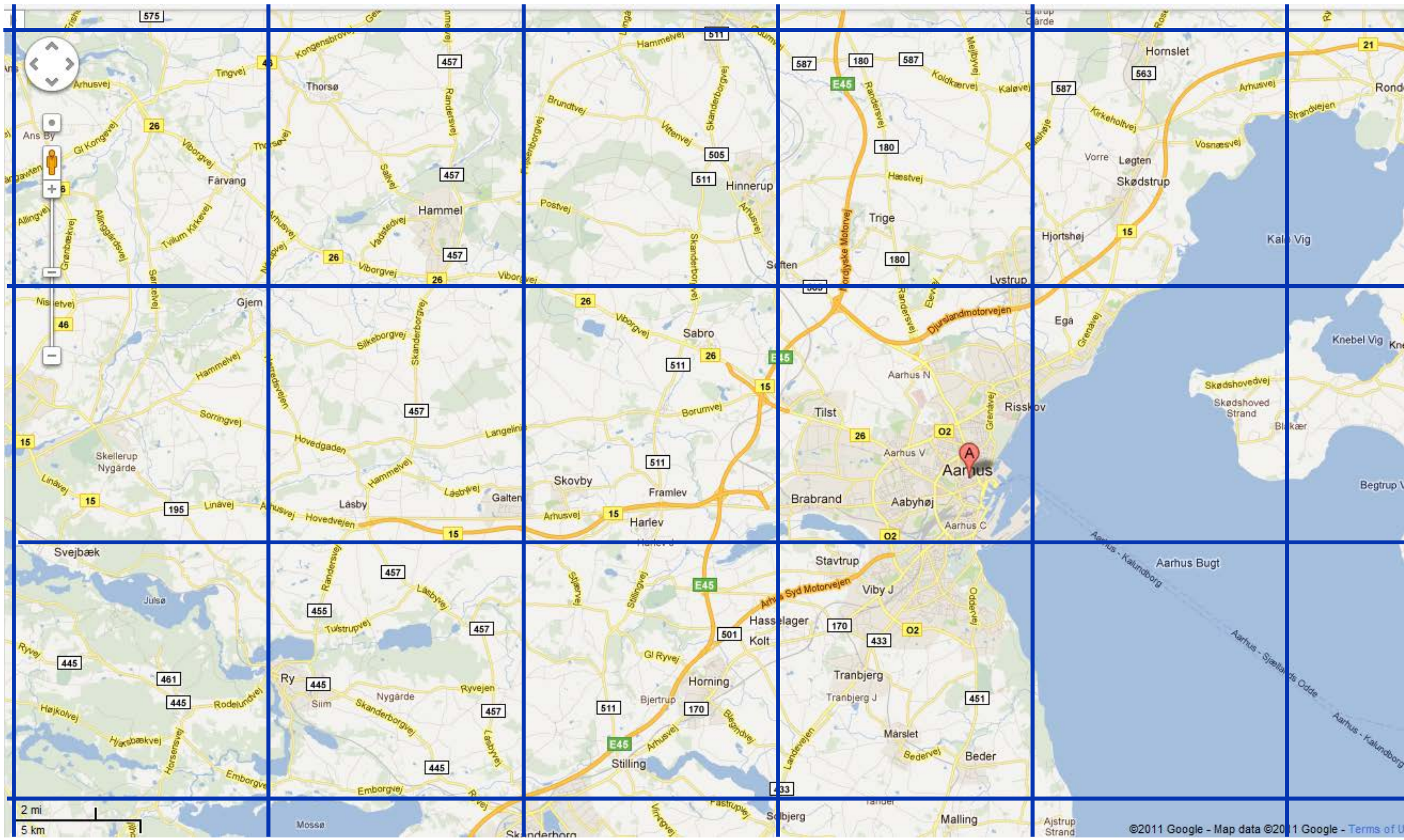
Bayesian classifier for Århus emsemble



One advantage of using a Bayesian classifier, compared to simple averaging of model predictions, is that the classifier can take into account different environmental quantities (precipitation, wind, temperature, etc.) that are of importance to the problem at hand

A 10 km grid on the Århus area

Approximately 5 rain gauges in area (DMI webpage)



Prior for node GCM-RCM from expert judgment

GCM-RCMEnsemble(C1)

C4IRCA3 HadC	0.040486
CNRM-RM5 AR	0.020243
DMI HIRHAM5	0.020243
DMI HIRHAM5	0.020243
DMI HIRHAM5	0.089069
ETHZ CLM SN	0.030364
ITCP REGCM3	0.089069
KNMI RACMO2	0.089069
KNMI RACMO2	0.089069
KNMI RACMO2	0.089069
KNMI RACMO2	0.089069
KNMI RACMO2	0.020243
METO-HC Had	0.040486
METO-HC Had	0
METO-HC Had	0.020243
MPI-M-REMO E	0.07085
SMHIRCA BCM	0.020243
SMHIRCA ECH	0.07085
SMHIRCA ECH	0.07085
SMHIRCA Had	0.020243

Predicted current climate from each model configuration

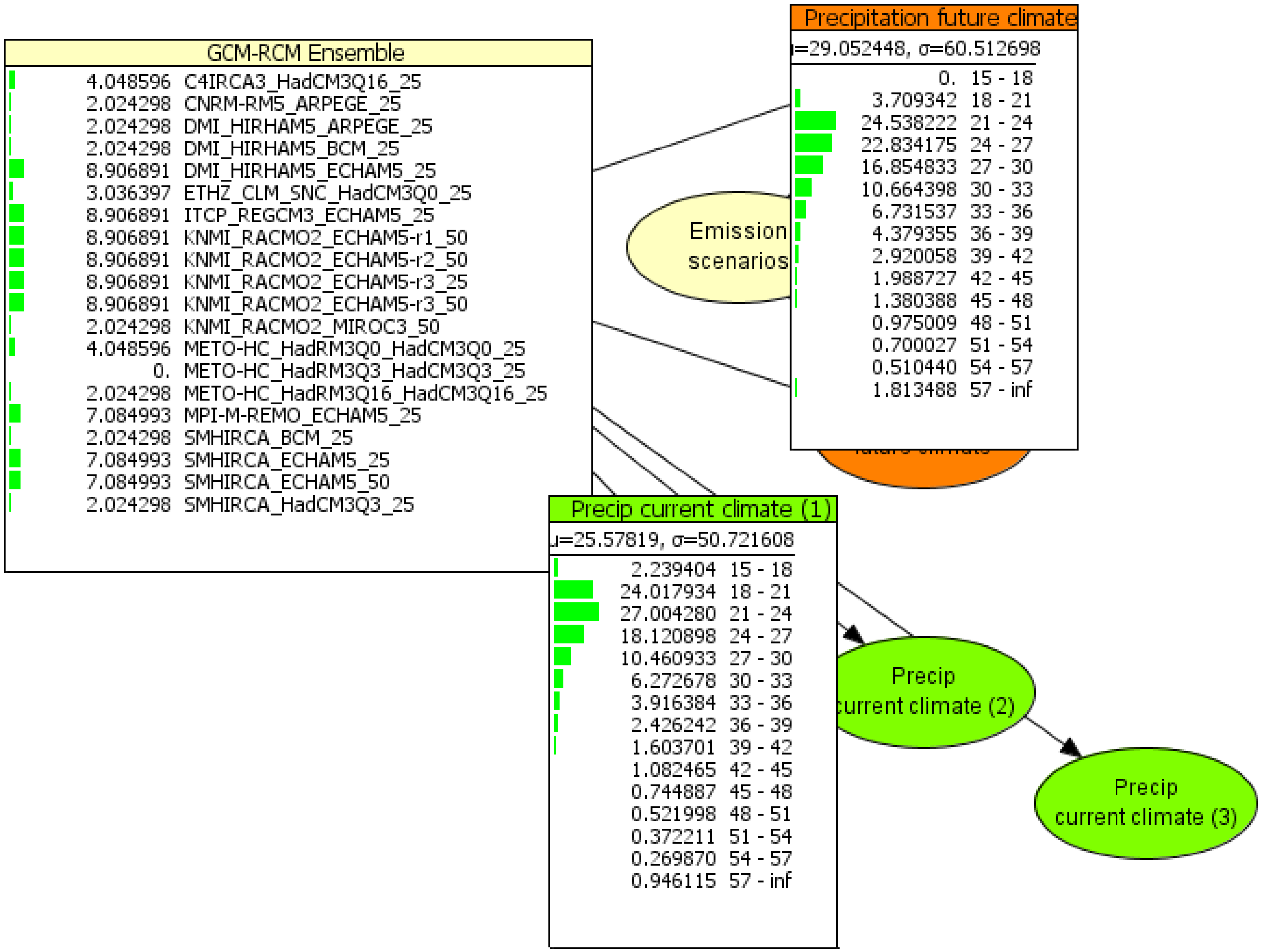
Precipitationcurrent climate(C2)

C1	C4IRCA3 H	CNRM-RM5	DMI HIRH/	DMI HIRH/	DMI HIRH/	ETHZ CLM	ITCP REG(	KNMI RAC	KNMI RAC	KNM
15 - 18	0.169864	0	0	0	0	0	0	0	0	
18 - 21	0.346434	0.396037	0.022205	0.202066	0	0.209642	0	0.175205	0.118284	
21 - 24	0.191243	0.24356	0.357609	0.262142	0.144045	0.254665	0.375241	0.316351	0.341713	
24 - 27	0.109838	0.13225	0.226902	0.178921	0.356538	0.170225	0.239297	0.186783	0.207859	
27 - 30	0.065307	0.076779	0.143939	0.121168	0.185644	0.114551	0.135163	0.113372	0.126987	
30 - 33	0.040035	0.047031	0.091291	0.081389	0.105259	0.077589	0.080909	0.070547	0.07791	
33 - 36	0.025218	0.030104	0.057888	0.054207	0.063714	0.052885	0.050778	0.044897	0.048001	
36 - 39	0.016275	0.019989	0.036699	0.035784	0.040617	0.036268	0.033145	0.029164	0.029696	
39 - 42	0.010736	0.013692	0.023262	0.023404	0.027006	0.025019	0.022364	0.019302	0.018446	
42 - 45	0.007223	0.009632	0.014741	0.015159	0.018593	0.017358	0.015524	0.012995	0.011503	
45 - 48	0.004948	0.006935	0.00934	0.00972	0.013182	0.01211	0.011044	0.008888	0.007202	
48 - 51	0.003446	0.005096	0.005916	0.006166	0.009583	0.008494	0.008027	0.006168	0.004526	
51 - 54	0.002436	0.003812	0.003747	0.003869	0.007119	0.005989	0.005945	0.004339	0.002855	
54 - 57	0.001746	0.002897	0.002372	0.002399	0.005389	0.004245	0.004478	0.00309	0.001808	
57 - inf	0.005252	0.012187	0.004089	0.003605	0.023312	0.010961	0.018086	0.008898	0.003209	

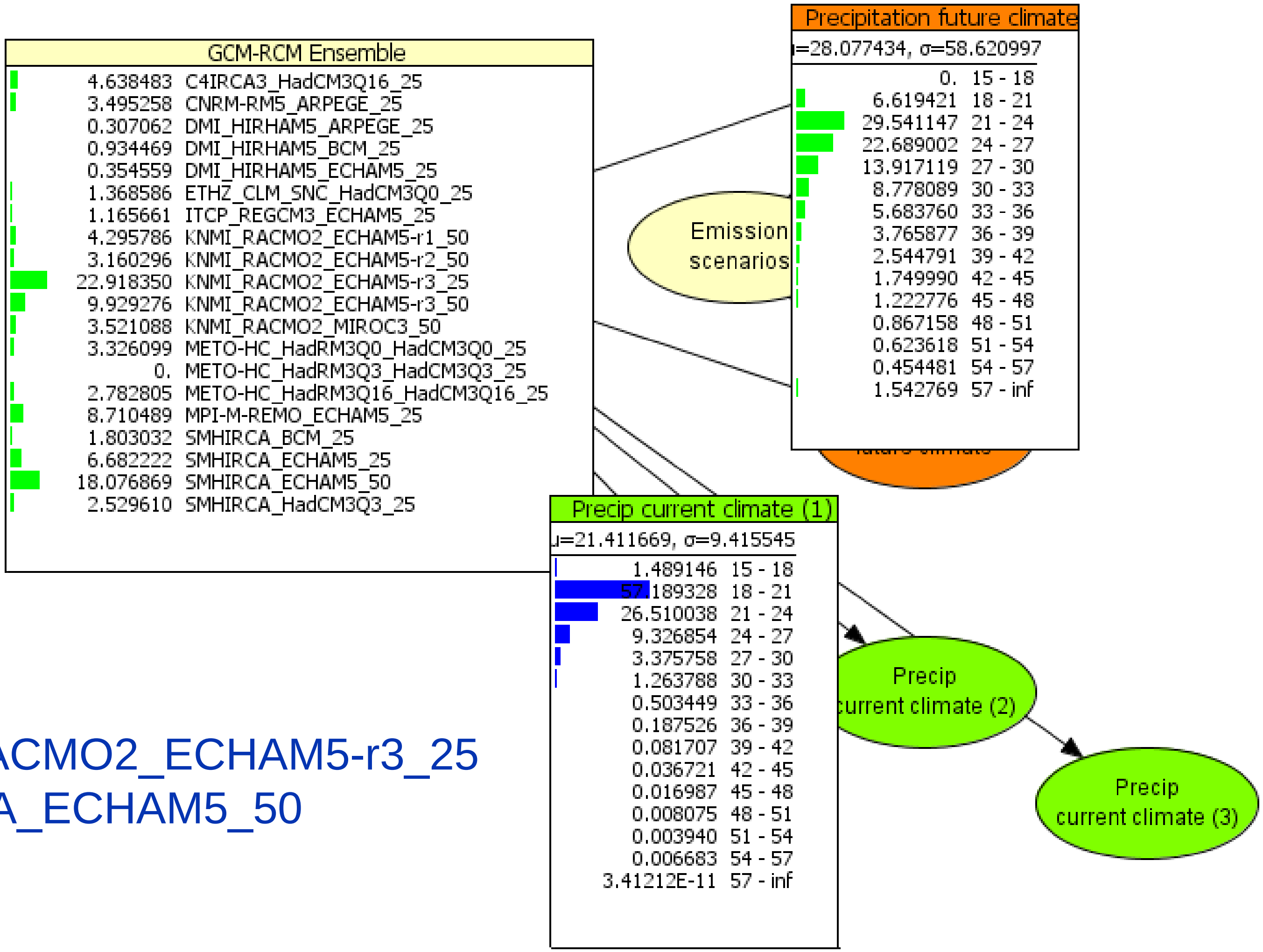
C1	KNMI RAC	KNMI RAC	KNMI RAC	METO-HC	METO-HC	METO-HC	MPI-M-REN	SMHIRCA	SMHIRCA	SMH
15 - 18	0.066482	0	0.030518	0	0	0.036114	0	0.192292	0	
18 - 21	0.475917	0.27319	0.412317	0.274054	0.438987	0.372476	0.321376	0.319319	0.296008	
21 - 24	0.204071	0.341382	0.214779	0.271036	0.248941	0.217403	0.273279	0.180081	0.255854	
24 - 27	0.100375	0.157385	0.120642	0.158597	0.137351	0.131084	0.157576	0.106491	0.160686	
27 - 30	0.054486	0.082292	0.071931	0.096875	0.076279	0.081347	0.093088	0.065553	0.101706	
30 - 33	0.031859	0.047097	0.045016	0.061376	0.042633	0.051795	0.056226	0.041766	0.064862	
33 - 36	0.019743	0.028854	0.029325	0.040128	0.023977	0.033748	0.03466	0.027417	0.041668	
36 - 39	0.01282	0.018643	0.019759	0.026966	0.013568	0.022451	0.021771	0.018474	0.026957	
39 - 42	0.00865	0.012571	0.013703	0.018562	0.007724	0.015221	0.013915	0.012739	0.01756	
42 - 45	0.006027	0.00878	0.009742	0.013053	0.004423	0.010498	0.009039	0.008966	0.011514	
45 - 48	0.004315	0.006315	0.007078	0.009355	0.002547	0.007355	0.00596	0.006428	0.007599	
48 - 51	0.003163	0.004656	0.005242	0.00682	0.001475	0.005229	0.003985	0.004685	0.005046	
51 - 54	0.002366	0.003507	0.003949	0.005049	0.000859	0.003767	0.0027	0.003466	0.003371	
54 - 57	0.001801	0.002691	0.00302	0.003791	0.000503	0.002747	0.001852	0.0026	0.002266	
57 - inf	0.007924	0.012638	0.01298	0.014338	0.000733	0.008764	0.004574	0.009722	0.004902	

C1	SMHIRCA	SMHIRCA
15 - 18	0.00296	0.204728
18 - 21	0.42432	0.351616
21 - 24	0.266945	0.191379
24 - 27	0.157382	0.106234
27 - 30	0.085142	0.060066
30 - 33	0.040884	0.034556
33 - 36	0.022367	0.020206
36 - 39	0	0.011998
39 - 42	0	0.007228
42 - 45	0	0.004415
45 - 48	0	0.002732
48 - 51	0	0.001711
51 - 54	0	0.001085
54 - 57	0	0.000695
57 - inf	0	0.001351

GCM Classifier



Updated

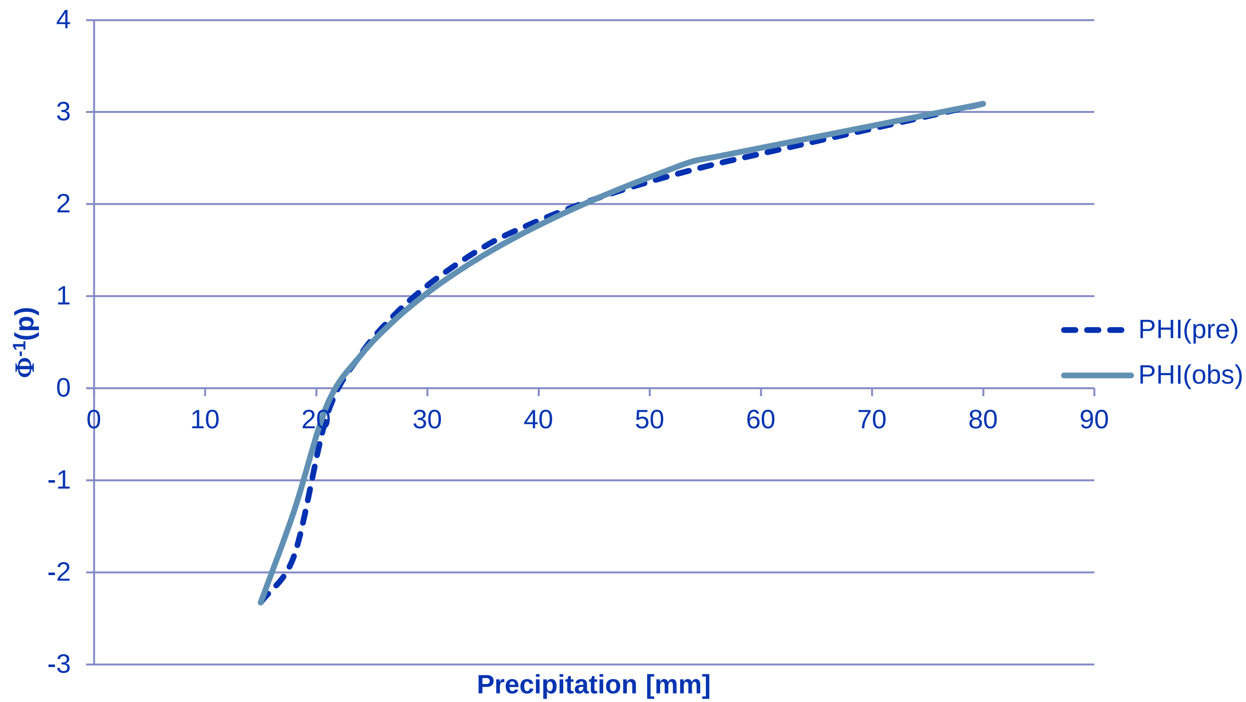


KNMI_RACMO2_ECHAM5-r3_25
SMHIRCA_ECHAM5_50

Bayesian Classifier – comparing observations and predictions

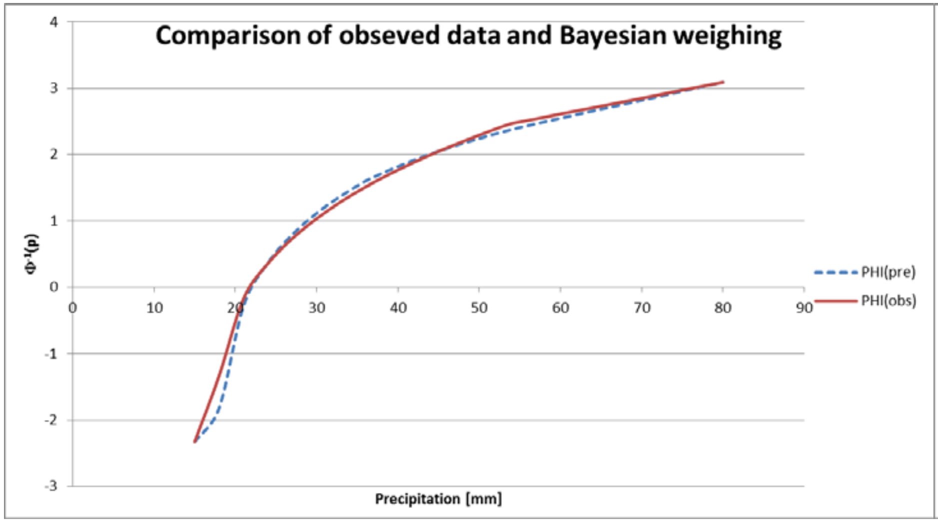
Weight		
22,92		KNMI_RACMO2_ECHAM5-r3_25
18,08		SMHIRCA_ECHAM5_50
9,93		KNMI_RACMO2_ECHAM5-r3_50
8,71		MPI-M-REMO_ECHAM5_25
6,68		SMHIRCA_ECHAM5_25
4,64		C4IRCA3_HadCM3Q16_25
4,30		KNMI_RACMO2_ECHAM5-r1_50
3,52		KNMI_RACMO2_MIROC3_50
3,50		CNRM-RM5_ARPEGE_25
3,33		METO-HC_HadRM3Q0_HadCM3Q0_25
3,16		KNMI_RACMO2_ECHAM5-r2_50
2,78		METO-HC_HadRM3Q16_HadCM3Q16_25
2,53		SMHIRCA_HadCM3Q3_25
1,80		SMHIRCA_BCM_25
1,37		ETHZ_CLM_SNC_HadCM3Q0_25
1,17		ITCP_REGCM3_ECHAM5_25
0,93		DMI_HIRHAM5_BCM_25
0,35		DMI_HIRHAM5_ECHAM5_25
0,31		DMI_HIRHAM5_ARPEGE_25
0,00		METO-HC_HadRM3Q3_HadCM3Q3_25

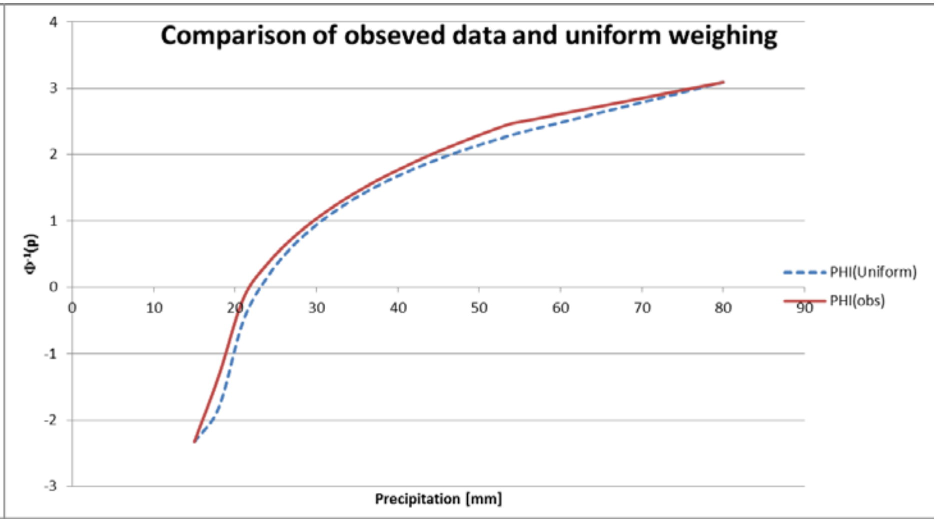
Solid line is observed precipitation
Dashed line results from the Bayesian classifier



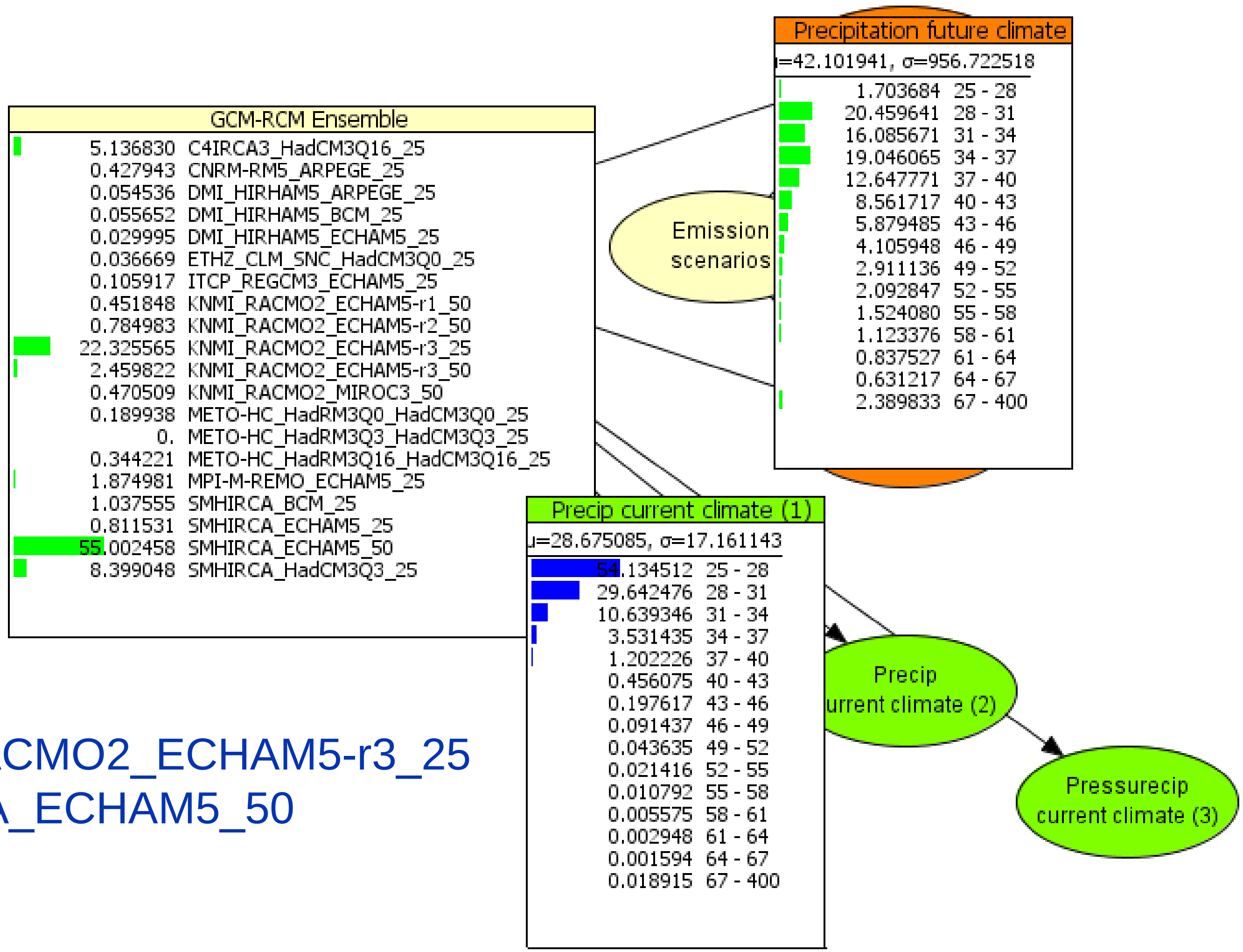
Comparing Bayesian and uniform weighing

Bayesian weighing yield better results, although the uniform weighing performs well, but slightly biased





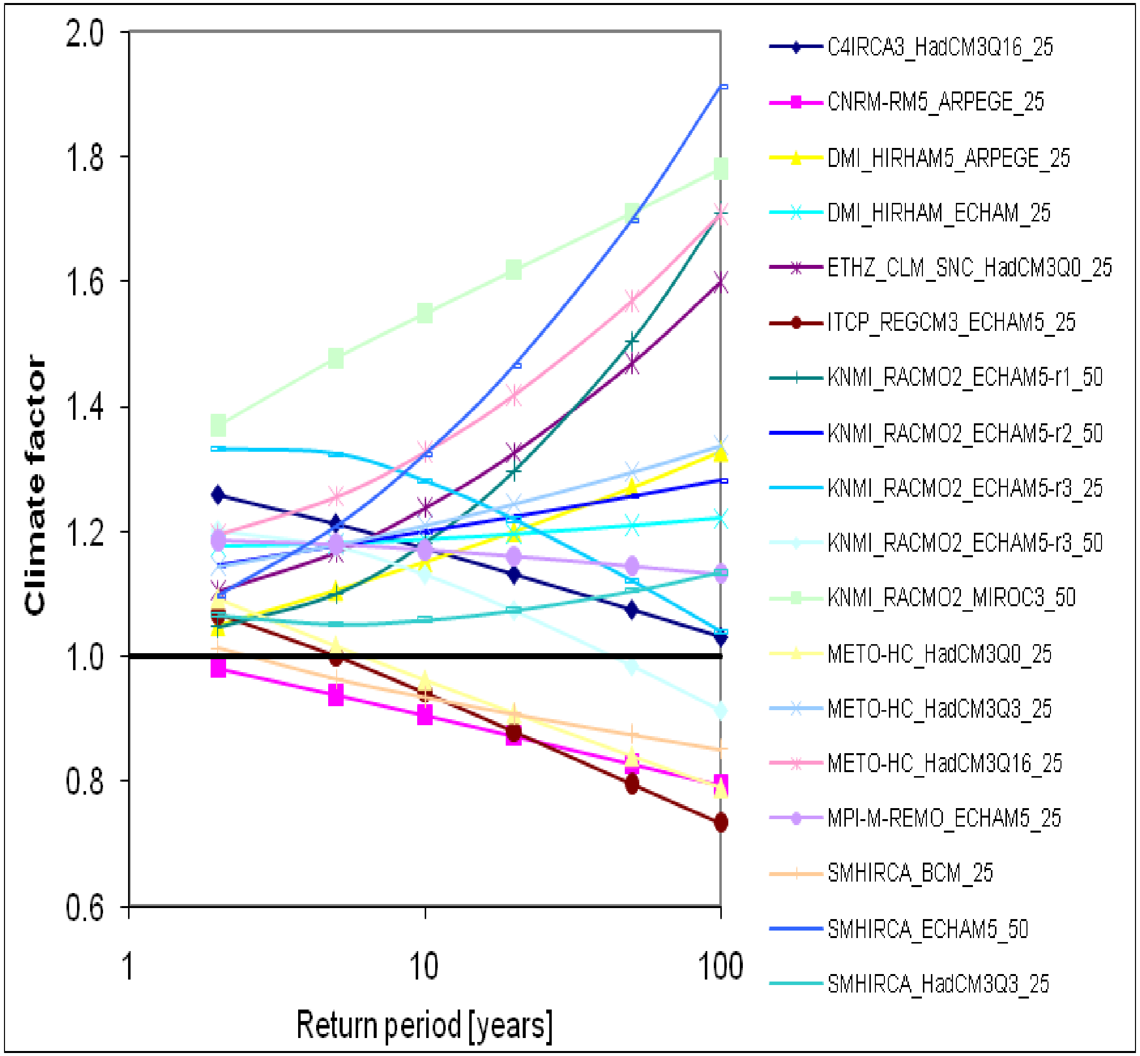
Truncation at 80% fractile – when the upper tail is important



KNMI_RACMO2_ECHAM5-r3_25
SMHIRCA_ECHAM5_50

Conclusion

- Bayesian network offers a powerful tool for modelling complex systems
- Bayesian network offers a good test bench for testing updating procedure
- The useful classifier-model (may/ will) depend on the task of the pursuing study



Safeguarding life, property and the environment

www.dnv.com