CSIC

Simulating lateral flow and its contribution to spatial variations of (rainfed) winter wheat yields

Research state, preliminary results and following steps

Tomás Roquette Tenreiro

Department of Agronomy (IAS-CSIC)

Promoters and supervisors: **José A. Gómez & Elías Fereres**

Conceptual background

Point-based model

(1st routine)

Mechanistic modelling of surface run-off and optimization of hydraulic parameters

+

30 year-sequence

Upscaling method

(LIF prediction)

MLR / ANN

(data-driven method)

training and testing sets

+

30 year-sequence

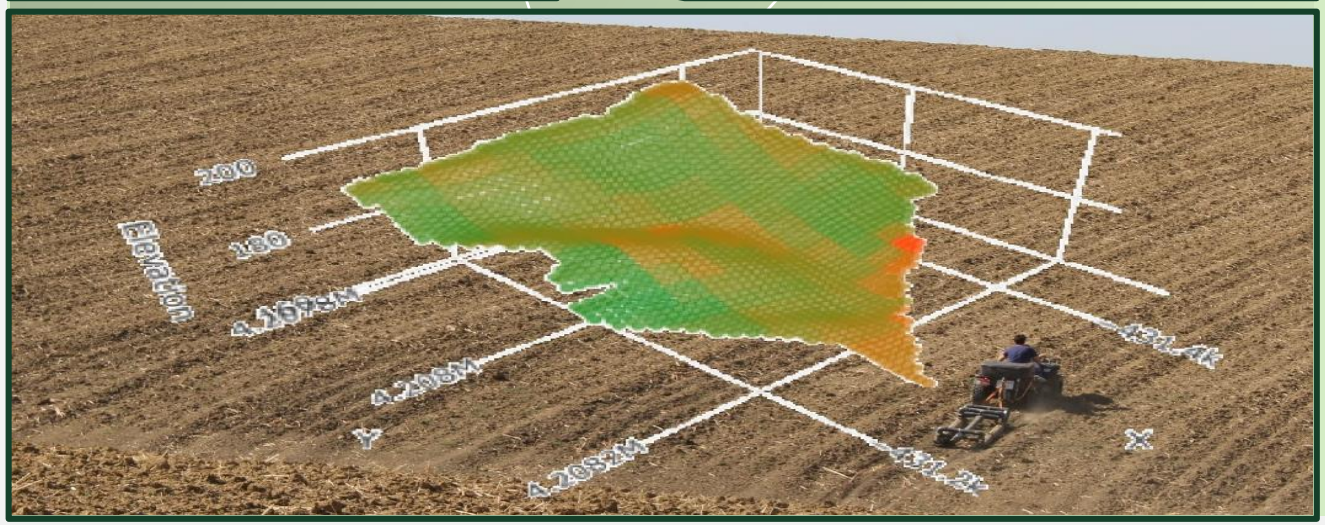
Crop modelling routine

(2nd routine)

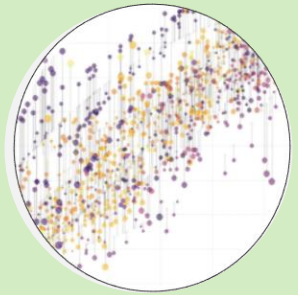
Yield simulations under the 1st vs. 2nd scenario

(with/without LIF)

Probability distribution functions of yield response to LIF

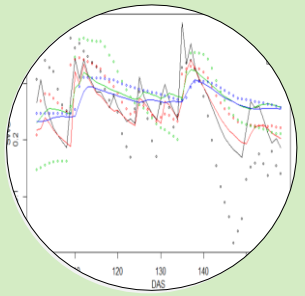


Preliminary results



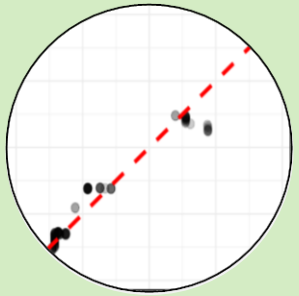
Meta-analysis

Assimilation of NDVI-CC models into HYDRUS-1D



Field datasets

Soil parameterization, root growth rates, phenological development, crop growth and crop management



Upscaling methods

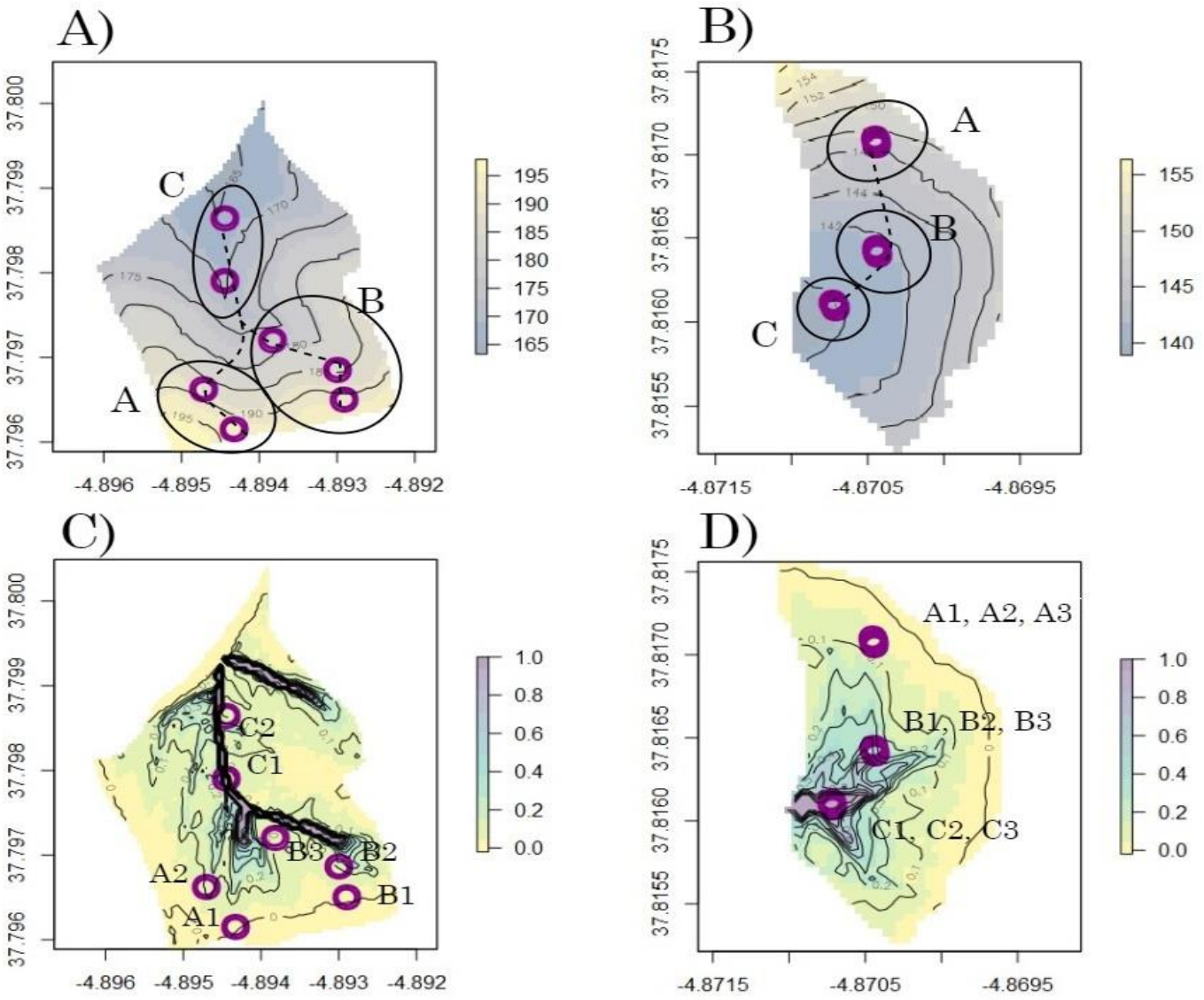
Multiple Linear Reg. models

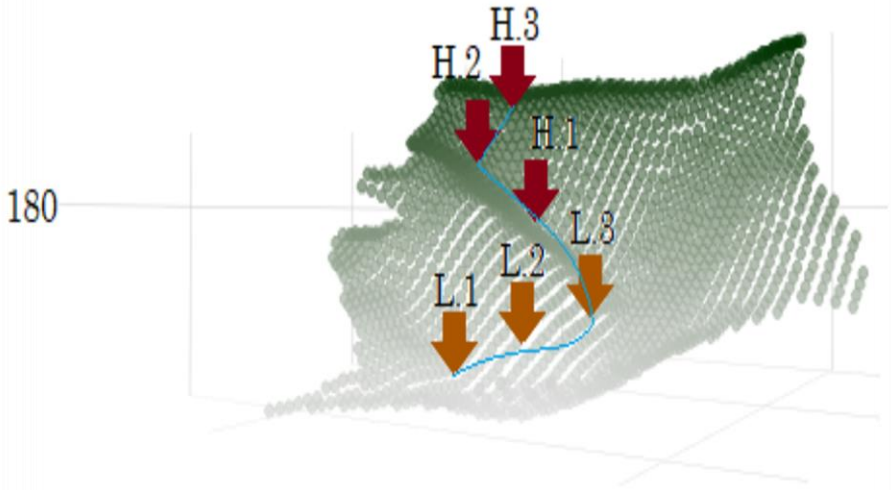
Artificial Neural Networks

"novel data assimilation algorithms"

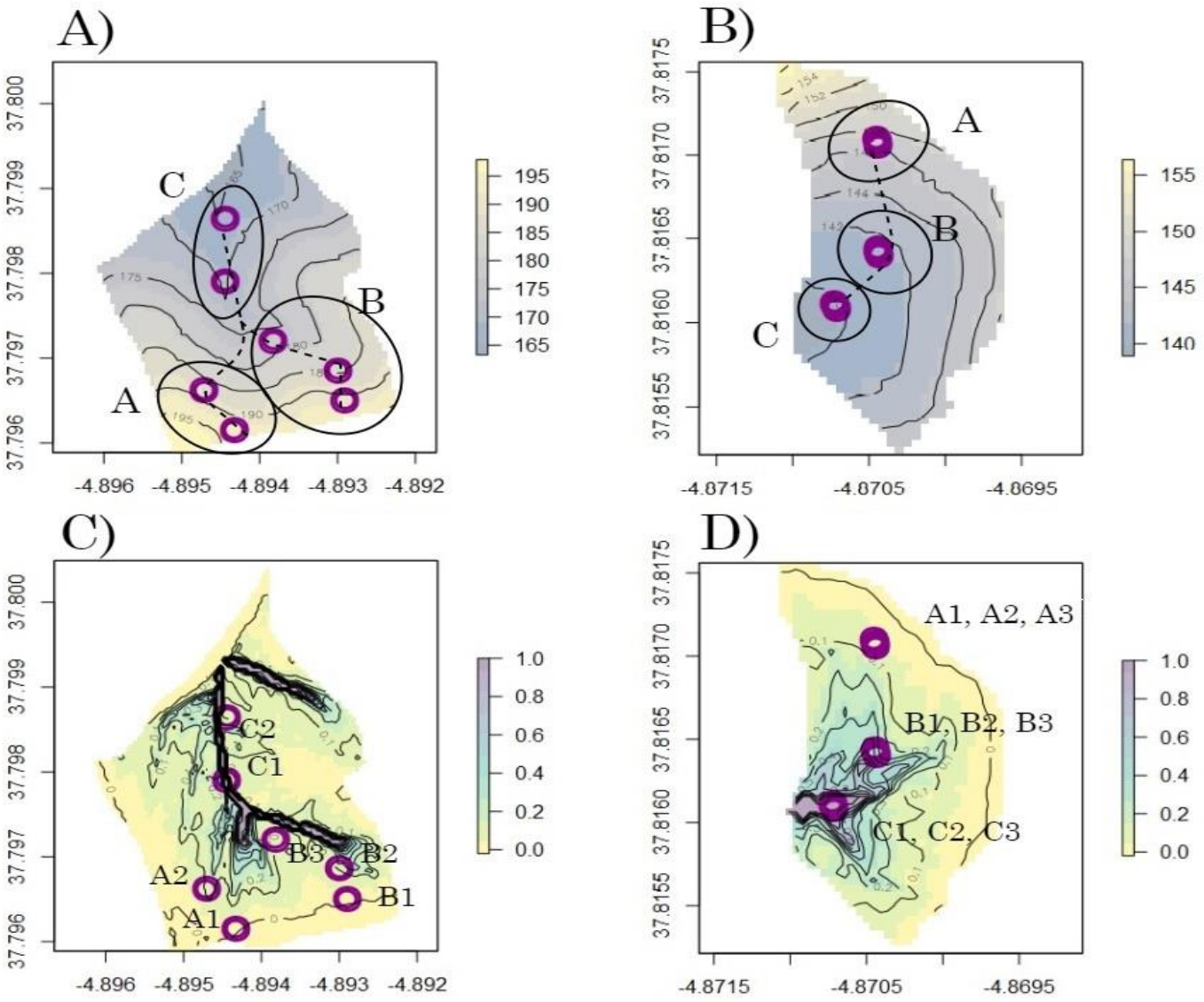
CSIC Consejo Superior de Investigaciones Científicas | Instituto de Agricultura Sostenible IAS

Experimental sites



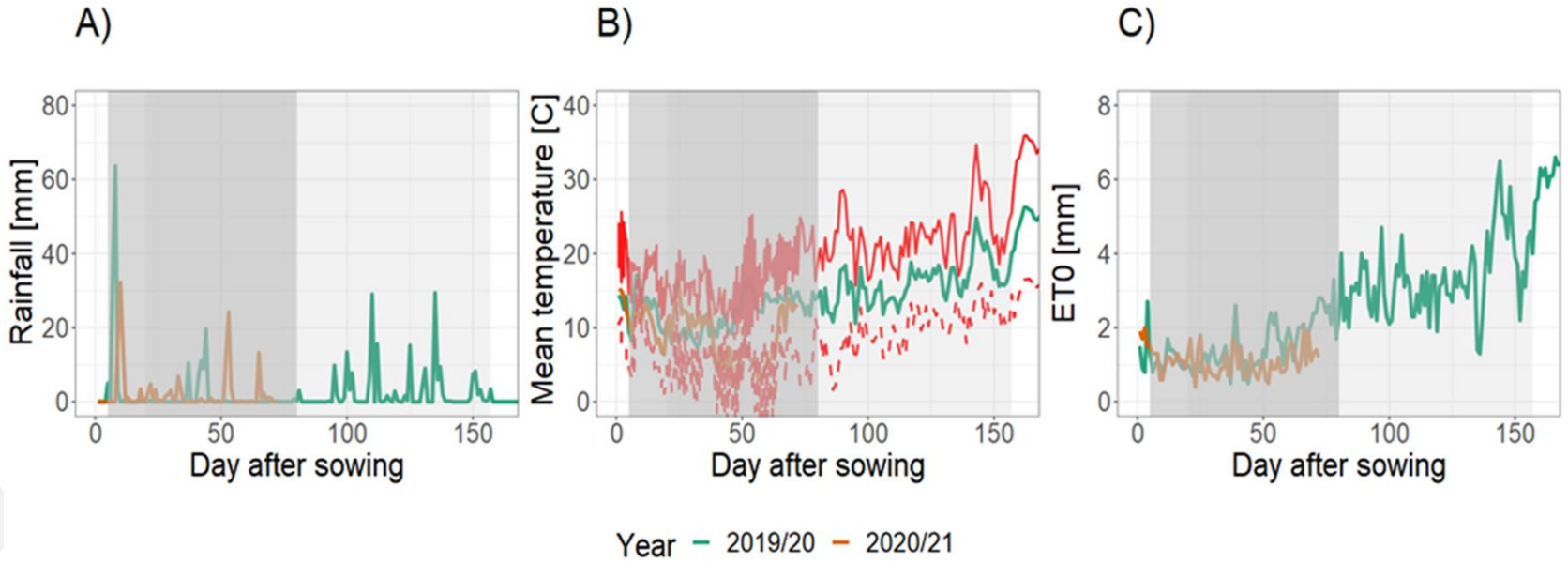


Experimental sites

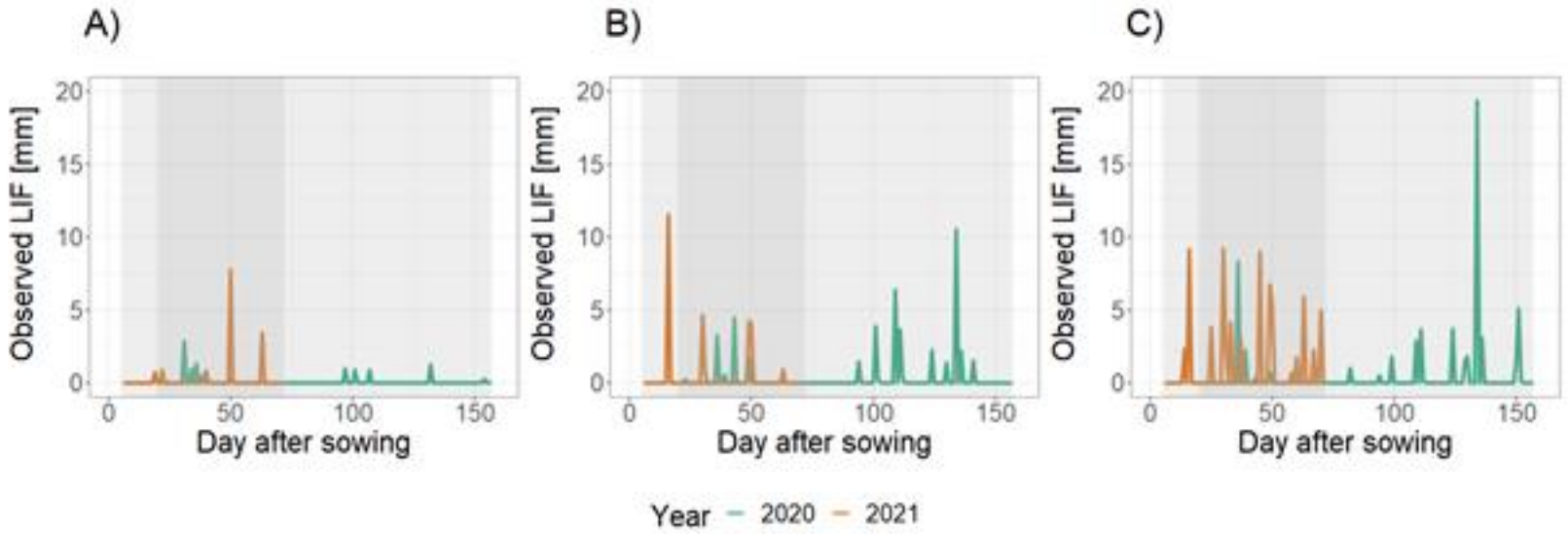


Parameter	Zone	Catchment-1 A	Catchment-1 B	Catchment-1 C	Catchment-2 A	Catchment-2 B	Catchment-2 C
	Sampling points [N]	2	3	2	3	3	3
	[Units]						
Elevation	m (amsl.)	187 (2.4)	184 (5.1)	168 (2.3)	147 (0.5)	141 (0.3)	139 (0.3)
ECa	dS/m	0.31 (0.08)	0.28 (0.09)	0.48 (0.06)	-	-	-
%Clay	%	45 (3.4)	42 (3.3)	50 (3.4)	44 (2.8)		
%Sand	%	18 (2.6)	22 (2.9)	15 (2.8)	22 (3.2)		
Bulk density	g/cm3	1.78 (0.06)	1.81 (0.04)	1.74 (0.05)	1.66 (0.05)		
FAI	[0;1]	0.09 (0.021)	0.12 (0.082)	0.72 (0.218)	0.07 (0.005)	0.42 (0.002)	0.98 (0.011)

Rainfall and meteorological data

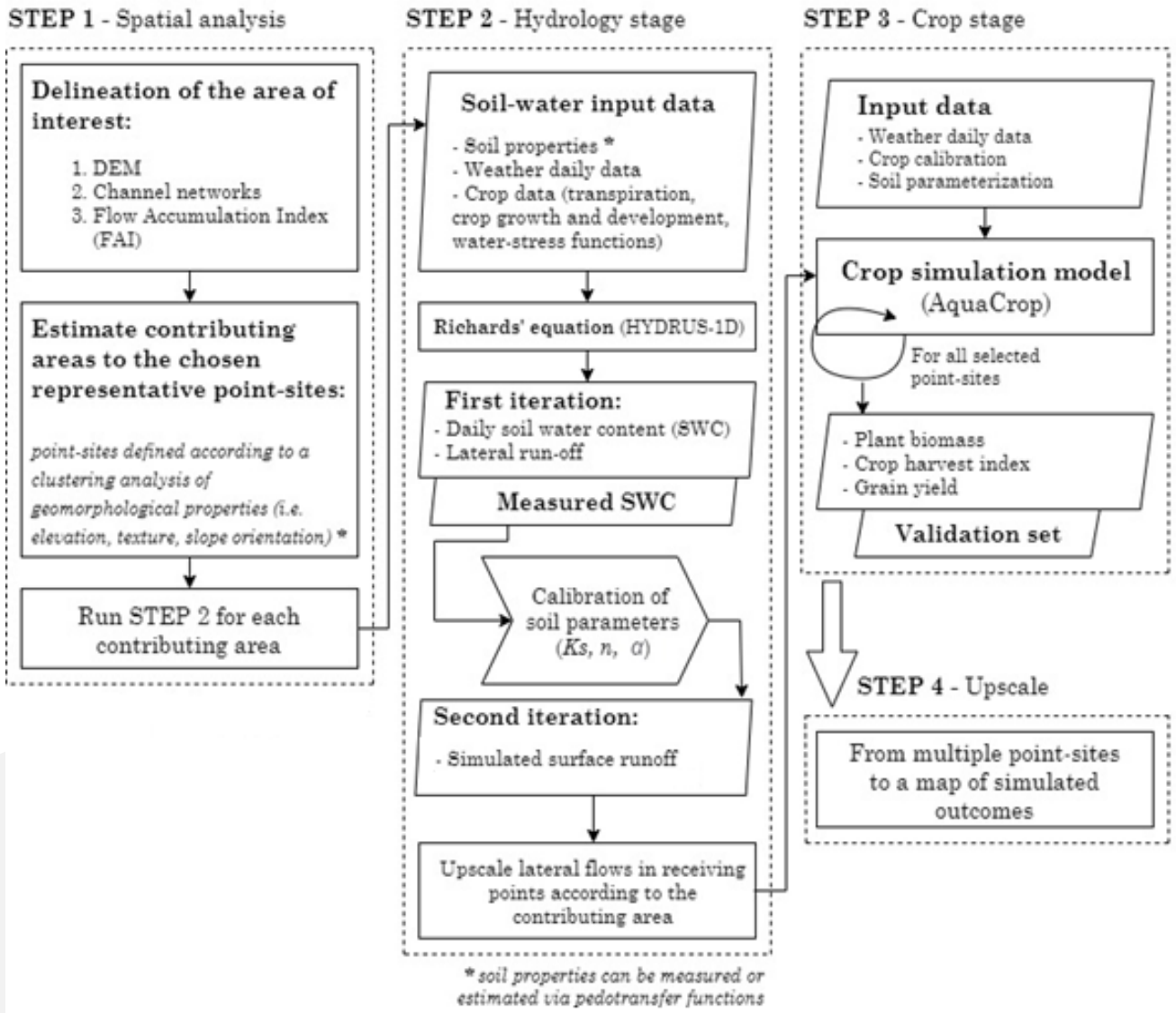


Lateral inflows
(inferred from field observations)

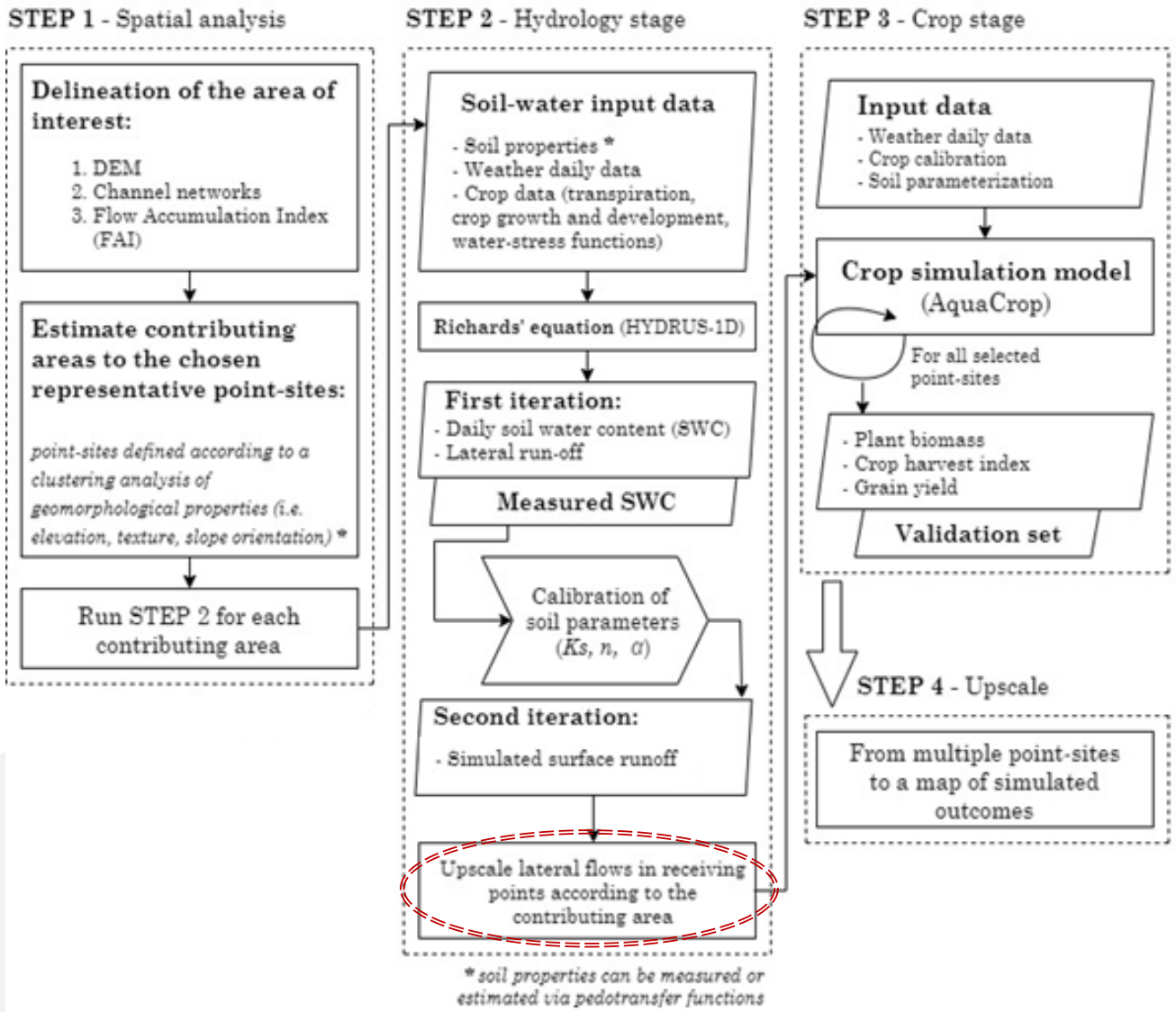


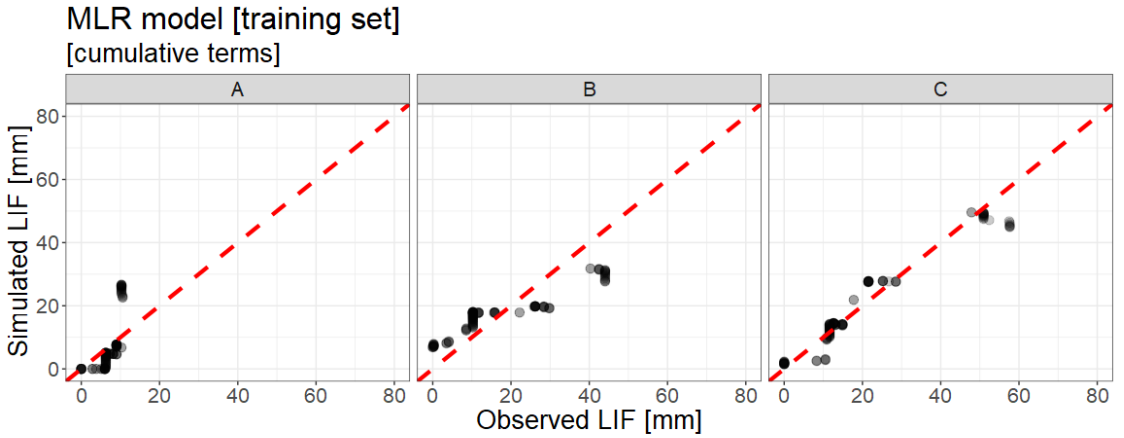
	Zone	Catchment-1 *2019/20*	Catchment-2 *2019/21*	Average *2019-21*
SWC [mm]	A	209.36 (39.6) c	327.46 (22.3) c	257.36 (67.1) c
	B	237.36 (34.4) b	352.06 (10.9) a	269.66 (60.7) b
	C	269.37 (63.1) a	343.41 (12.6) b	299.65 (67.4) a
LIF_n [mm]	A	0.08 (0.46) b	0.21 (1.22) b	0.13 (0.9) b
	B	0.31 (1.52) ab	0.41 (2.3) b	0.35 (1.8) b
	C	0.42 (2.09) a	1.11 (3.3) a	0.70 (2.7) a
$CUM.LIF_n$ [mm]	A	6.98 (3.9) b	4.43 (6.1) c	6.15 (5.5) c
	B	17.88 (15.7) a	16.75 (15.3) b	18.52 (15.8) b
	C	19.76 (16.9) a	34.45 (27.3) a	25.77 (22.9) a

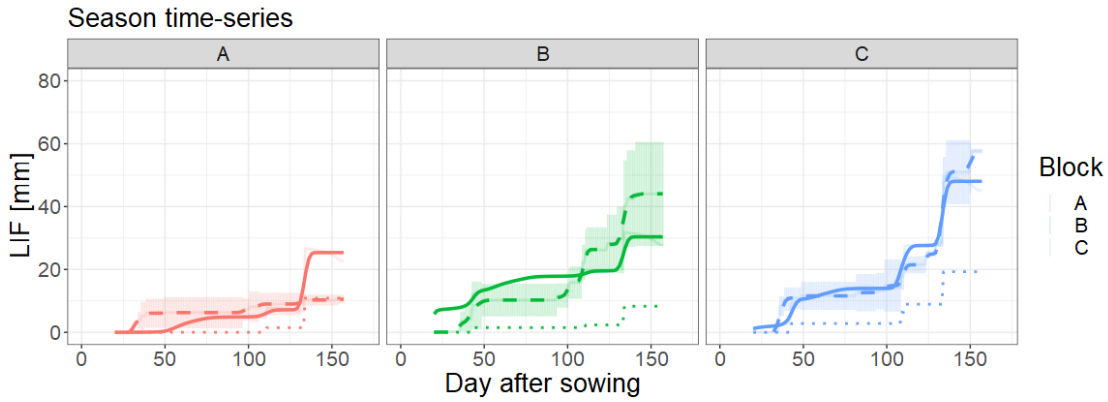
Methodological design



Upscaling methods – "MLR"



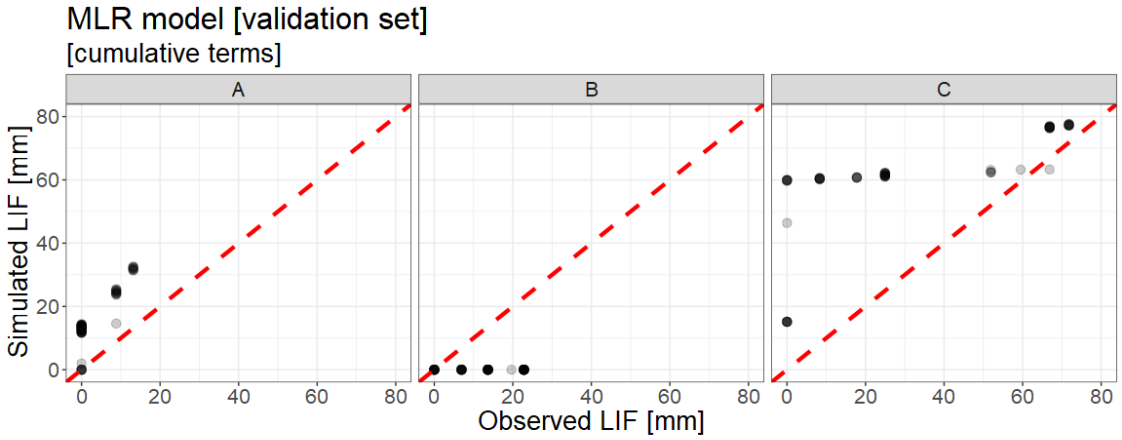


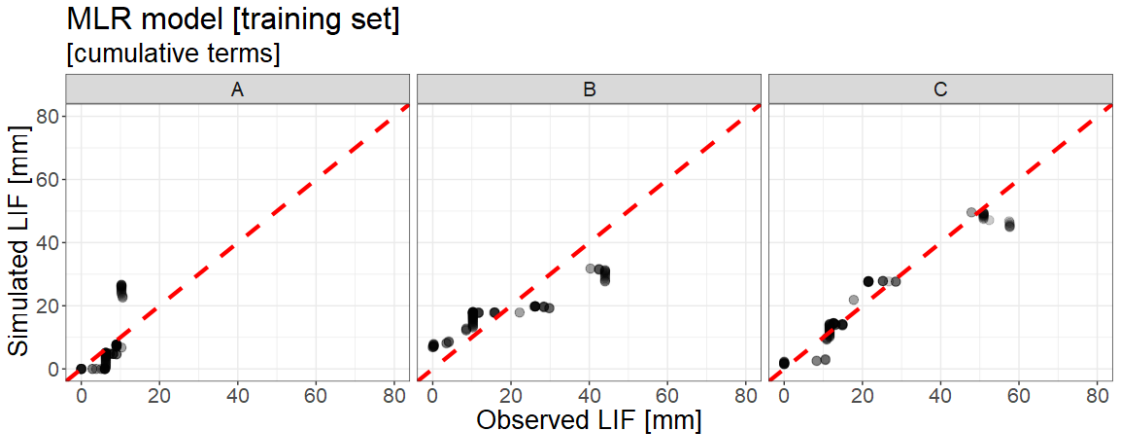


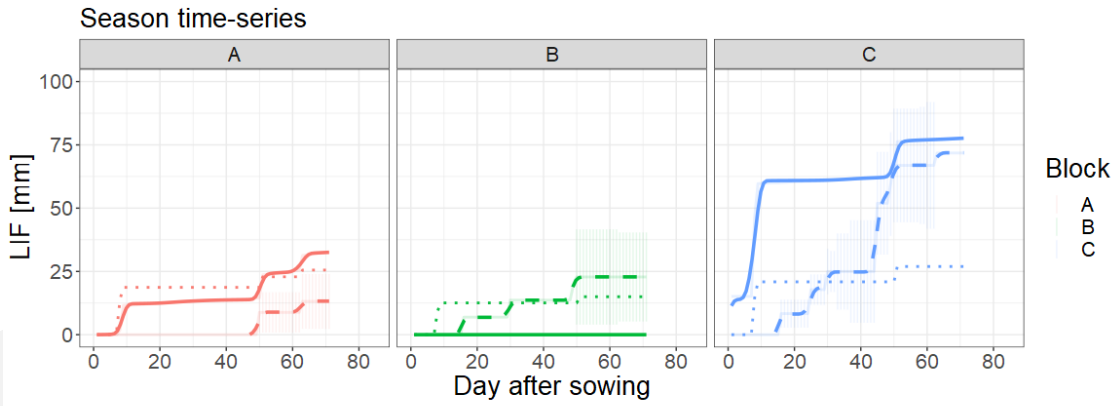
R-sq=0.78 (training)

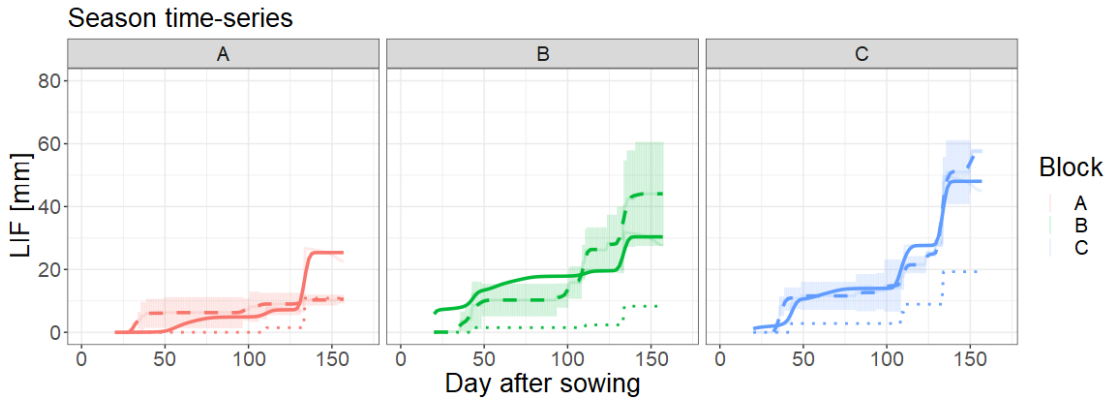
MLR model

(RMSE=21 mm; R-sq=0.78 (training); R-sq=0.48 (testing); Pearson=0.68; d-Willmott index=0.7)



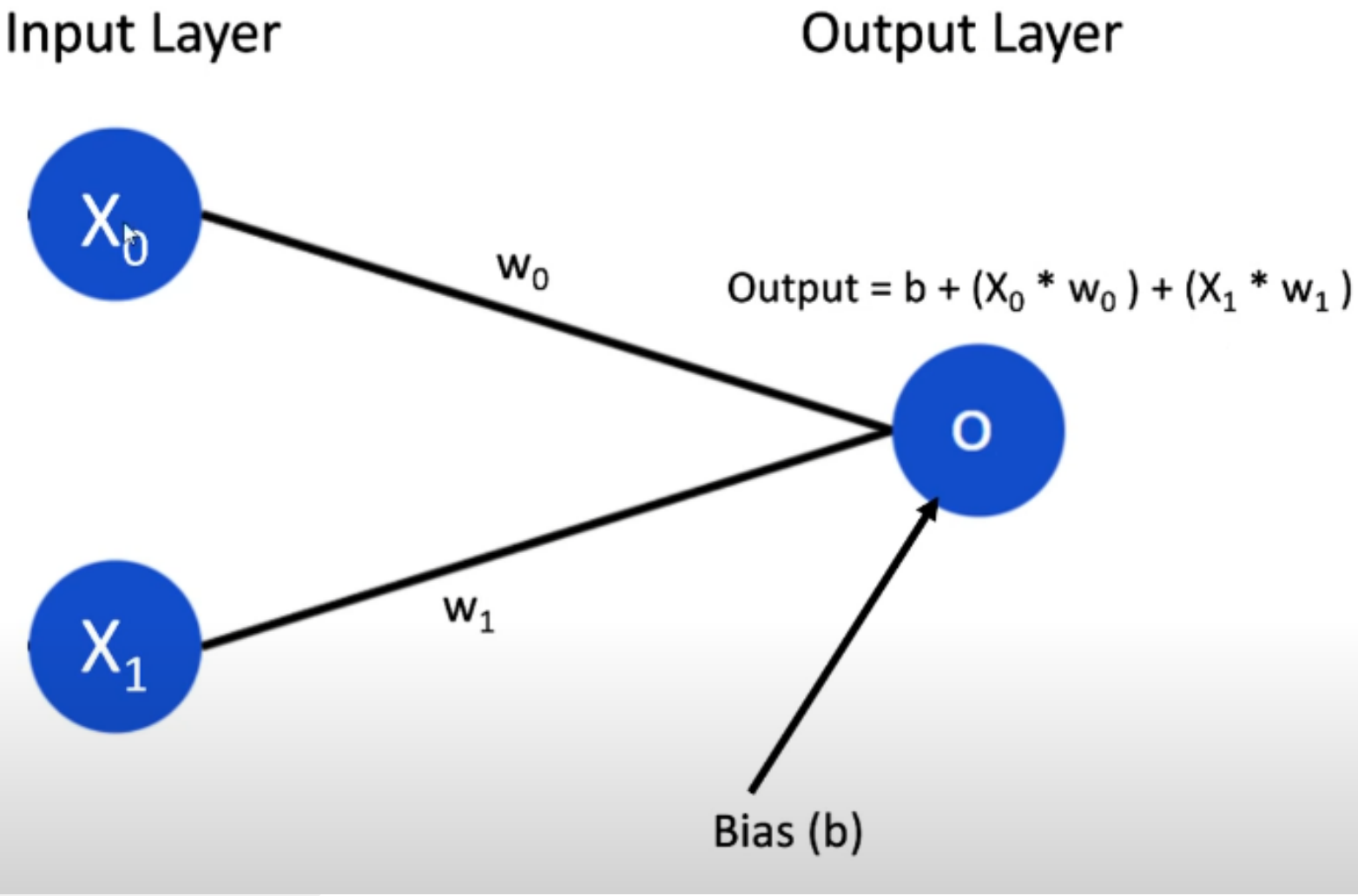






MLR -> ANN

Neural network with no hidden layers is just a linear model

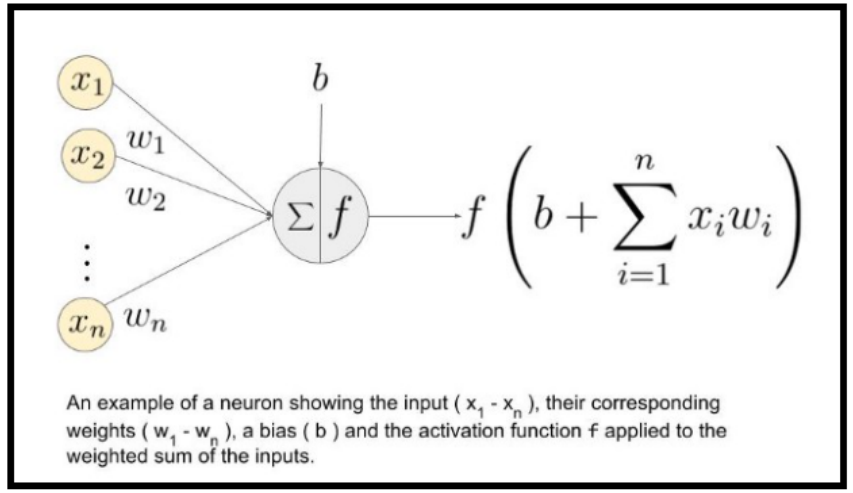


MLR -> ANN

MLR

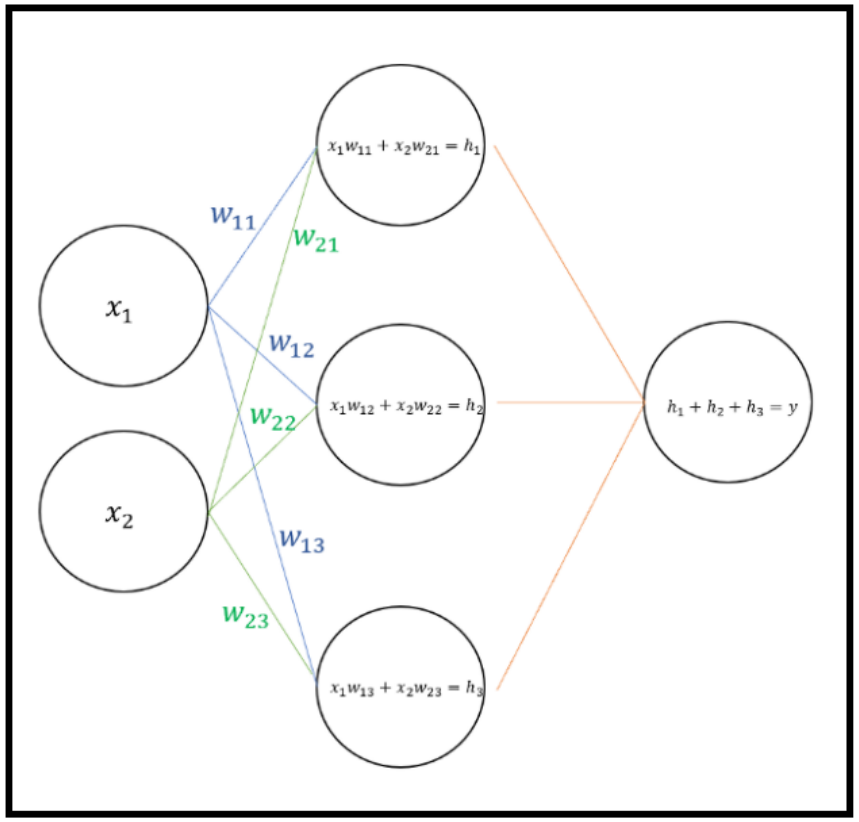
$$y_{pi} = X_i \beta_p + e_{p_i}.$$

The simplest ANN (single neuron)



An example of a neuron showing the input (x_1 - x_n), their corresponding weights (w_1 - w_n), a bias (b) and the activation function f applied to the weighted sum of the inputs.

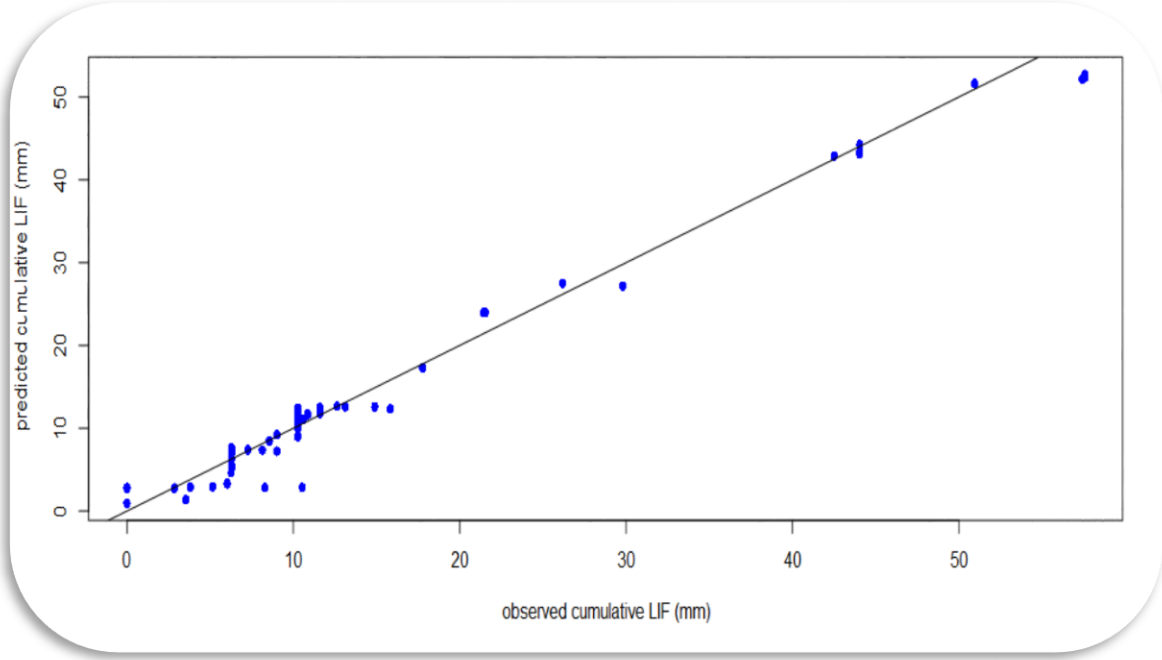
ANN w/3 neurons

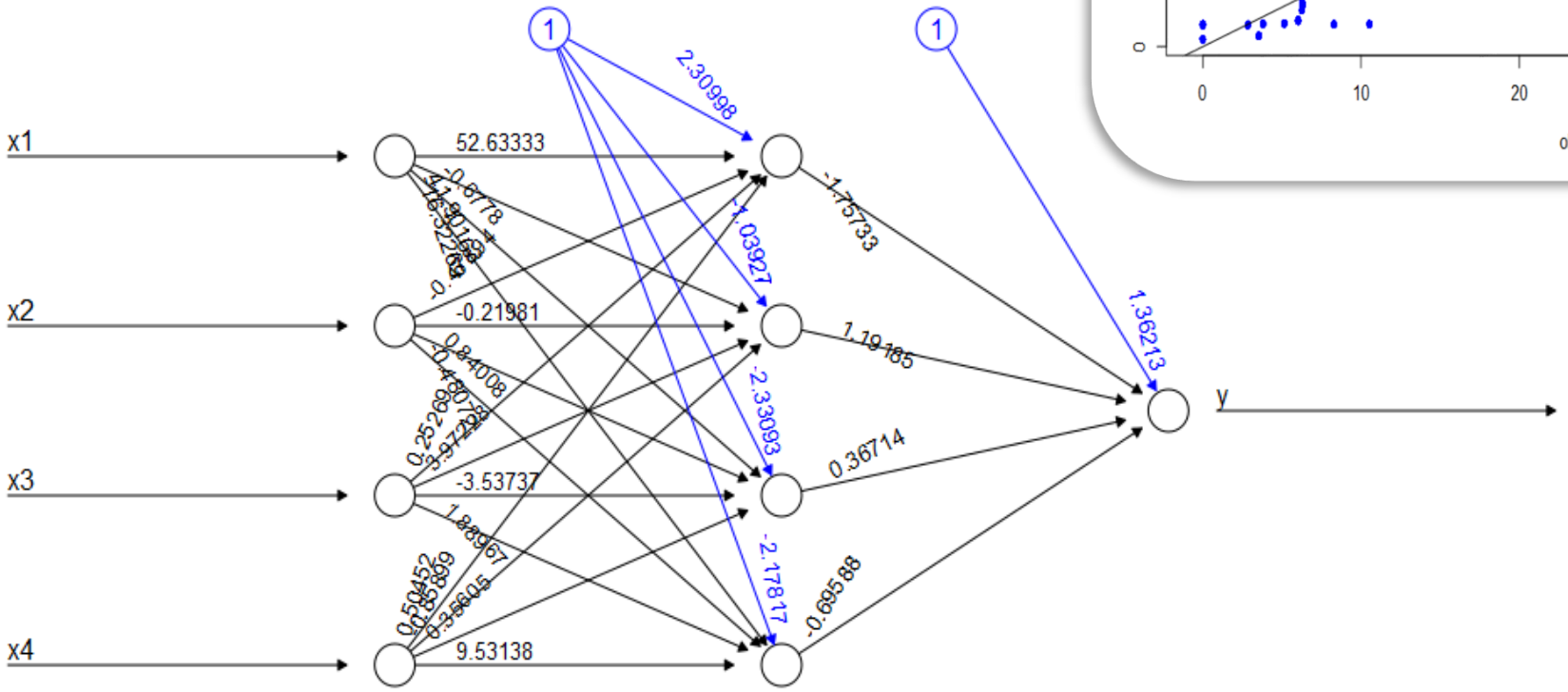


$$[x_1 \quad x_2] \begin{bmatrix} w_{11} & w_{12} & w_{13} \\ w_{21} & w_{22} & w_{23} \end{bmatrix} = \begin{bmatrix} x_1 w_{11} + x_2 w_{21} \\ x_1 w_{12} + x_2 w_{22} \\ x_1 w_{13} + x_2 w_{23} \end{bmatrix}' = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}'$$

ANN – random initialization

(RMSE=3 mm; R-sq=0.93)

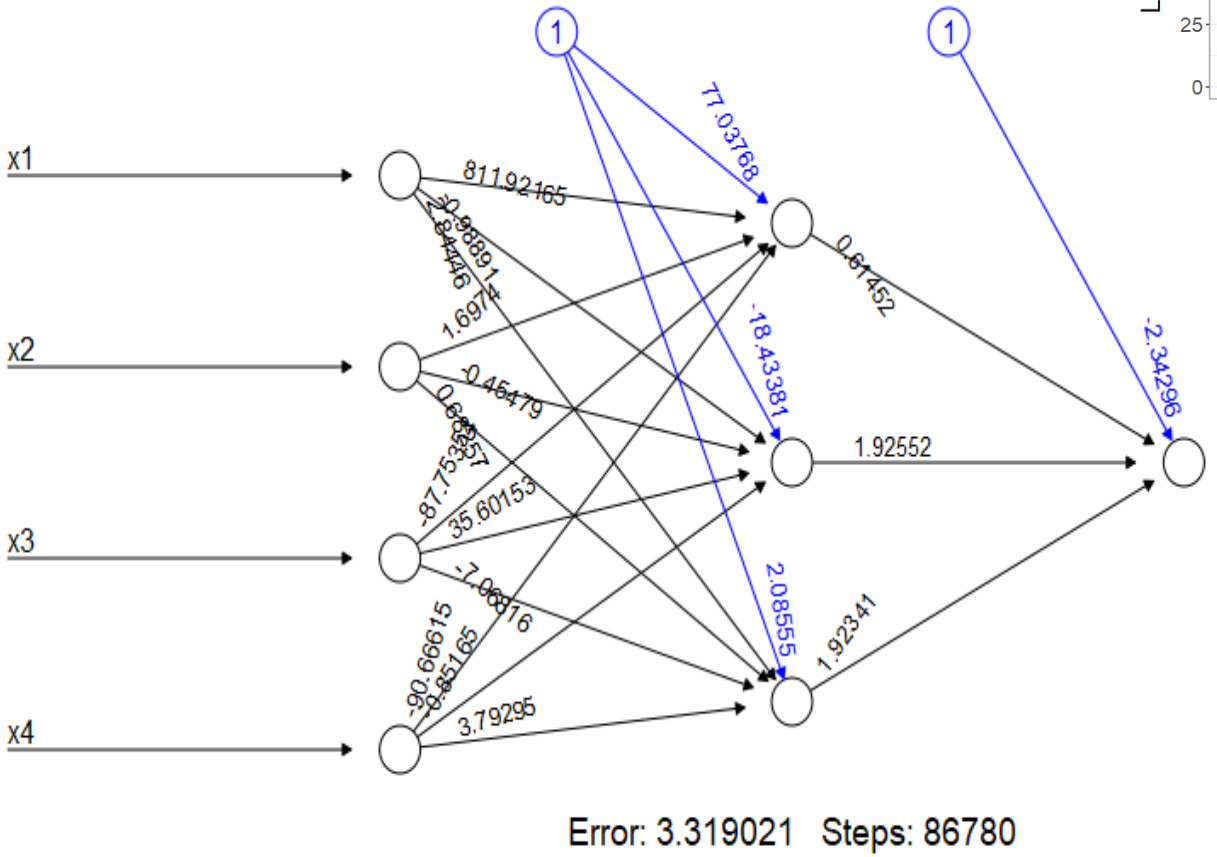


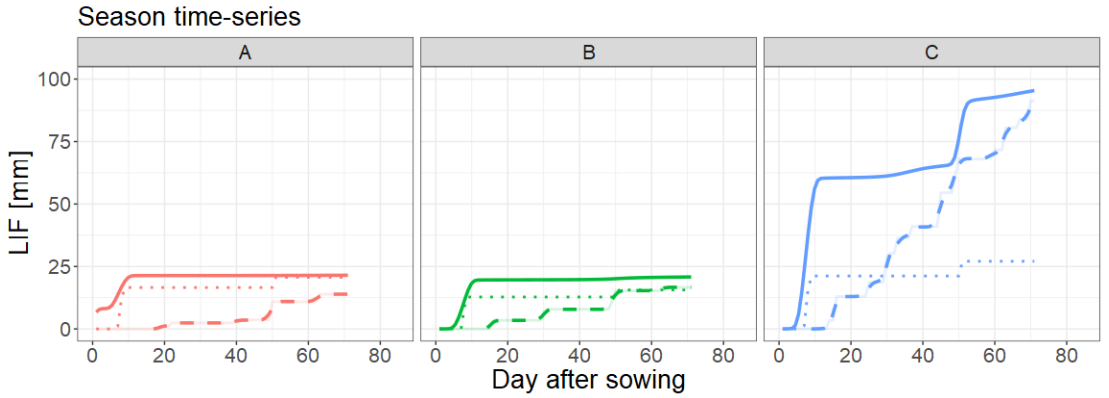


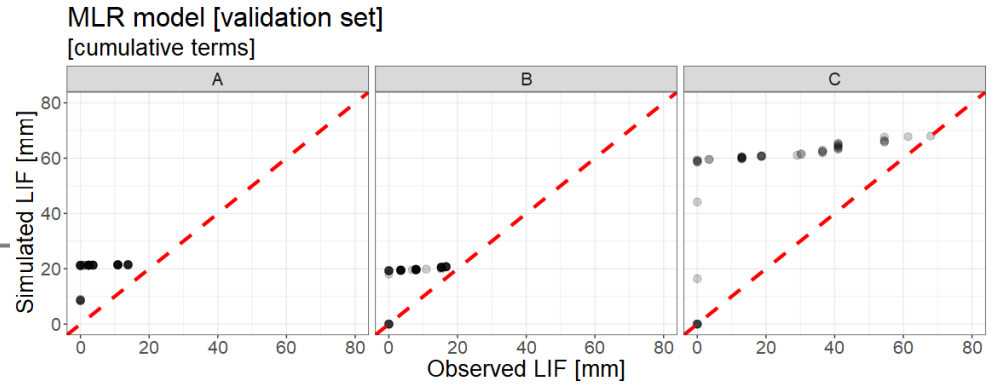
Error: 0.195424 Steps: 1369

ANN – 1st/2nd year sets [train/test]

(RMSE=21 mm; R-sq=0.76; Pearson=0.86; d-Willmott index=0.83)

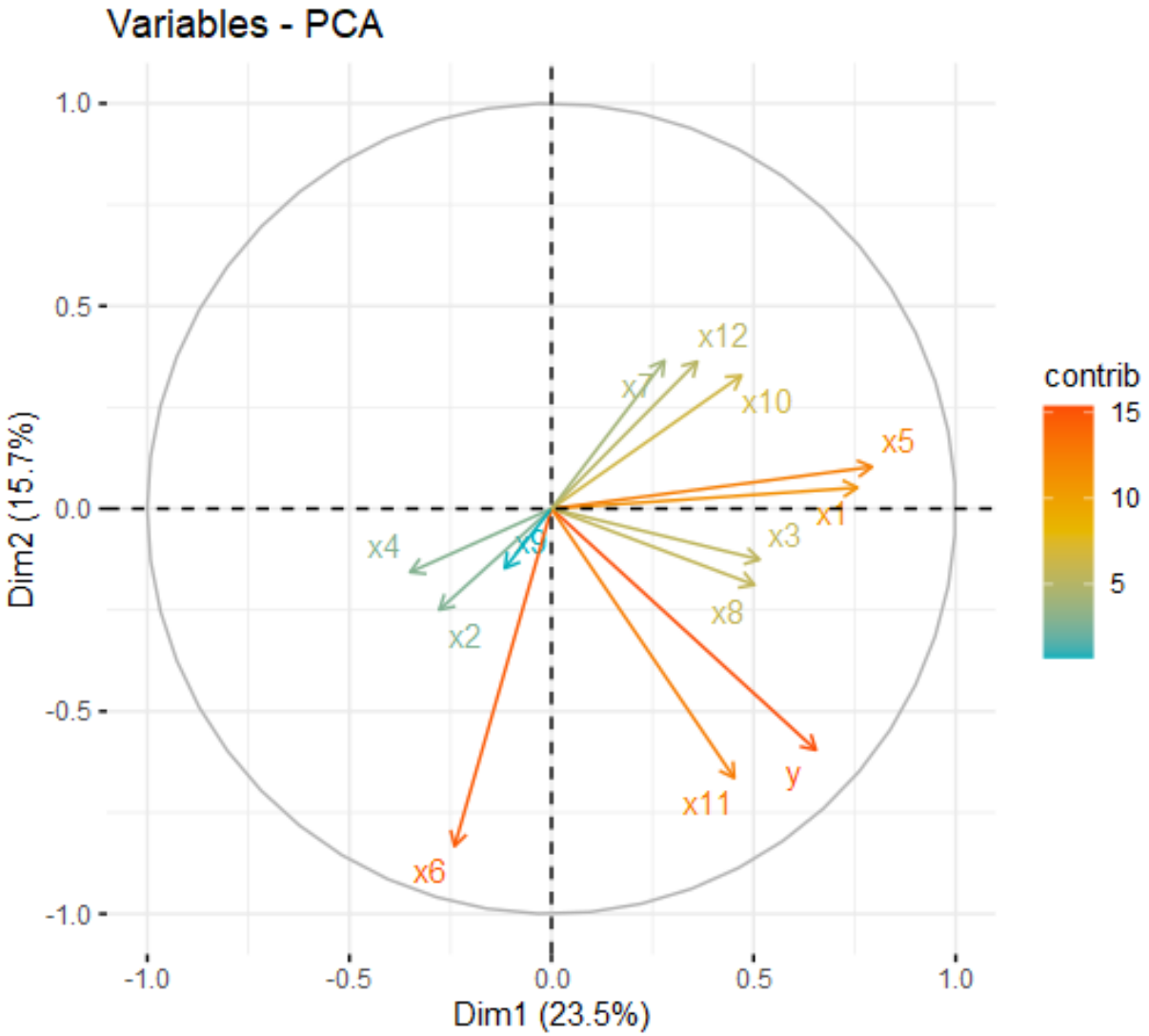






ANN – 1st/2nd year sets [train/test] + X_n (??)

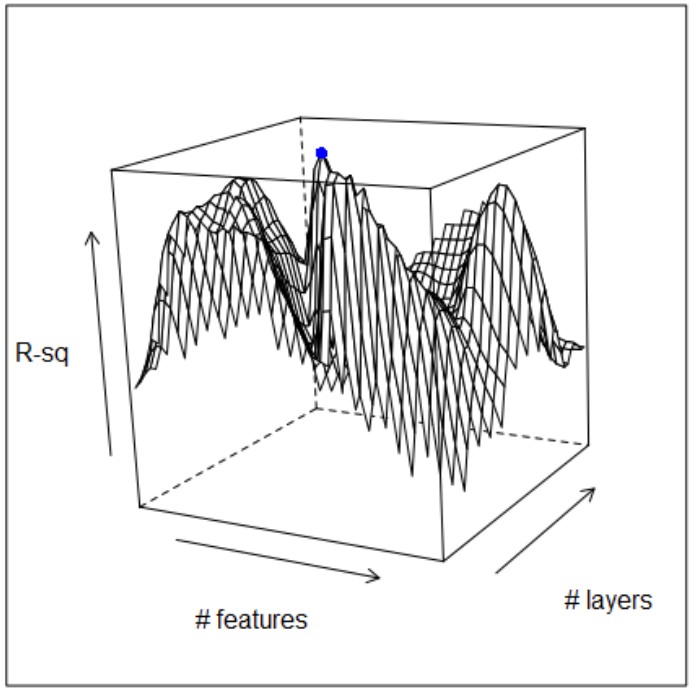
(RMSE=??; R-sq=??; Pearson=??; d-Willmott index=??)



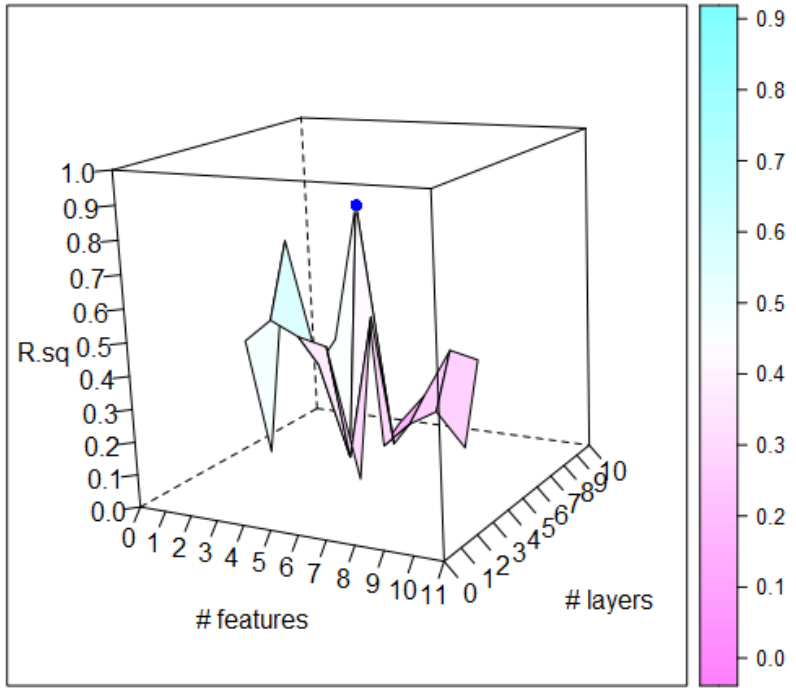
ANN – 1st/2nd year sets [train/test] -> *X1 + X3 + X5 + X6 + X10 + X11*

(*Train RMSE=7.9 mm; Test RMSE=9.1 mm*)

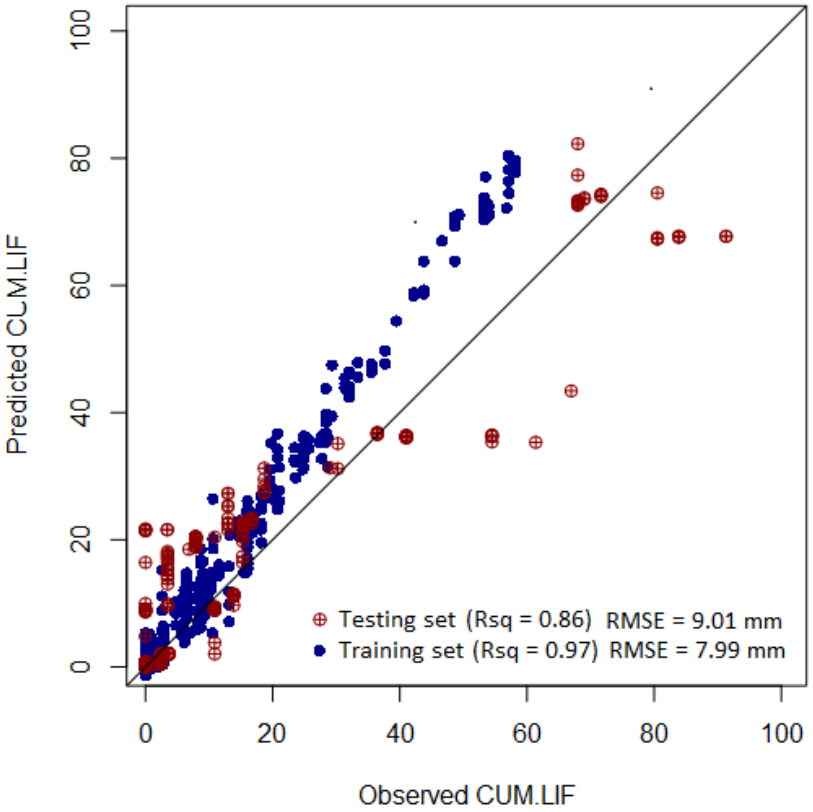
A) **ANN R-sq surface plot**



B) **ANN structure and performance assessment**



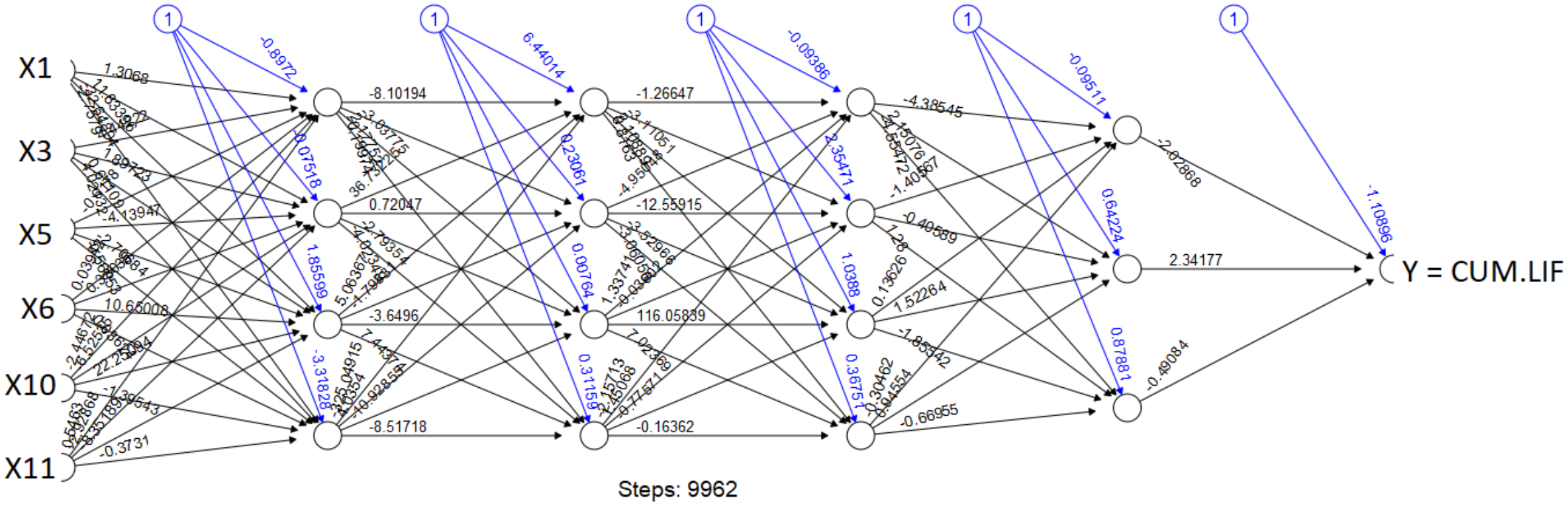
C)



- static multi-layer feedforward network;
- 6 input features;
- 4 hidden layers (w/logistic transformations);
- 3-4 neurons per layer [4,4,4,3]

ANN – 1st/2nd year sets [train/test] -> *X1 + X3 + X5 + X6 + X10 + X11*

(RMSE=9.8 mm; R-sq=0.86; Pearson=0.93; d-Willmott index=0.95)



Best performing **ANN structure**:

- static multi-layer feedforward network;
- **6 input features**;
- **4 hidden layers** (w/logistic transformations);
- **3-4 neurons** per layer [4,4,4,3]

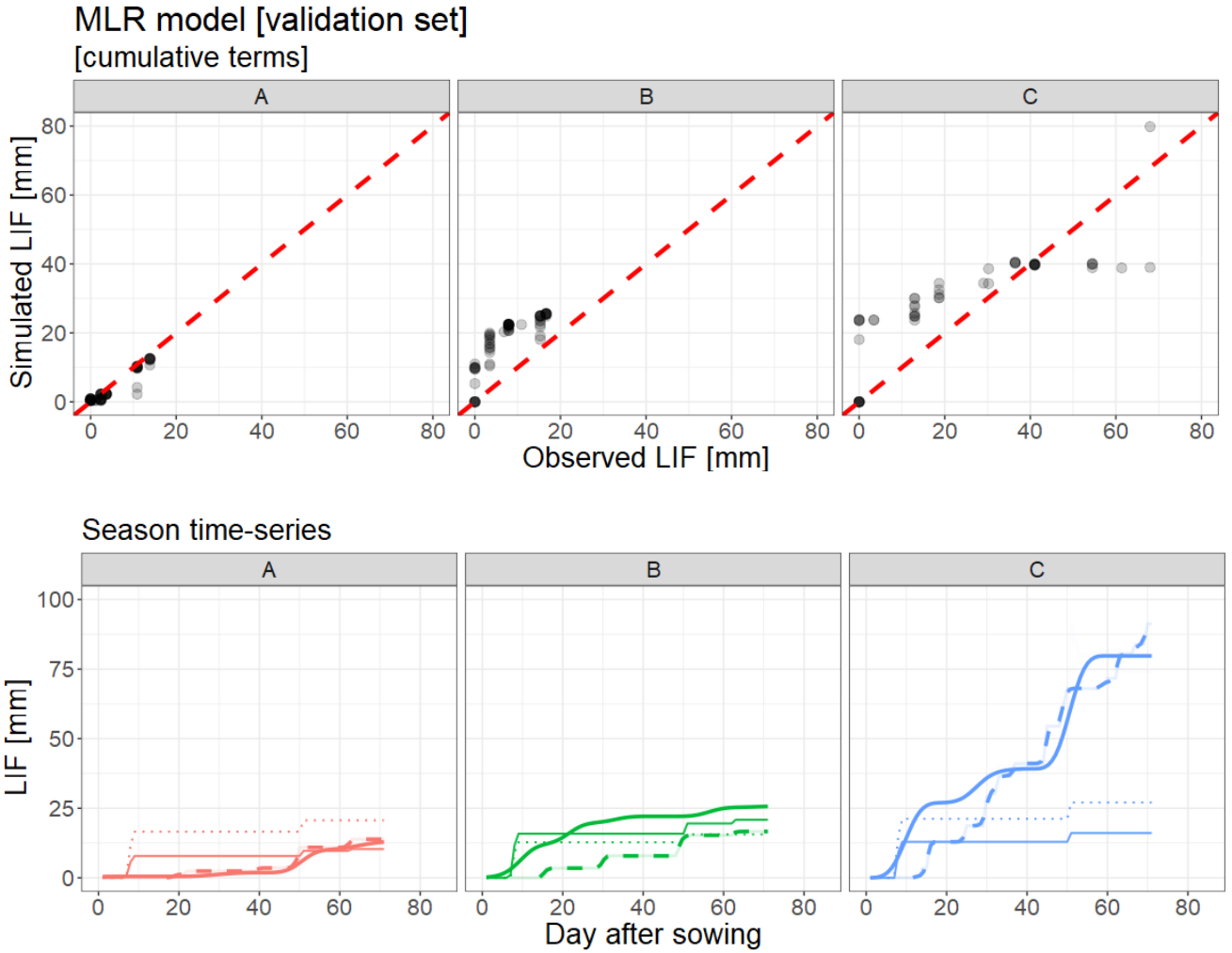
ANN – 1st/2nd year sets [train/test] -> *X1 + X3 + X5 + X6 + X10 + X11*

(RMSE=9.8 mm; R-sq=0.86; Pearson=0.93; d-Willmott index=0.95)

```
data=data.frame(CVUT.LIF$Zone,                    #Zone
                CVUT.LIF$Year,                    #Year
                CVUT.LIF$DAS,                     #DAS
                CVUT.LIF$LIF.C.CUM,               #y
                CVUT.LIF$HYD.10min.CUM,           #x1
                CVUT.LIF$DAR.off,                 #x2
                normalize(CVUT.LIF$FAI),          #x3
                CVUT.LIF$slope,                   #x4
                CVUT.LIF$SR.contribution,         #x5
                CVUT.LIF$CC,                      #x6
                CVUT.LIF$Ks.10min.surface,        #x7
                CVUT.LIF$DAR.off.c,               #x8
                CVUT.LIF$DAPre,                   #x9
                CVUT.LIF$K.sat,                   #x10
                CVUT.LIF$P.CUM,                   #x11
                CVUT.LIF$satunsat                 #x12
                )
```

Best performing ANN structure:

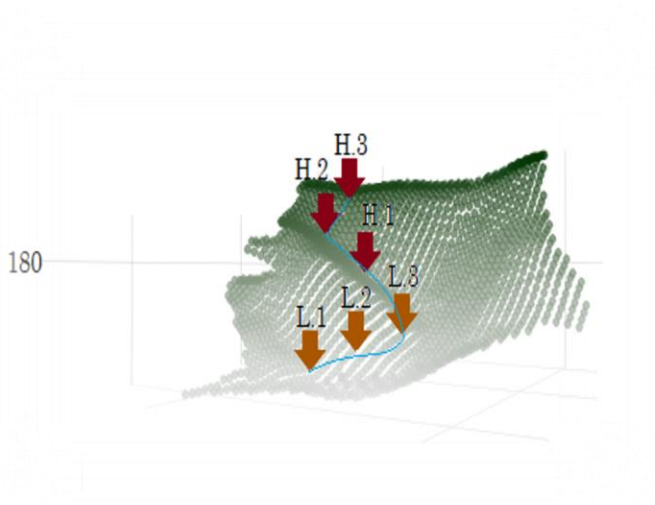
- static multi-layer feedforward network;
- 6 input features;
- 4 hidden layers (w/logistic transformations);
- 3-4 neurons per layer [4,4,4,3]

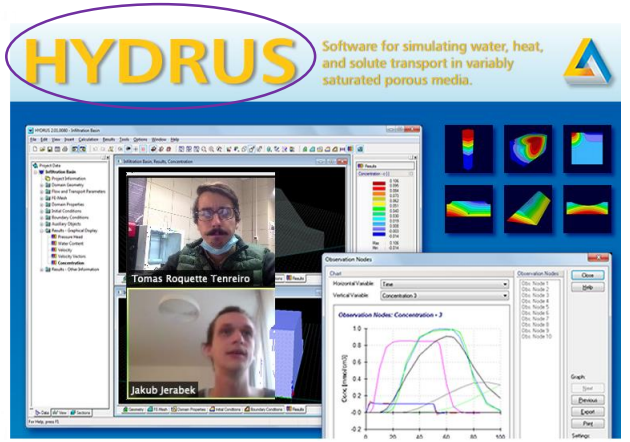


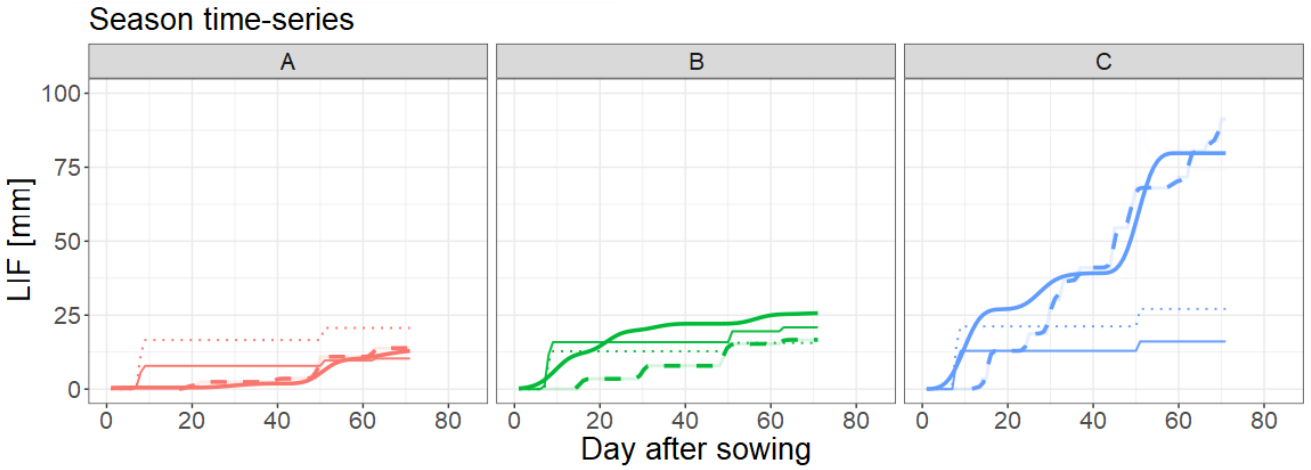
ANN – 1st/2nd year sets [train/test] -> X1 + X3 + X5 + X6 + X10 + X11

(RMSE=9.8 mm; R-sq=0.86; Pearson=0.93; d-Willmott index=0.95)

```
data=data.frame(CVUT.LIF$Zone,                  #Zone
                CVUT.LIF$Year,                  #Year
                CVUT.LIF$DAS,                   #DAS
                CVUT.LIF$LIF.C.CUM,             #y
                CVUT.LIF$HYD.10min.CUM,         #x1
                CVUT.LIF$DAR.off,               #x2
                normalize(CVUT.LIF$FAI),        #x3
                CVUT.LIF$slope,                 #x4
                CVUT.LIF$SR.contribution,       #x5
                CVUT.LIF$CC,                    #x6
                CVUT.LIF$Ks.10min.surface,      #x7
                CVUT.LIF$DAR.off.c,             #x8
                CVUT.LIF$DAPre,                 #x9
                CVUT.LIF$K.sat,                 #x10
                CVUT.LIF$P.CUM,                 #x11
                CVUT.LIF$satunsat               #x12
                )
```





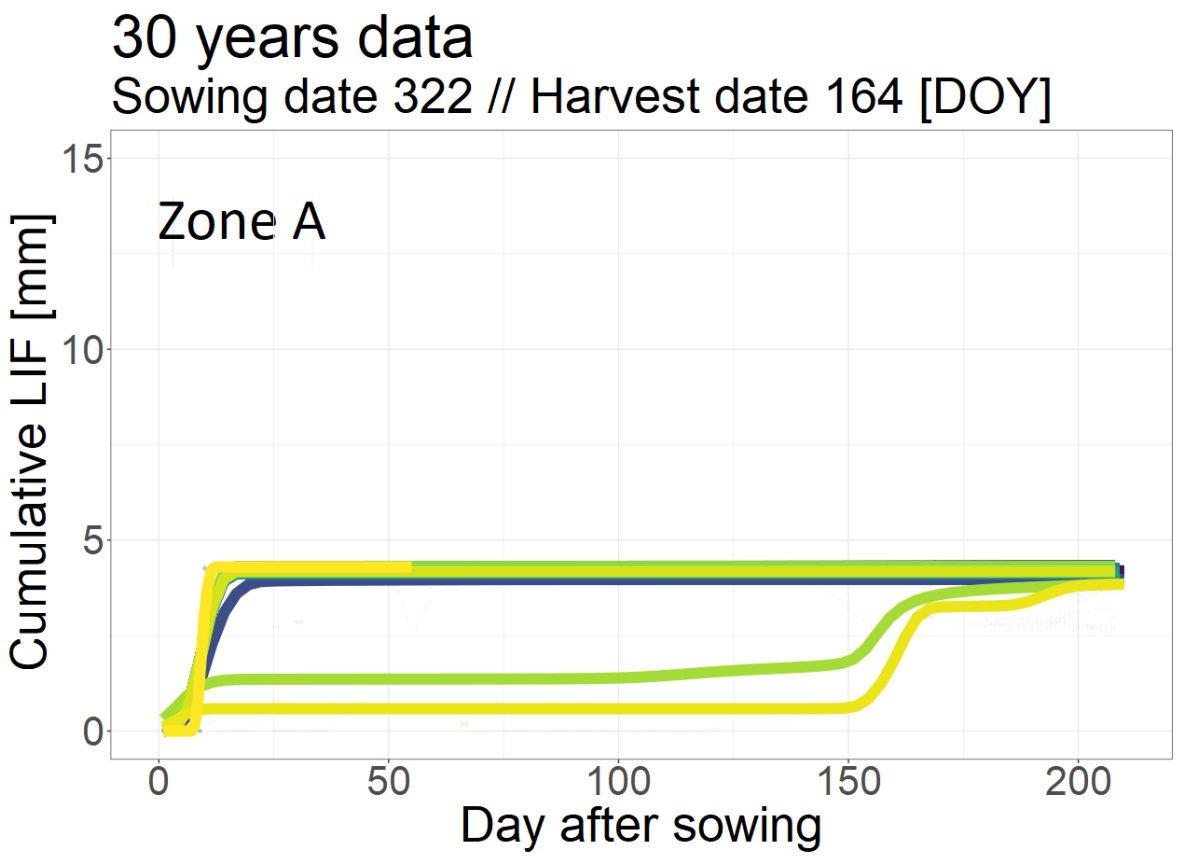


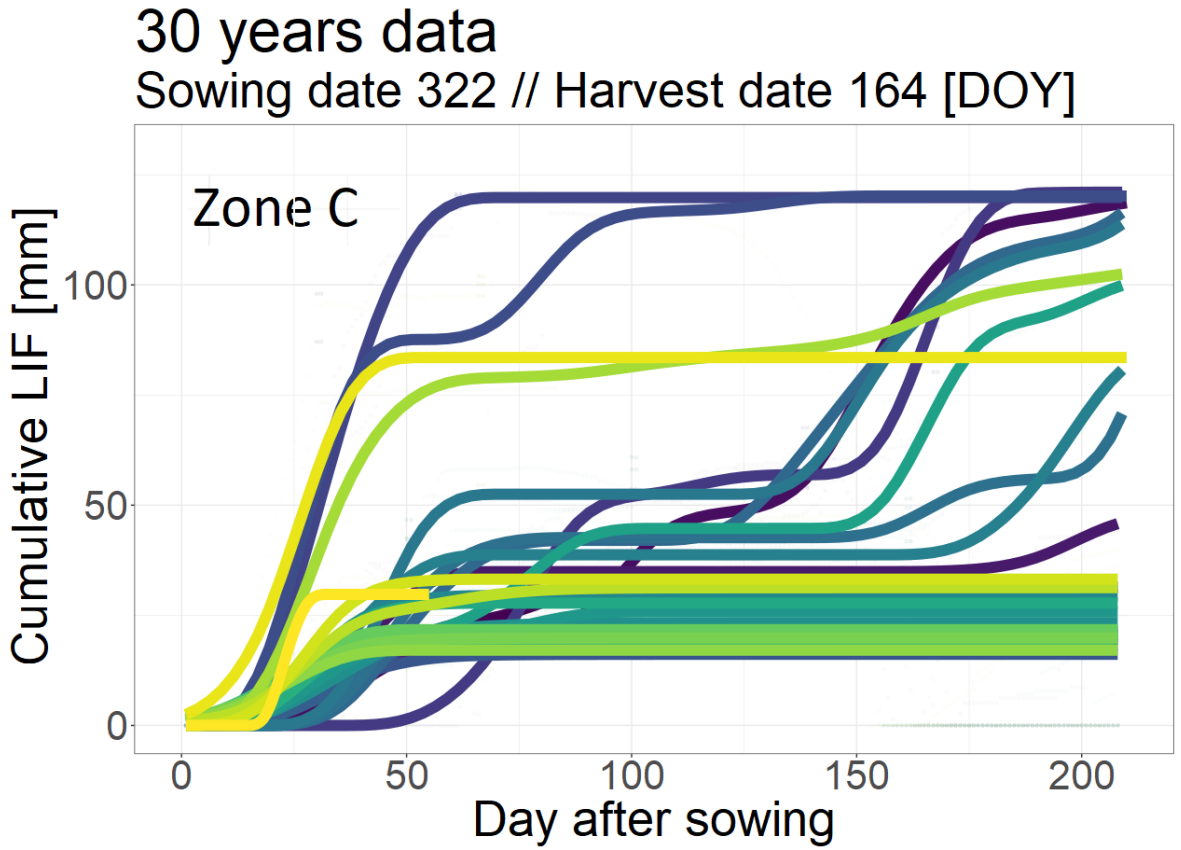
Best performing ANN structure:

- static multi-layer feedforward network;
- 6 input features;
- 4 hidden layers (w/logistic transformations);
- 3-4 neurons per layer [4,4,4,3]

ANN $X1 + X3 + X5 + X6 + X10 + X11$

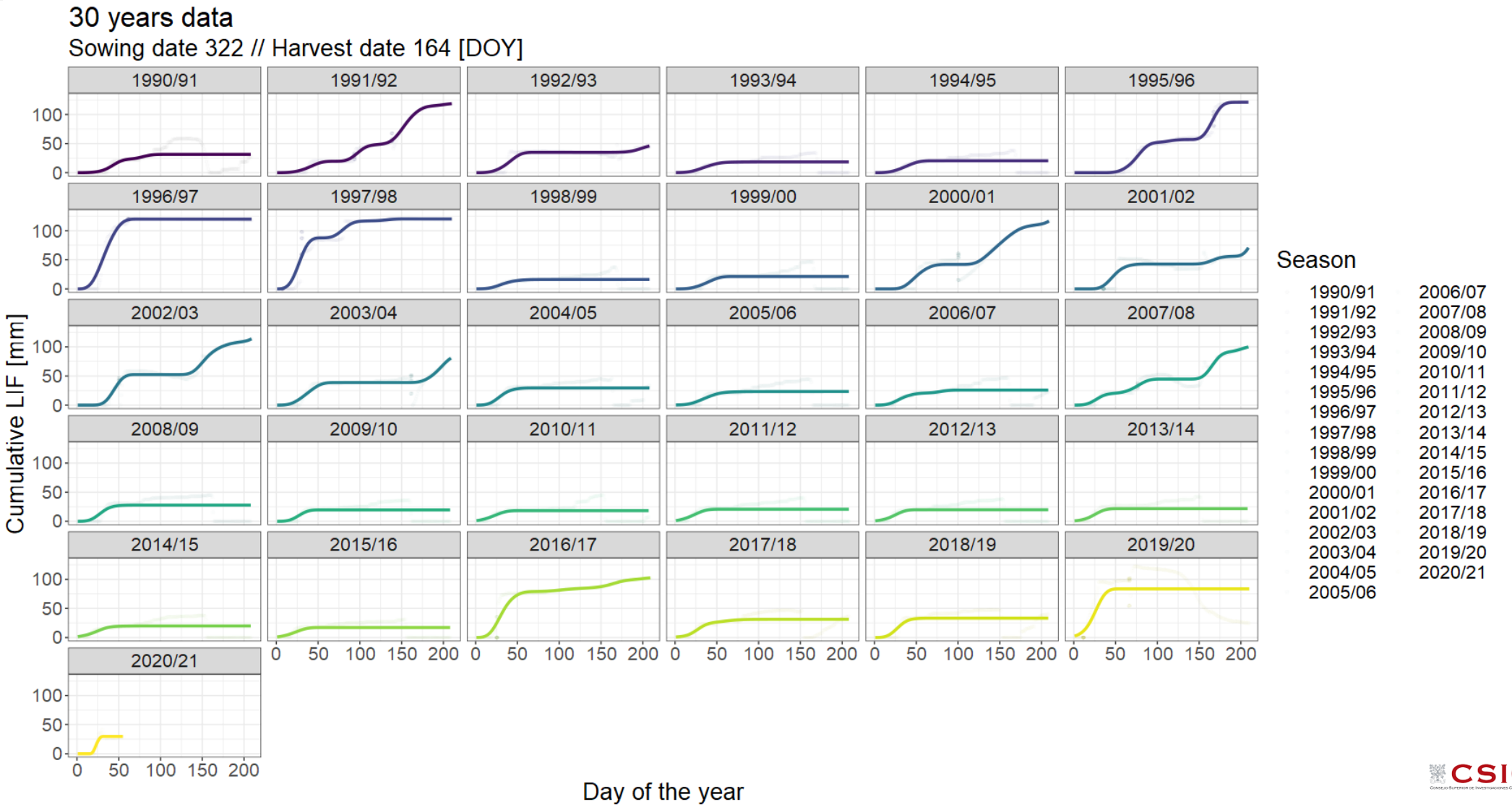
Lateral inflow in cumulative terms (CUM.LIF) predicted by the ANN for a sequence of 30 years





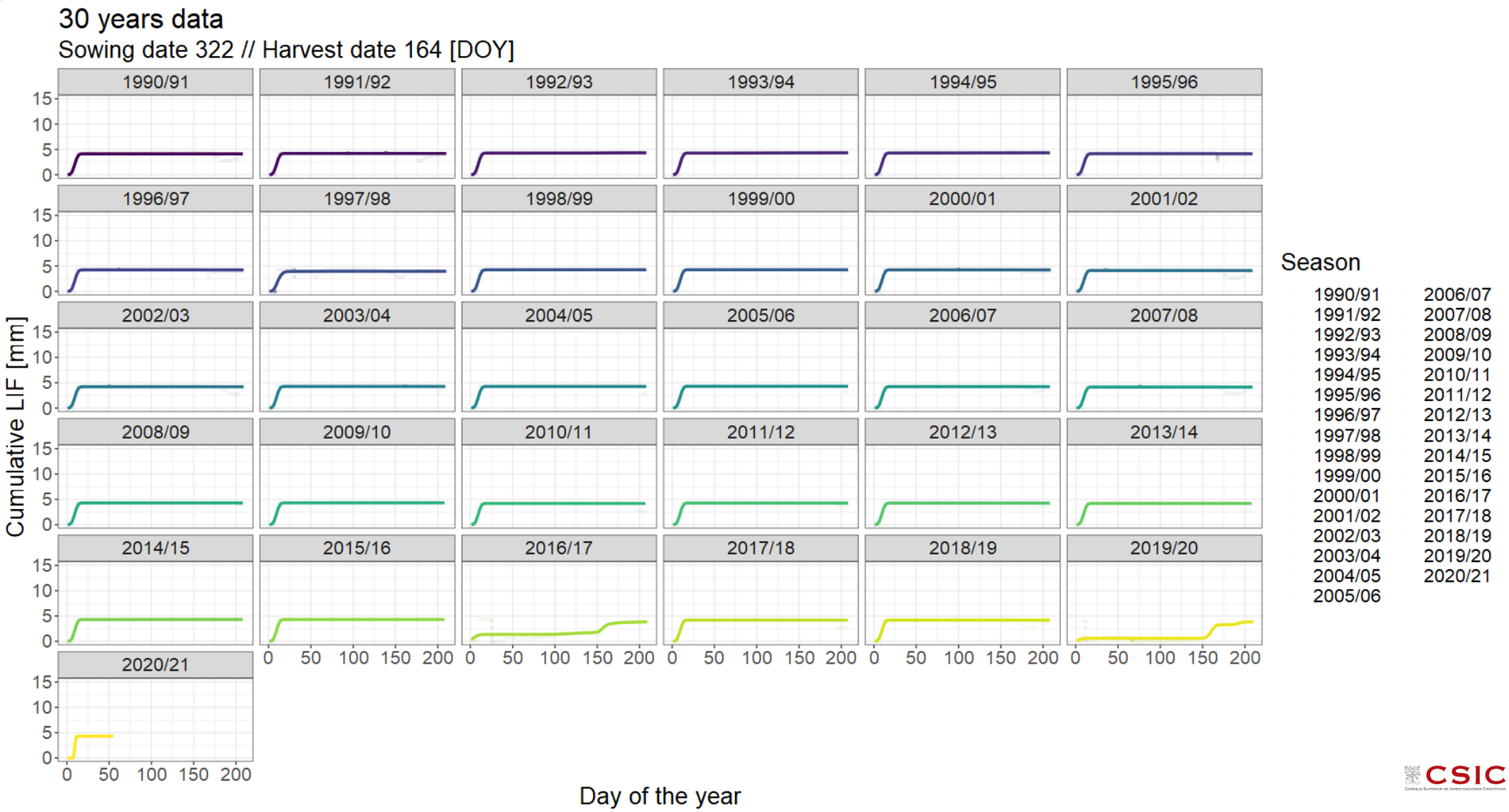
ANN – 1st/2nd year sets [train/test] -> *X1 + X3 + X5 + X6 + X10 + X11*

Lateral inflow in cumulative terms (CUM.LIF) predicted by the ANN for a sequence of 30 years



ANN – 1st/2nd year sets [train/test] -> *X1 + X3 + X5 + X6 + X10 + X11*

Lateral inflow in cumulative terms (CUM.LIF) predicted by the ANN for a sequence of 30 years



ANN – 1st/2nd year sets [train/test] -> *X1 + X3 + X5 + X6 + X10 + X11*

(RMSE=9.8 mm; R-sq=0.86; Pearson=0.93; d-Willmott index=0.95)

Post-prediction inference correction ??

Wang, S., McCormick, T. H., & Leek, J. T. (2020).
Methods for correcting inference based on outcomes predicted by machine learning.
Proceedings of the National Academy of Sciences, *117*(48), 30266-30275.