Prediction of turbidity using neural networks: application on chalk aquifer of Normandy (France)

Michaël Savary[1,2], Anne Johannet[2], Nicolas Massei[1], Jean-Paul Dupont[1], Emmanuel Hauchard[1,3]

[1]Rouen University, M2C Laboratory France

[2]Ecole des mines d'Alès, LGEI, France

[3] Communauté d'Agglomération Havraise, France

Yport cliff

Abstract N°2338

Early Career Hydrogeologists' Network

25-29th September 2016

Montpellier, France

43rd IAH congress

Turbidity

Definition, origins and properties

- Ability of a liquid to disrupt a light beam.
- Turbidity essentially comes from two processes; the surface run-off or/and the underground matter resuspension.
- Could be an indicator of surface contaminations (fecal particles, phytosanitary products,…).
- Non linear response to rainfalls.

Yport localization

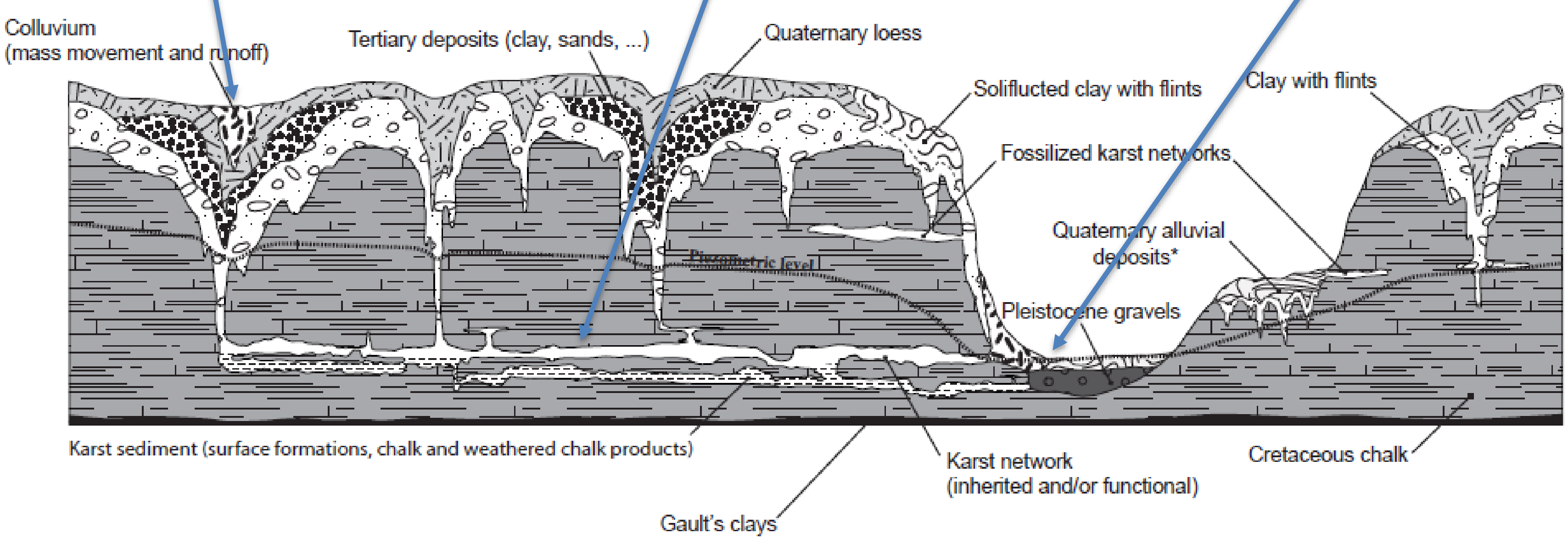
ECHN Early Career Hydrogeologists' Network
iah • aih 60 years
25-29th September 2016
Montpellier, France
CORUM CONFERENCE CENTER
43rd IAH congress

Normandie hydrogeological context

Bétoires

Karst duct

Yport fountain



Schematic cut of the superficial formation in the Seine Maritime (Chédeville, S. & al. 2014a)

Yport fresh water pumping well

Stakes

- Responsible for the half of Le Havre/Great Le Havre drinking water suppling (180,000 inhabitant)
- Regularly contaminated by turbid water (potability norms < 1 NTU)
- Risk of phytosanitary product contamination

100 NTU

Managers needs

- Forecast 12h and 24h ahead a 100NTU threshold overtaking.
- Water storage to avoid turbid water pumping
- Clean processing chain to maximize the filtration if pumping is necessary

Yport drinking water station

Database

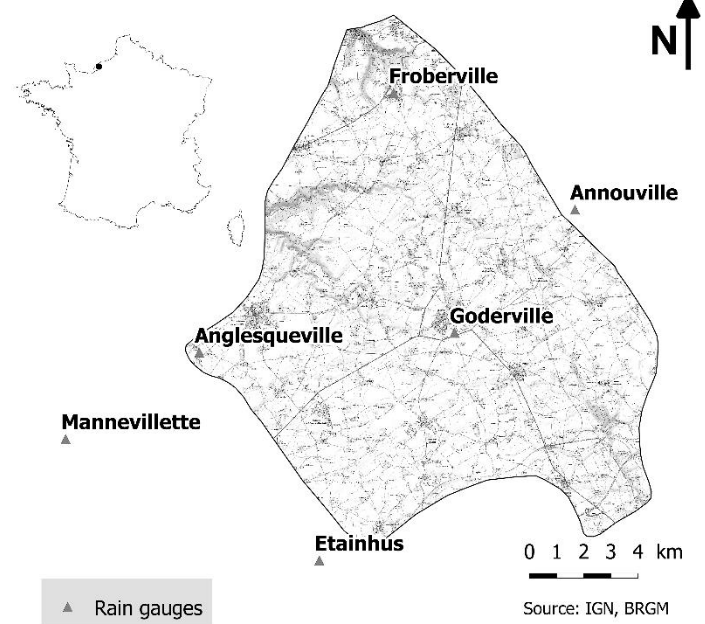
6 rain gauges

- Data since 2009 at hourly time step (with gaps)

1 turbidimeter

- At the input of the station : data since 1993 at hourly time step (with gaps)

Underground discharge cannot be measured



22 events selected

Event without turbidity peak					Event with turbidity peak				
Event	Duration (hours)	Turbidity (NTU)		Rain (mm)	Event	Duration (hours)	Turbidity (NTU)		Rain (mm)
		Max	Min	Cumul			Max	Min	Cumul
1	288	7.07	1	15.9	2	624	302.48	1.54	41.3
4	384	9.82	0.91	14.1	3	1008	135.03	0	26.7
5	336	7.71	1.52	17.5	6	720	245.38	1.53	42.0
8	360	26.87	0.97	22.5	7	744	84.67	0.05	19.2
9	384	9	1.00	20.2	10	576	256.15	0.92	24.9
12	456	12	0.84	28.7	11	744	307.89	0.87	54.8
13	576	13	0.86	30.8	16	648	405.25	0.81	53.8
14	384	14	0.86	23.3	17	744	157.45	0.49	50.7
15	600	15	0.85	31.7	18	744	86.67	2.18	42.8
19	504	19	1.50	30.3	22	623	53.91	0.80	44.2
20	576	20	0.89	40.8					
21	600	48.44	0.93	48.5					

Turbidity forecasting: state of the art

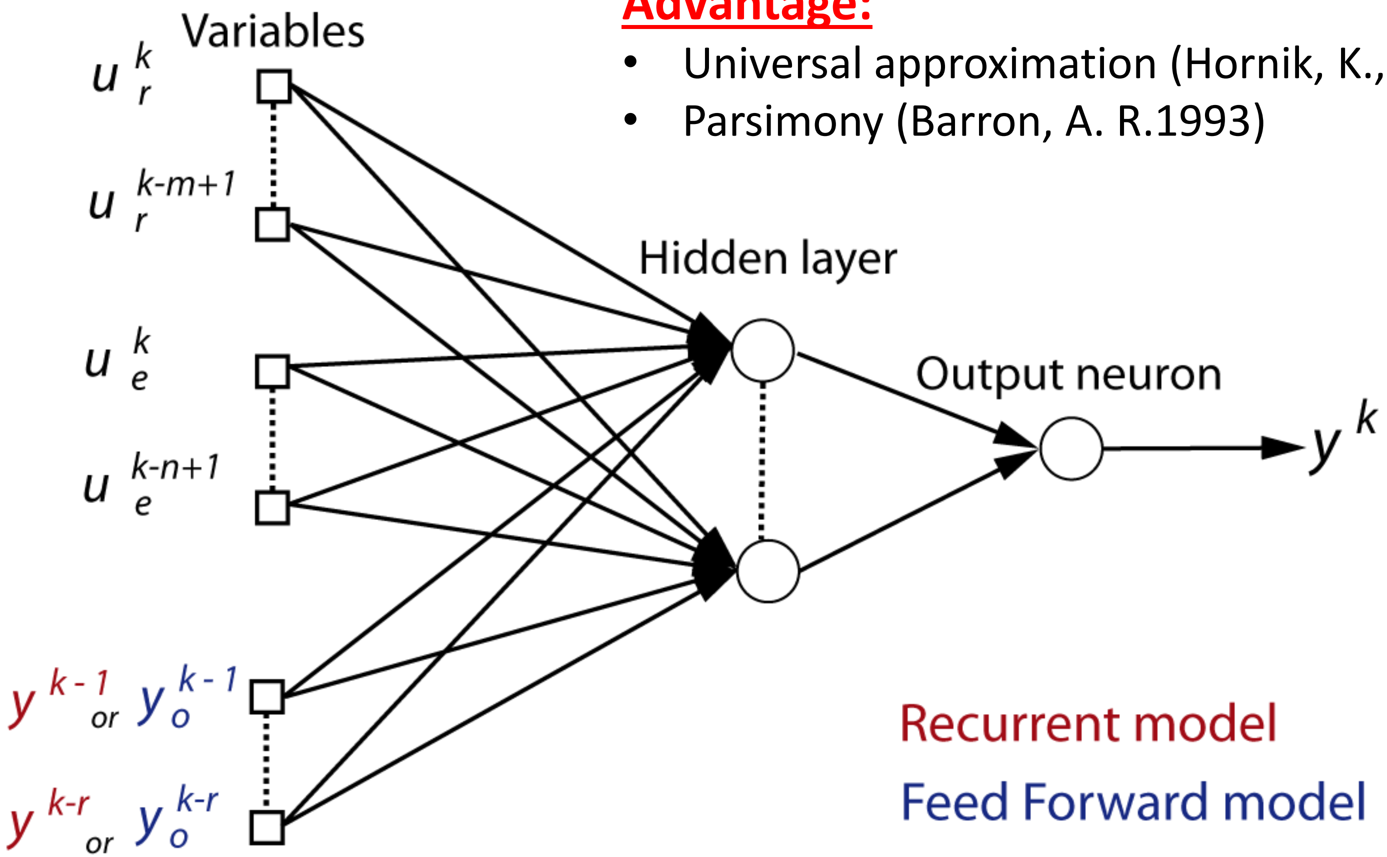
- Turbidity results from complex and difficult to observe processes
  - ➢ No physical model
  - ➢ Black box
- Discharge/Turbidity relation models
  - ➢ Neural network and particle swarm optimization with daily data (Houria, B & al 2014)
  - ➢ Neural network with hourly data (Mulia, I. E. & al. 2013)
- Rainfall/Turbidity relation models
  - ➢ Local logistic regression with monthly data (Towler ,E. & al 2010)
  - ➢ Adjust probabilistic distribution to daily data (Y. Tramblay. 2007)

Goals of the present study

Forecast turbidity using rainfall at 12h lead time

Neural network models

Multilayer perceptron architecture



Advantage:

- Universal approximation (Hornik, K., & *al* 1989)
- Parsimony (Barron, A. R.1993)

Recurrent model

Feed Forward model

Limits:

- Bias-variance dilemna (Geman, S. & *al* 1992)
- Overfiting (Sjöberg, J., & *al* 1995)

u_r: Rainfalls

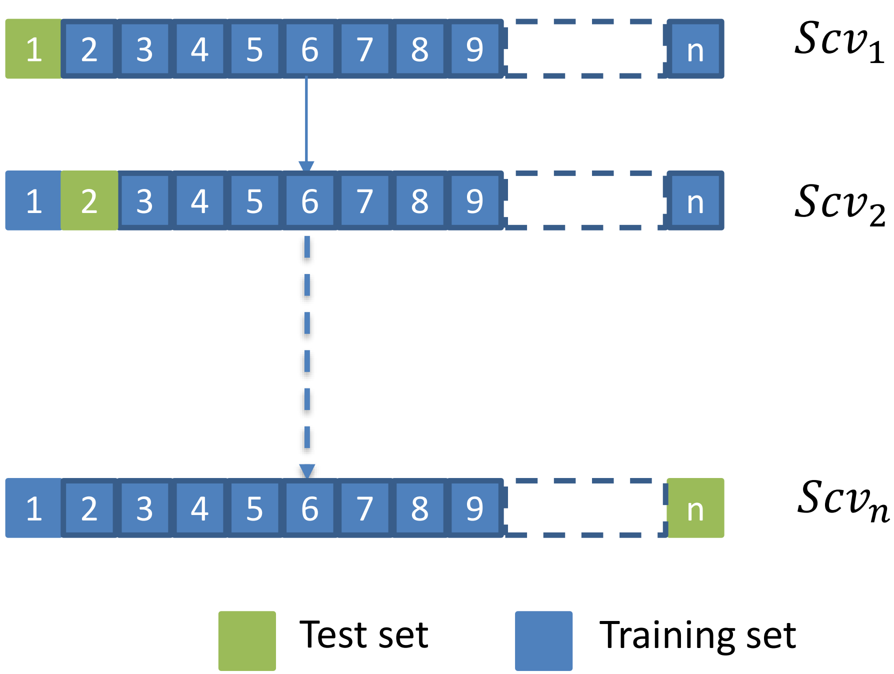
u_e: ETP

y: Turbidity

Regularization methods

To deal with bias-variance dilemma

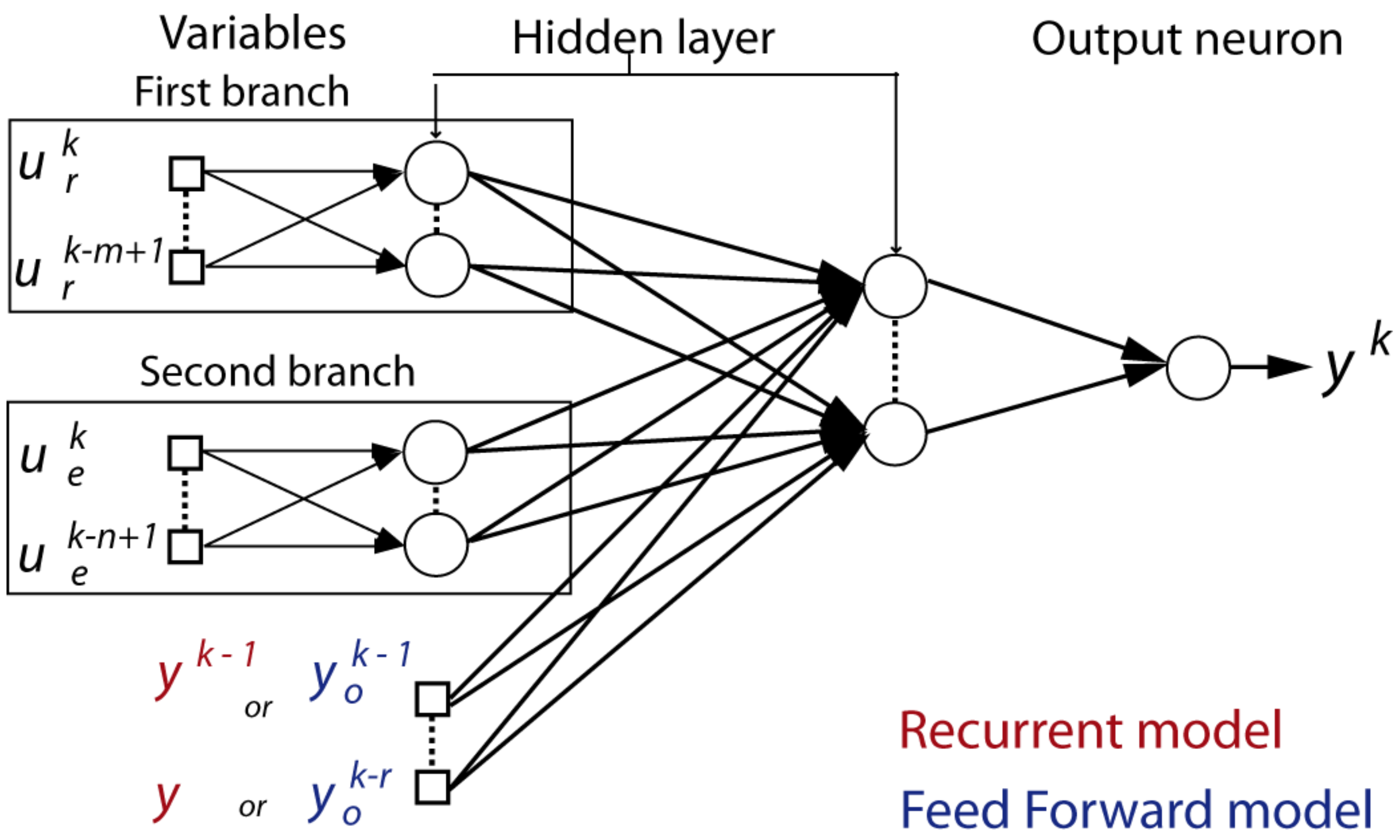
- Cross-validation (Stone, M. 1974)
- Early stopping (Morgan, N., & Bourlard, H. 1989, Sjöberg, J., & *al* 1995).



$$Scv_{validation} = \frac{\sum_{i=1}^{n} Scv_i}{n}$$

Neural network models

Multilayer perceptron based architecture: Two-branches model (Johannet, A., & *al* 2008)



$\mathbf{u_r}$: Rainfalls

$\mathbf{u_e}$: ETP

y: Turbidity

Designed models

6 rain gauges

- Data at hourly time step (with gaps)

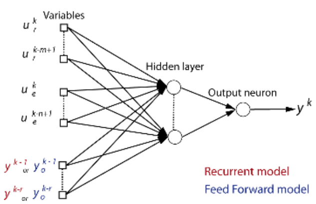
1 turbidimeter

- At the input of the station : hourly time step

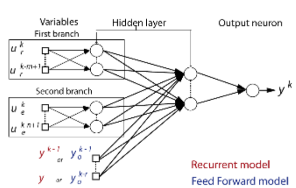
Goal of the models:

- Complexity decreasing by six hours sampling
- Evaluation of the spatialization of rainfalls interest

- Multilayer perceptron
  - Feed-forward
    - 1 hour time step
      - Mean rainfall
      - Spatialized rainfall
    - 6 hours time step
      - Mean rainfall
      - Spatialized rainfall
  - Recurrent
    - 1 hour time step
      - Mean rainfall
      - Spatialized rainfall
    - 6 hours time step
      - Mean rainfall
      - Spatialized rainfall



- Two-branches model
  - Feed-forward
    - 1 hour time step
      - Mean rainfall
      - Spatialized rainfall
    - 6 hours time step
      - Mean rainfall
      - Spatialized rainfall
  - Recurrent
    - 1 hour time step
      - Mean rainfall
      - Spatialized rainfall
    - 6 hours time step
      - Mean rainfall
      - Spatialized rainfall



16 models

Quality criteria

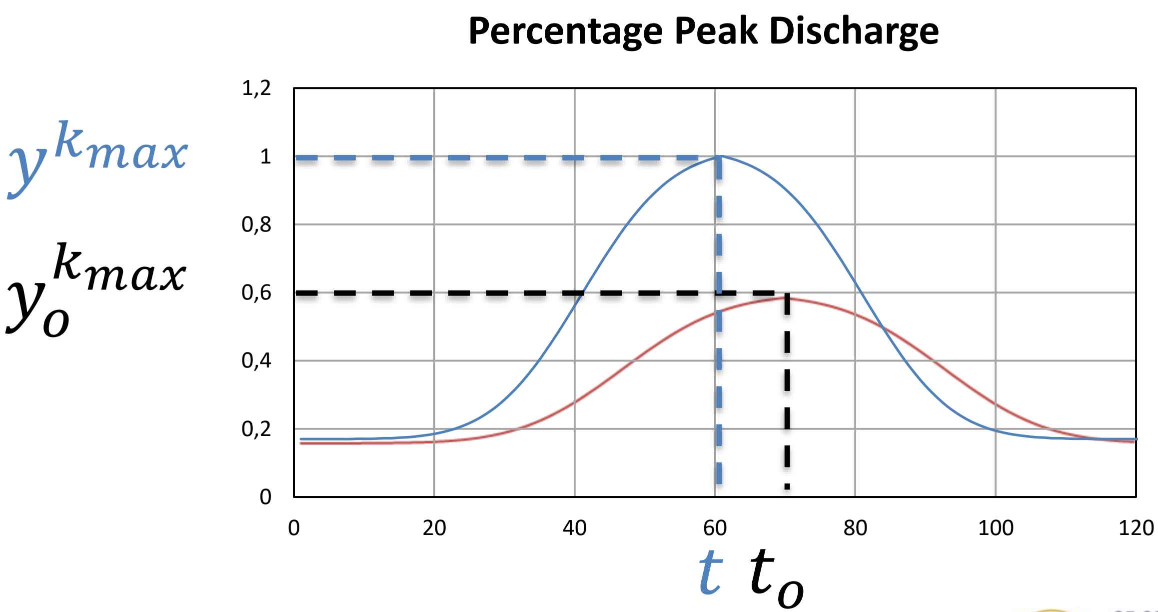
Percentage of the Peak Discharge and the Peak Delay

$$P_{PD} = 100 \frac{y^{k_{\max}}}{y_o^{k_{\max}}}$$

PPD (Percentage of the Peak Discharge)
(Artigue G., & *al* (2012))

$$P_D = t - t_o$$

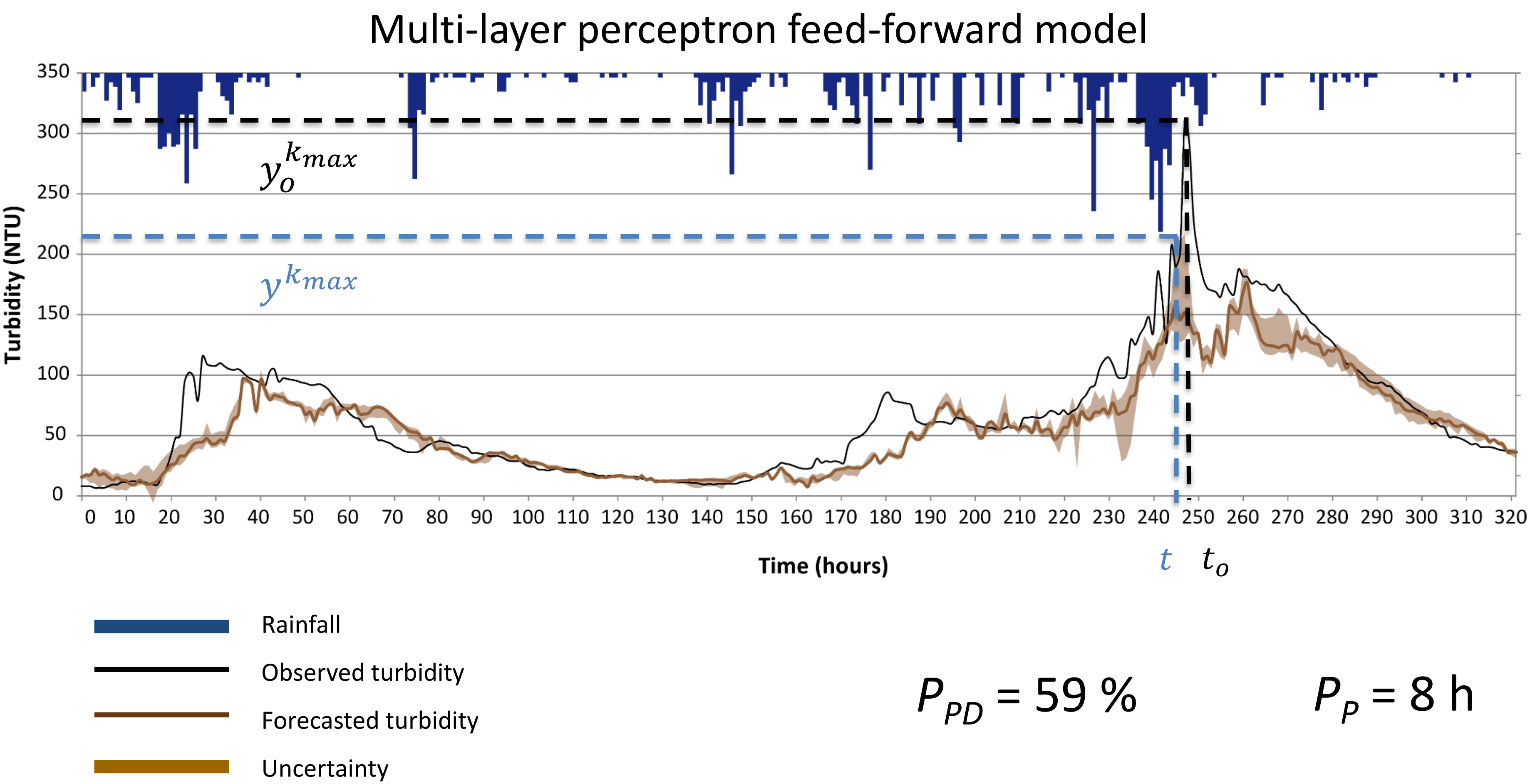
PD (Peak Delay)



Results

Illustration on test event n°11

(01/12/2011 to 30/12/2011)



P_{PD} = 59 %

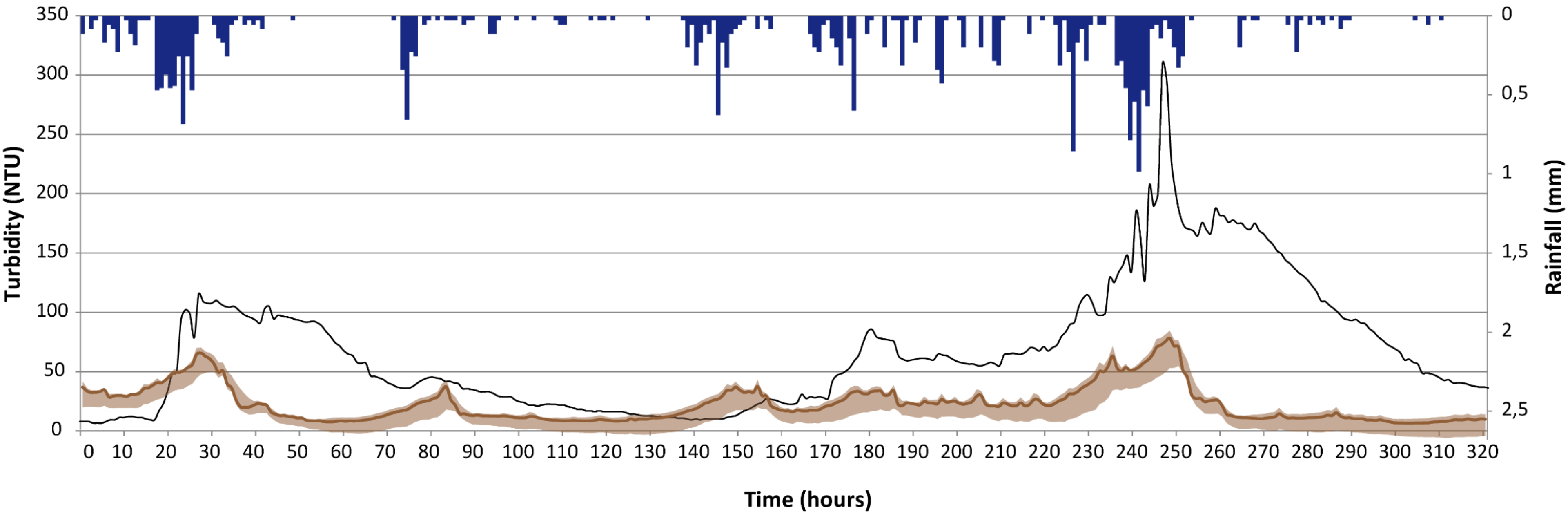
P_P = 8 h

Results

Illustration on test event n°11

(01/12/2011 to 30/12/2011)

Multi-layer perceptron recurrent model



Rainfall

Observed turbidity

Forecasted turbidity

Uncertainty

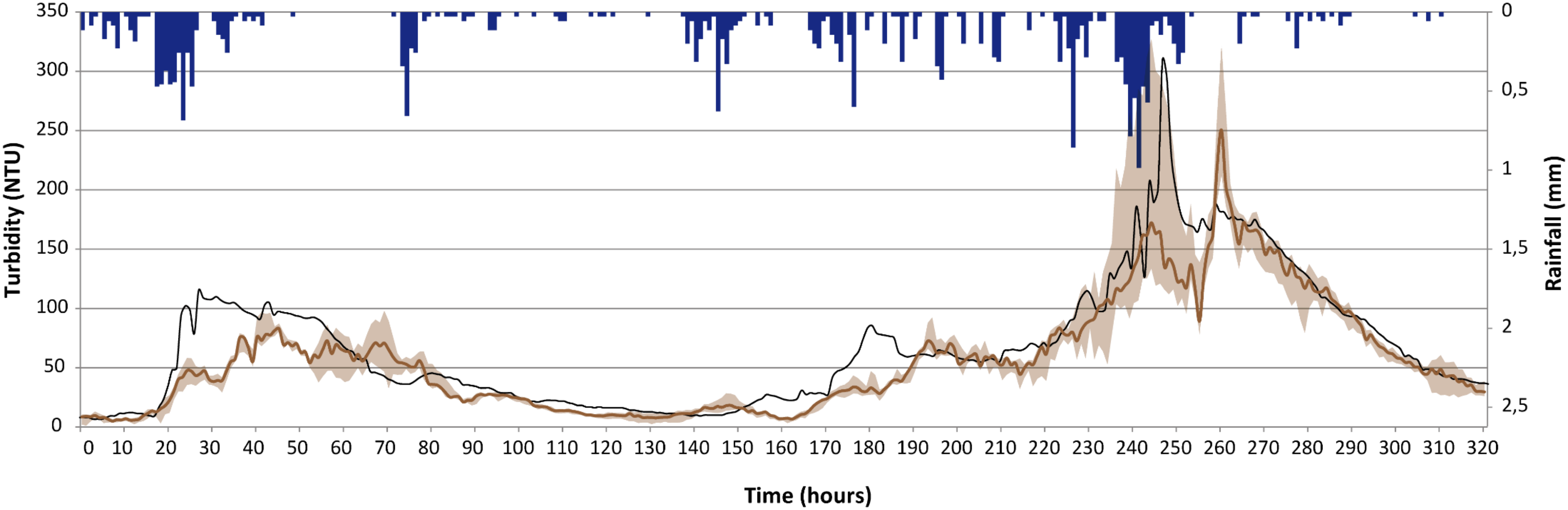
$P_{PD} = 25\ \%$ $P_P = 3\ h$

Results

Illustration on test event n°11

(01/12/2011 to 30/12/2011)

Two branches feed-forward model



Rainfall

Observed turbidity

Forecasted turbidity

Uncertainty

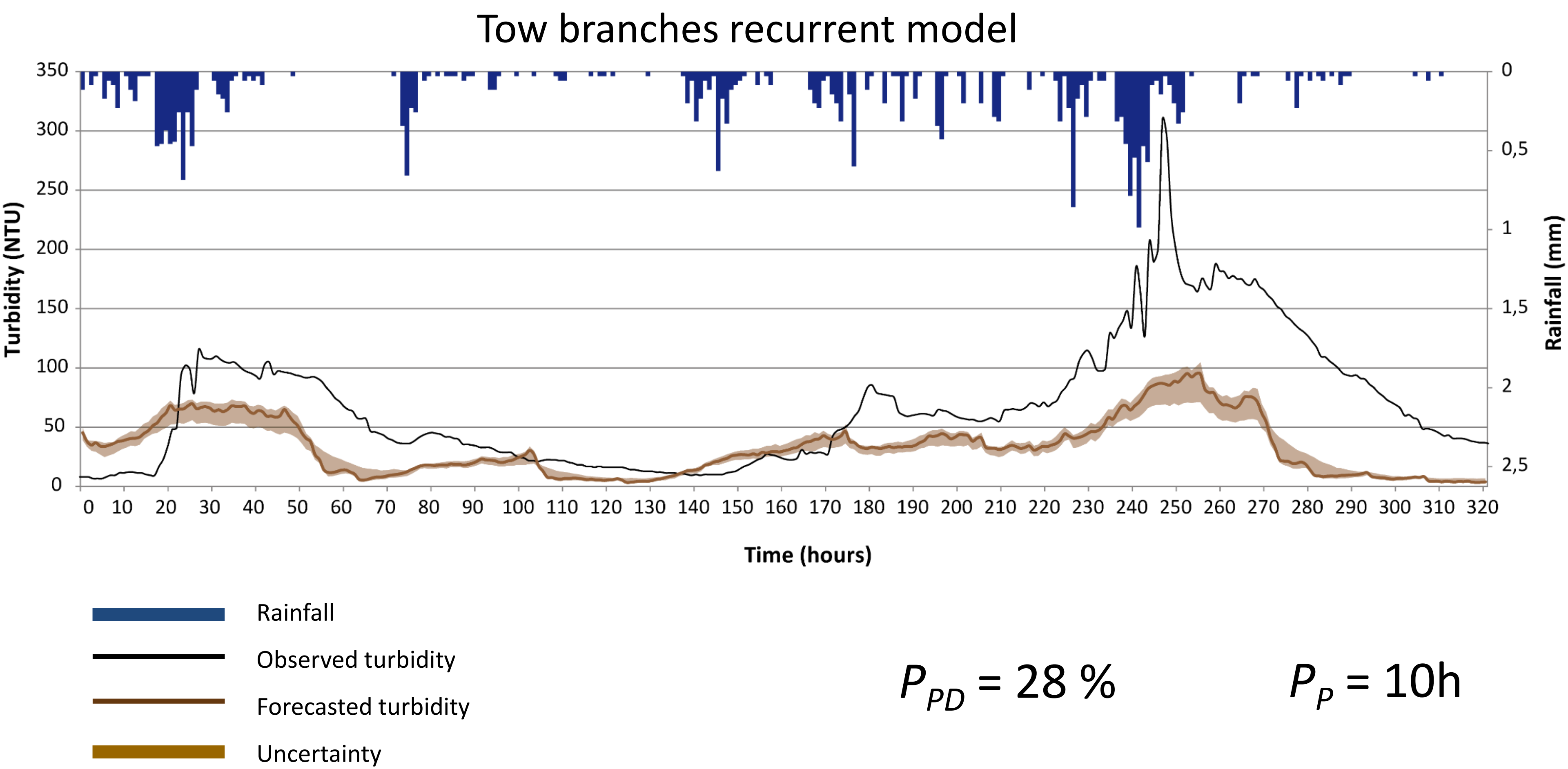
P_{PD} = 80 %

P_P = 15 h

Results

Illustration on test event n°11

(01/12/2011 to 30/12/2011)

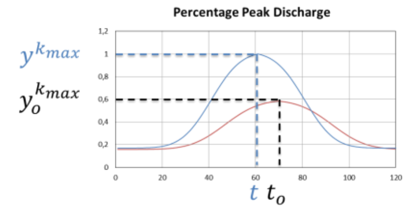


P_{PD} = 28 %

P_P = 10h

Results: Model with hourly data and mean rainfall

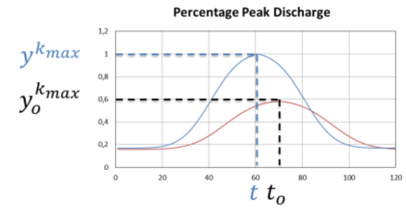
Cross test



	Two branches feed-forward		Two branches recurrent		MLP feed-forward		MLP recurrent	
	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)
T_2	79.95	4	38.01	18	80.62	19	39.58	21
T_3	112.95	4	60.99	15	108.67	4	70.63	3
T_6	217.41	8	79.97	10	93.30	14	69.70	10.50
T_7	Stop set		75.06	2	82.89	14	75.44	3
T_{10}	86.04	7	40.92	4	50.38	14	Stop set	
T_{11}	80.36	15	28.48	10	59.46	8	25.69	3
T_{16}	106.39	2	Stop set		65.52	14	44.16	10
T_{17}	105.55	5	87.55	101	Stop set		50.21	9
T_{18}	98.17	97	152.86	3	89.84	102	79.53	0
T_{22}	143.54	3	152.33	14	169.18	4	149.17	6
Median	105	5	75	10	83	14	70	6

Results: Model with hourly data and mean rainfall

Cross test

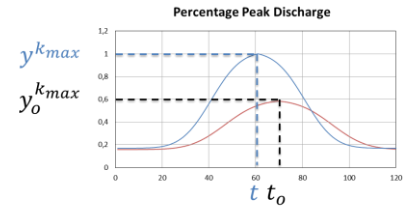


	Two branches feed-forward		Two branches recurrent		MLP feed-forward		MLP recurrent	
	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)
T_2	79.95	4	38.01	18	80.62	19	39.58	21
T_3	112.95	4	60.99	15	108.67	4	70.63	3
T_6	217.41	8	79.97	10	93.30	14	69.70	10.50
T_7	Stop set		75.06	2	82.89	14	75.44	3
T_{10}	86.04	7	40.92	4	50.38	14.50	Stop set	
T_{11}	80.36	15	28.48	10	59.46	8.50	25.69	3
T_{16}	106.39	2	Stop set		65.52	14	44.16	10
T_{17}	105.55	5	87.55	101	Stop set		50.21	9
T_{18}	98.17	97	152.86	3	89.84	102	79.53	0
T_{22}	143.54	3	152.33	14	169.18	4	149.17	6
Median	105	5	75	10	83	14	70	6

ECHN Early Career Hydrogeologists' Network
25-29th September 2016
Montpellier, France
CORUM CONFERENCE CENTER
43rd IAH congress

Results: Model with hourly data and spatialized rainfalls

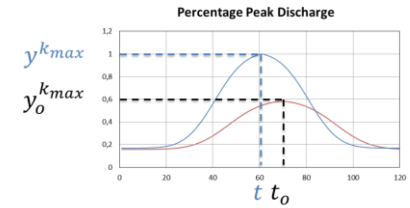
Cross test



	Two branches feed-forward		Two branches recurrent		MLP feed-forward		MLP recurrent	
	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)
T_2	128,2	17	36,4	27	89,6	16	32,2	20
T_3	101,6	4	80,3	84	105,0	8	77,2	3
T_6	173,5	11	56,3	13	117,0	11	46,4	12
T_7	Stop set		61,7	8	80.5	14	81,5	4
T_{10}	85,0	9	63,7	3	72,6	14	Stop set	
T_{11}	65,7	11	31,0	11	67,6	14	28,9	3
T_{16}	219,5	9	Stop set		61,5	15	40,1	12
T_{17}	94,5	7	61,1	8	Stop set		49,3	11
T_{18}	135,0	103	93,1	13	115,2	14	83,0	69
T_{22}	229,8	10	181,5	8	128,5	10	172,3	7
Median	137,0	20	73,9	19	94,6	13	67,9	15

Results: Model with hourly data and spatialized rainfalls

Cross test

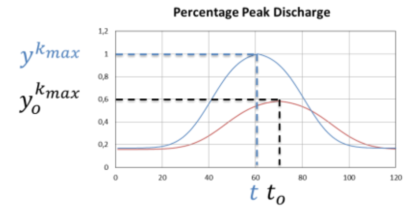


	Two branches feed-forward		Two branches recurrent		MLP feed-forward		MLP recurrent	
	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)	P_{PD}	Peak delay (h)
T_2	128,2	17	36,4	27	**89,6**	**16**	32,2	20
T_3	**101,6**	4	80,3	84	105,0	8	77,2	**3**
T_6	173,5	11	56,3	13	**117,0**	**11**	46,4	12
T_7	Stop set		61,7	8	80.5	14	**81,5**	**4**
T_{10}	**85,0**	9	63,7	**3**	72,6	14	Stop set	
T_{11}	65,7	11	31,0	11	**67,6**	14	28,9	**3**
T_{16}	**219,5**	**9**	Stop set		61,5	15	40,1	12
T_{17}	**94,5**	**7**	61,1	8	Stop set		49,3	11
T_{18}	135,0	103	**93,1**	**13**	115,2	14	83,0	69
T_{22}	229,8	10	181,5	8	**128,5**	10	172,3	**7**
Median	137,0	20	73,9	19	**94,6**	**13**	67,9	15

ECHN Early Career Hydrogeologists' Network
25-29th September 2016
Montpellier, France
CORUM CONFERENCE CENTER
43rd IAH congress

Results: Model with six hours sampling data and mean rainfall

Cross test

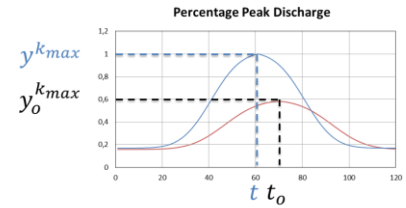


	Two branches feed-forward		Two branches recurrent		MLP feed-forward		MLP recurrent	
	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)
T_2	147,6	4	39,0	3	67,9	3	41,2	4
T_3	110,9	2	68,3	3	96,3	3	78,2	2
T_6	242,5	3	81,2	4	117,3	3	71,0	2
T_7	Stop set		75,0	2	81	4	66,6	1
T_{10}	75,1	3	48,4	2	60,1	4	Stop set	
T_{11}	62,8	3	28,3	2	65,6	4	34,3	2
T_{16}	132,0	3	Stop set		62,5	4	35,1	4
T_{17}	121,6	2	53,9	2	Stop set		43,8	2
T_{18}	92,4	2	75,2	5	91,3	4	82,5	7
T_{22}	187,3	3	161,8	2	148,6	2	166,8	3
Median	130,3	3	70,1	3	87,8	3	68,8	3

Early Career Hydrogeologists' Network
25-29th September 2016
Montpellier, France
43rd IAH congress

Results: Model with six hours sampling data and mean rainfall

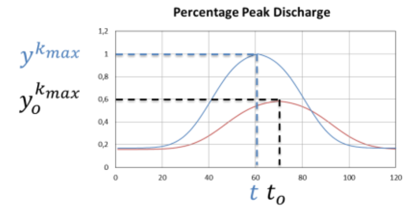
Cross test



	Two branches feed-forward		Two branches recurrent		MLP feed-forward		MLP recurrent	
	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)
T_2	147,6	4	39,0	3	67,9	3	41,2	4
T_3	110,9	2	68,3	3	96,3	3	78,2	2
T_6	242,5	3	81,2	4	117,3	3	71,0	2
T_7	Stop set		75,0	2	81	4	66,6	1
T_{10}	75,1	3	48,4	2	60,1	4	Stop set	
T_{11}	62,8	3	28,3	2	65,6	4	34,3	2
T_{16}	132,0	3	Stop set		62,5	4	35,1	4
T_{17}	121,6	2	53,9	2	Stop set		43,8	2
T_{18}	92,4	2	75,2	5	91,3	4	82,5	7
T_{22}	187,3	3	161,8	2	148,6	2	166,8	3
Median	130,3	3	70,1	3	87,8	3	68,8	3

Results: Model with six hours sampling data and spatialized rainfall

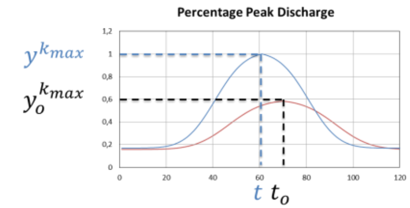
Cross test



	Two branches feed-forward		Two branches recurrent		MLP feed-forward		MLP recurrent	
	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)
T_2	186,4	6	38,2	6	67,0	9	41,1	3
T_3	114,9	2	41,4	3	93,9	3	63,5	1
T_6	196,9	4	60,9	3	85,2	3	59,7	3
T_7	Stop set		73,4	2	63	5	52,1	6
T_{10}	116,2	3	56,5	2	65,4	4	Stop set	
T_{11}	66,3	3	33,7	3	67,9	4	30,8	3
T_{16}	250,1	2	Stop set		59,8	4	40,4	3
T_{17}	120,1	2	56,3	3	Stop set		48,9	1
T_{18}	89,2	5	113,0	4	105,3	17	93,5	5
T_{22}	292,7	3	192,8	2	89,3	3	115,1	2
Median	159,2	3	74,0	3	77,4	6	60,6	3

Results: Model with six hours sampling data and spatialized rainfall

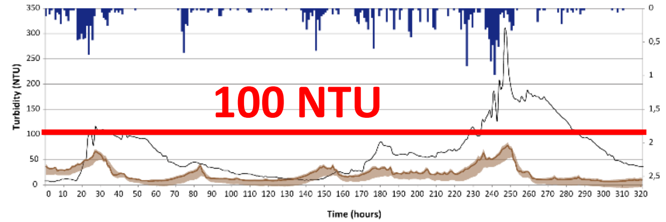
Cross test



	Two branches feed-forward		Two branches recurrent		MLP feed-forward		MLP recurrent	
	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)	P_{PD}	Peak delay (6h)
T_2	186,4	6	38,2	6	67,0	9	41,1	3
T_3	114,9	2	41,4	3	93,9	3	63,5	1
T_6	196,9	4	60,9	3	85,2	3	59,7	3
T_7	Stop set		73,4	2	63	5	52,1	6
T_{10}	116,2	3	56,5	2	65,4	4	Stop set	
T_{11}	66,3	3	33,7	3	67,9	4	30,8	3
T_{16}	250,1	2	Stop set		59,8	4	40,4	3
T_{17}	120,1	2	56,3	3	Stop set		48,9	1
T_{18}	89,2	5	113,0	4	105,3	17	93,5	5
T_{22}	292,7	3	192,8	2	89,3	3	115,1	2
Median	159,2	3	74,0	3	77,4	6	60,6	3

Results: Model with hourly data and mean rainfall

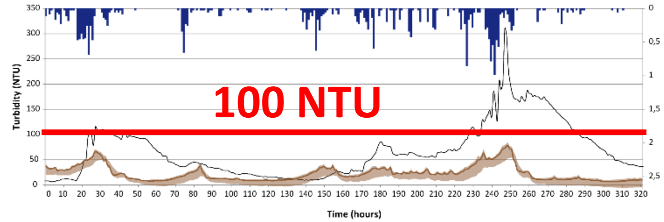
Cross test (100 NTU threshold overtaking)



	Two branches feed-forward				Two branches recurrent				MLP feed-forward				MLP recurrent			
	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)
T_2	0	0	6	11	0	0	6	-76	0	0	6	8	0	0	6	-53
T_3	1	0	0	3	1	0	-11	-4	1	0	7	4	0	0	0	-22
T_6	0	0	-4	10	0	0	-5	6	0	0	-4	9	0	0	-4	4
T_7	Stop set				X	X	X	X	X	X	X	X	X	X	X	X
T_{10}	0	1	-2	10	0	0	-1	9	0	0	-9	11	Stop set			
T_{11}	0	1	3	6	0	1	-17	-31	0	0	-12 / 8	-3 / 4	0	2	X	X
T_{16}	0	0	14	18	Stop set				0	0	1	12	0	0	6	-19
T_{17}	1	1	X	X	0	1	X	X	Stop set				0	1	X	X
T_{18}	0	0	X	X	0	0	X	X	0	0	X	X	1	0	X	X
T_{22}	1	0	X	X	0	0	X	X	0	0	X	X	1	0	X	X
Median	0.2	0.3	1.5	10	0.1	0.3	-5	-4	0.2	0	-1.5	8	0.3	0.4	3	-20

Results: Model with hourly data and mean rainfall

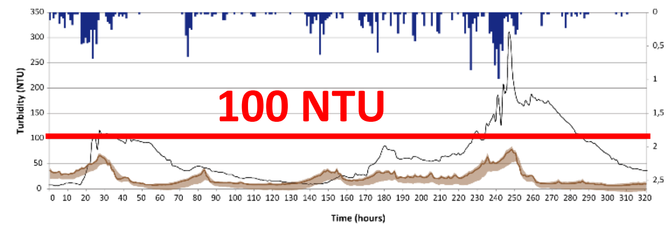
Cross test (100 NTU threshold overtaking)



	Two branches feed-forward				Two branches recurrent				MLP feed-forward				MLP recurrent			
	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)
T_2	0	0	6	11	0	0	6	-76	0	0	6	8	0	0	6	-53
T_3	1	0	0	3	1	0	-11	-4	1	0	7	4	0	0	0	-22
T_6	0	0	-4	10	0	0	-5	6	0	0	-4	9	0	0	-4	4
T_7	Stop set				X	X	X	X	X	X	X	X	X	X	X	X
T_{10}	0	1	-2	10	0	0	-1	9	0	0	-9	11	Stop set			
T_{11}	0	1	3	6	0	1	-17	-31	0	0	-12 / 8	-3 / 4	0	2	X	X
T_{16}	0	0	14	18	Stop set				0	0	1	12	0	0	6	-19
T_{17}	1	1	X	X	0	1	X	X	Stop set				0	1	X	X
T_{18}	0	0	X	X	0	0	X	X	0	0	X	X	1	0	X	X
T_{22}	1	0	X	X	0	0	X	X	0	0	X	X	1	0	X	X
Median	0.2	0.3	1.5	10	0.1	0.3	-5	-4	0.2	0	-1.5	8	0.3	0.4	3	-20

Results: Model with hourly data and spatialized rainfalls

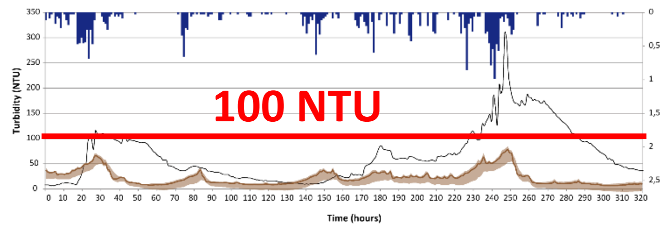
Cross test (100 NTU threshold overtaking)



	Two branches feed-forward				Two branches recurrent				MLP feed-forward				MLP recurrent			
	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)
T_2	0	0	-3	-8	0	0	-9	-32	0	0	-5	8	0	0	0	-88
T_3	0	0	6	-3	1	1	X	X	1	0	4	-24	1	0	4	-24
T_6	0	0	4	-10	1	0	4	-9	1	0	-1	-15	1	0	1	-15
T_7	Stop set				0	0	X	X	0	0	X	X	0	0	X	X
T_{10}	0	0	-4	-11	0	0	-3	17	0	0	3	17	Stop set			
T_{11}	0	0	7/9	-1/-1	0	1	X/-25	X/-20	0	2	X	X	0	2	X	X
T_{16}	0	0	-6	-13	Stop set				0	0	-6	-19	0	0	-6	-19
T_{17}	0	2	13	-13	0	2	10	-8	Stop set				0	3	X	X
T_{18}	1	0	X	X	0	0	X	X	0	0	X	X	0	0	X	X
T_{22}	2	0	X	X	1	0	X	X	1	0	X	X	1	0	X	X
Median	0	0	5	-9	0	0	-3	-9	0	0	-1	-15	0	1	1	-22

Results: Model with hourly data and spatialized rainfalls

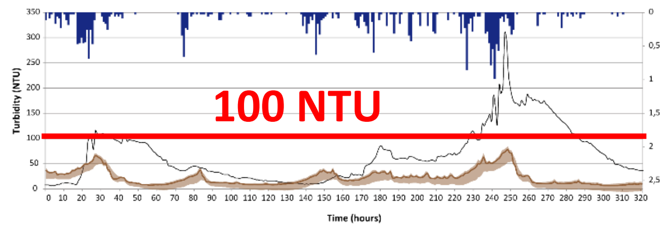
Cross test (100 NTU threshold overtaking)



	Two branches feed-forward				Two branches recurrent				MLP feed-forward				MLP recurrent			
	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)	F_p	F_n	d_r (h)	d_d (h)
T_2	0	0	-3	-8	0	0	-9	-32	0	0	-5	8	0	0	0	-88
T_3	0	0	6	-3	1	1	X	X	1	0	4	-24	1	0	4	-24
T_6	0	0	4	-10	1	0	4	-9	1	0	-1	-15	1	0	1	-15
T_7	Stop set				0	0	X	X	0	0	X	X	0	0	X	X
T_{10}	0	0	-4	-11	0	0	-3	17	0	0	3	17	Stop set			
T_{11}	0	0	7/9	-1/-1	0	1	X/-25	X/-20	0	2	X	X	0	2	X	X
T_{16}	0	0	-6	-13	Stop set				0	0	-6	-19	0	0	-6	-19
T_{17}	0	2	13	-13	0	2	10	-8	Stop set				0	3	X	X
T_{18}	1	0	X	X	0	0	X	X	0	0	X	X	0	0	X	X
T_{22}	2	0	X	X	1	0	X	X	1	0	X	X	1	0	X	X
Median	0	0	5	-9	0	0	-3	-9	0	0	-1	-15	0	1	1	-22

Results: Model with six hours sampling data and mean rainfall

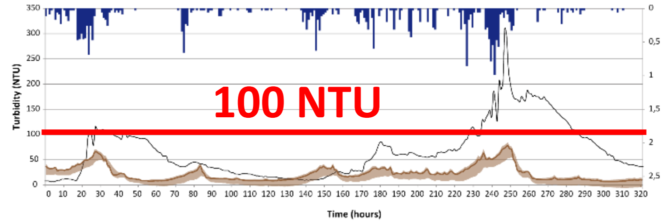
Cross test (100 NTU threshold overtaking)



	Two branches feed-forward				Two branches recurrent				MLP feed-forward				MLP recurrent			
	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)
T_2	0	0	0	3	0	0	0	-10	0	0	0	2	0	0	0	-6
T_3	1	0	0	0	1	0	1	-2	1	0	0	0	0	0	0	-4
T_6	0	0	1	1	0	0	2	1	0	0	2	1	0	0	1	-2
T_7	Stop set				0	0	X	X	0	0	X	X	0	0	X	X
T_{10}	0	0	3	3	0	0	1	2	0	0	1	3	Stop set			
T_{11}	0	0	2/1	1/1	0	2	X	X	0	0	3/3	0/1	0	1	X/3	X/-7
T_{16}	0	0	-1	3	Stop set				0	0	1	3	0	0	0	4
T_{17}	0	2	3	3	0	2	5	1	Stop set				0	3	X	X
T_{18}	0	0	X	X	0	0	X	X	0	0	X	X	0	0	X	X
T_{22}	1	0	X	X	0	0	X	X	0	0	X	X	0	0	X	X
Median	0	0	2	2	0	0	1	1	0	0	0	1	0	0	0	-4

Results: Model with six hours sampling data and mean rainfall

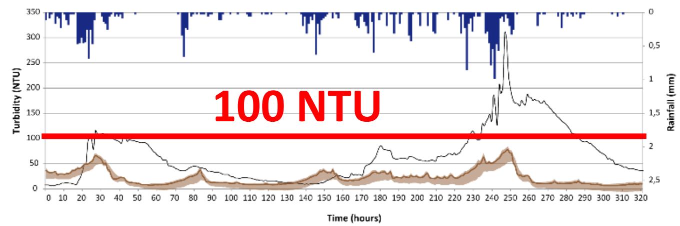
Cross test (100 NTU threshold overtaking)



	Two branches feed-forward				Two branches recurrent				MLP feed-forward				MLP recurrent			
	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)
T_2	0	0	0	3	0	0	0	-10	0	0	0	2	0	0	0	-6
T_3	1	0	0	0	1	0	1	-2	1	0	0	0	0	0	0	-4
T_6	0	0	1	1	0	0	2	1	0	0	2	1	0	0	1	-2
T_7	Stop set				0	0	X	X	0	0	X	X	0	0	X	X
T_{10}	0	0	3	3	0	0	1	2	0	0	1	3	Stop set			
T_{11}	0	0	2/1	1/1	0	2	X	X	0	0	3/3	0/1	0	1	X/3	X/-7
T_{16}	0	0	-1	3	Stop set				0	0	1	3	0	0	0	4
T_{17}	0	2	3	3	0	2	5	1	Stop set				0	3	X	X
T_{18}	0	0	X	X	0	0	X	X	0	0	X	X	0	0	X	X
T_{22}	1	0	X	X	0	0	X	X	0	0	X	X	0	0	X	X
Median	0	0	2	2	0	0	1	1	0	0	0	1	0	0	0	-4

Results: Model with six hours sampling data and spatialized rainfall

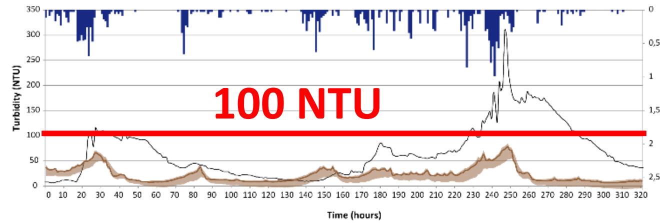
Cross test (100 NTU threshold overtaking)



	Two branches feed-forward				Two branches recurrent				MLP feed-forward				MLP recurrent			
	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)
T_2	0	0	-3	2	0	0	3	-3	0	0	2	2	0	0	0	-12
T_3	0	0	0	0	0	0	1	-2	0	0	1	0	0	1	X	X
T_6	1	0	-1	0	1	0	1	-1	0	0	3	-1	0	0	2	-2
T_7	Stop set				0	0	X	X	0	0	X	X	0	0	X	X
T_{10}	0	0	-1	3	0	0	0	4	0	0	3	3	Stop set			
T_{11}	1	0	-1/2	-1/1	0	0	0/3	-3/-2	0	0	-2/-3	-1/1	0	1	-1/X	-1/X
T_{16}	0	0	-1	2	Stop set				0	0	0	1	0	0	1	1
T_{17}	1	1	2	3	0	3	X	X	Stop set				1	3	X	X
T_{18}	0	0	X	X	2	0	X	X	0	0	X	X	0	0	X	X
T_{22}	2	0	X	X	2	0	X	X	0	0	X	X	0	0	X	X
Median	1	0	-1	1	1	0	1	-2	0	0	1	1	0	1	1	-2

Results: Model with six hours sampling data and spatialized rainfall

Cross test (100 NTU threshold overtaking)



	Two branches feed-forward				Two branches recurrent				MLP feed-forward				MLP recurrent			
	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)	F_p	F_n	d_r (6h)	d_d (6h)
T_2	0	0	-3	2	0	0	3	-3	0	0	2	2	0	0	0	-12
T_3	0	0	0	0	0	0	1	-2	0	0	1	0	0	1	X	X
T_6	1	0	-1	0	1	0	1	-1	0	0	3	-1	0	0	2	-2
T_7	Stop set				0	0	X	X	0	0	X	X	0	0	X	X
T_{10}	0	0	-1	3	0	0	0	4	0	0	3	3	Stop set			
T_{11}	1	0	-1/2	-1/1	0	0	0/3	-3/-2	0	0	-2/-3	-1/1	0	1	-1/X	-1/X
T_{16}	0	0	-1	2	Stop set				0	0	0	1	0	0	1	1
T_{17}	1	1	2	3	0	3	X	X	Stop set				1	3	X	X
T_{18}	0	0	X	X	2	0	X	X	0	0	X	X	0	0	X	X
T_{22}	2	0	X	X	2	0	X	X	0	0	X	X	0	0	X	X
Median	1	0	-1	1	1	0	1	-2	0	0	1	1	0	1	1	-2

Conclusion

Goal of the study:

- 12 hour lead time operational turbidity forecasting
  - Turbidity peak forecasting
  - 100 NTU threshold overtaking forecast

Method:

- Neural network (16 models)
  - Multi-layer Perceptron / Two Branches
  - 1 hour/ 6 hours sampling
  - Recurrent / Feed-forward
  - Mean/spatialized Rainfall

Conclusion

Results:

- Spatialized rainfalls increase the complexity but not the efficiency
- Generally feed forward models are better to forecast the **amplitude** of peaks when recurrent models are better to forecast the **synchronization** of peaks
- The threshold overtaking is predicted with very good efficiency

Future work

- 24 hours lead time forecast
- Multiresolution analysis
- Combine the multilayer perceptron and the two branches to forecast peak and threshold overtaking

References

- Artigue G., Johannet A., Borrell V., Pistre S. (2012) Flash flood forecasting in poorly gauged basins using neural networks: case study of the Gardon de Mialet basin (southern France). Nat. Hazards Earth Syst. Sci., 12, 3307-3324.
- Barron, A. R. (1993). Universal approximation bounds for superpositions of a sigmoidal function. Information Theory, IEEE Transactions on, 39(3), 930-945.
- Chédeville, S., Laignel, B., Rodet, J., Todisco, D., Fournier, M., Dupuis, E., Girot, G., Hanin, G. (2014a) The sedimentary filling in the chalk karst of the Northwestern Paris Basin (Normandy, France): Characterization, origin and hydro-sedimentary behaviour. Z. Für Geomorphol. 58(1)
- Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. Neural computation, 4(1), 1-58.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. Neural networks, 2(5), 359-366.
- Johannet, A., Vayssade, B., & Bertin, D. (2008). Neural networks: from black box towards transparent box. Application to evapotranspiration modeling. International Journal of Computational Intelligence, 4(3), 163-170.
- Morgan, N., & Bourlard, H. (1989). Generalization and parameter estimation in feedforward nets: Some experiments. International Computer Science Institute.
- Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P. Y., ... & Juditsky, A. (1995). Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, *31*(12), 1691-1724.
- Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society. Series B (Methodological), 111-147.

Thank you for your attention

Yport cliffs

ECHN Early Career Hydrogeologists' Network
iah • aih 60 years
25-29th September 2016
Montpellier, France
CORUM CONFERENCE CENTER
43rd IAH congress

Quality criteria

Nash criterion and Percentage Peak Discharge

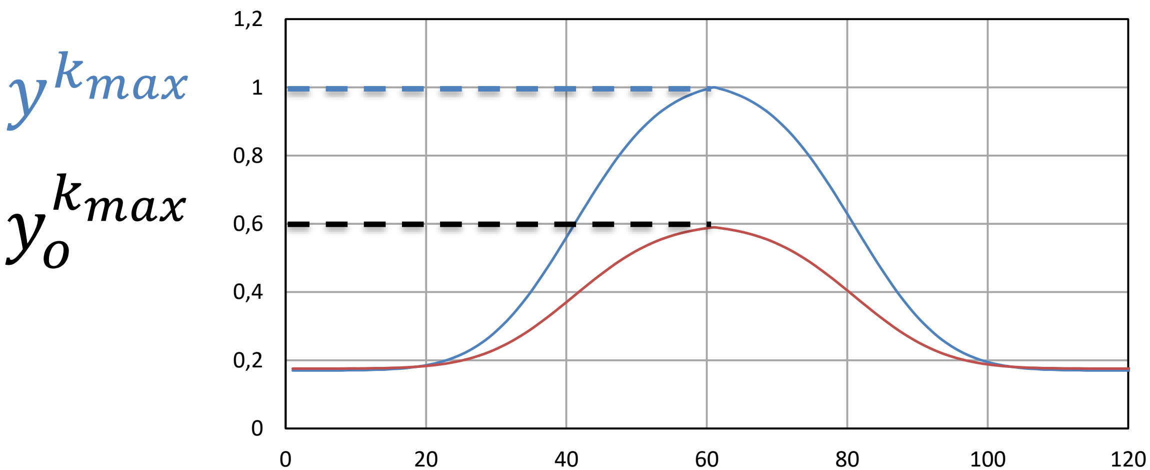
$$R^2 = 1 - \frac{\sum_{k=1}^{n} (y_o^k - y^k)^2}{\sum_{k=1}^{n} (y_o^k - \overline{y_o})^2}$$

Nash Criterion (Ref Nash)

$$P_{PD} = 100 \frac{y^{k_{max}}}{y_o^{k_{max}}}$$

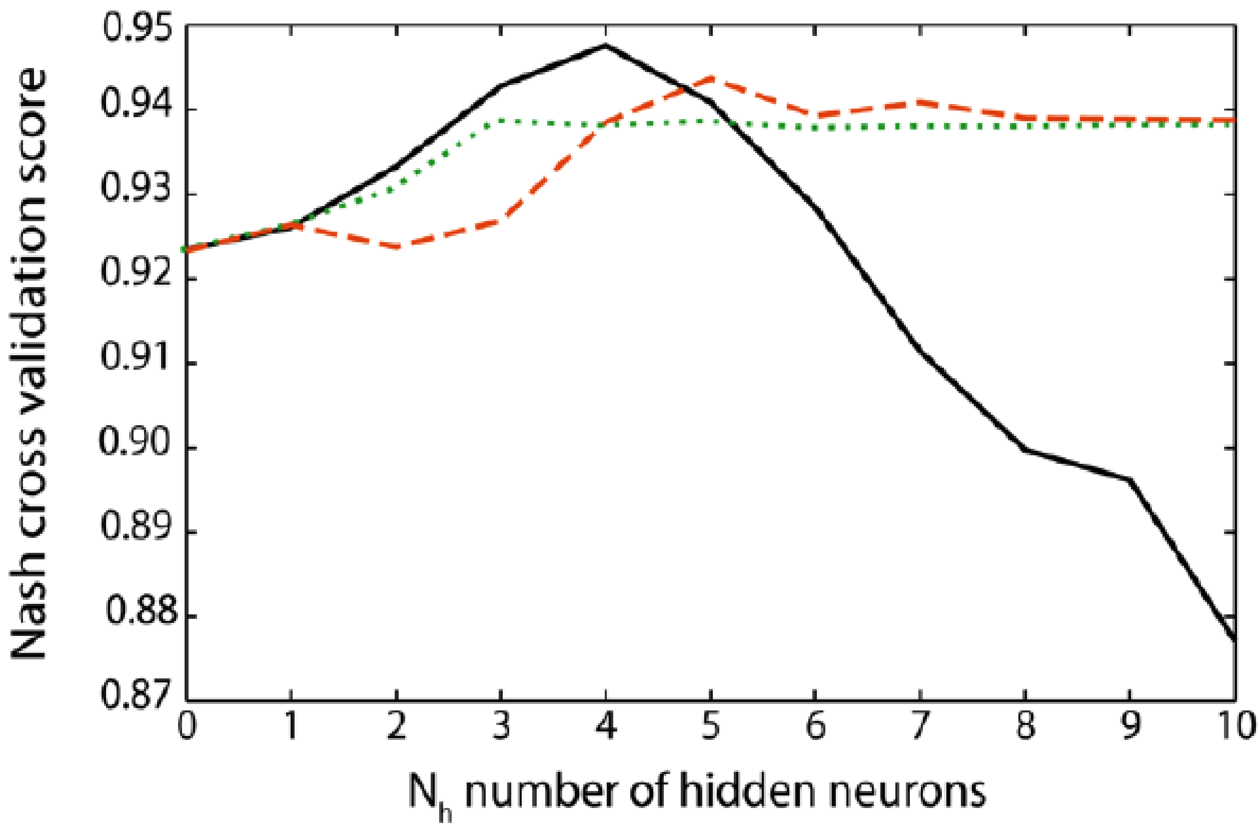
PPD (Percentage Peak Discharge) (Ref guillaume article nhess)

Percentage Peak Discharge



Bias variance dilemna

Nash criterion and Percentage Peak Discharge



Kong A. S., & *al (2011b)*

Score without regularization method

Score with weight decay

Score with early stopping

Correlations coefficient

	Froberville rainfall	Annouville rainfall	Goderville rainfall	Anglesqueville rainfall	Mannevillette rainfall	Etainhus rainfall
Yport turbidity	0,19	0,20	0,25	0,19	0,21	0,19
Froberville rainfall	1,00	0,50	0,49	0,48	0,38	0,42
Annouville rainfall		1,00	0,59	0,48	0,38	0,46
Goderville rainfall			1,00	0,60	0,42	0,52
Anglesqueville rainfall				1,00	0,48	0,53
Mannevillette rainfall					1,00	0,45
Etainhus rainfall						1,00