Model type II regression for lagrangian validation of HF Radar velocities in the NW Iberian Peninsula

Adrián Martínez-Fernández, Waldo Redondo-Caride, Fernando Alonso-Pérez, Silvia Piedracoba, Pablo Lorente, Silvia Allen-Perkins, Pedro Montero, Garbiñe Ayensa, Silvia Torres-López, Adrián Fernández-Baladrón, Ramiro A. Varela, Antón Velo, Miguel Gilcoto

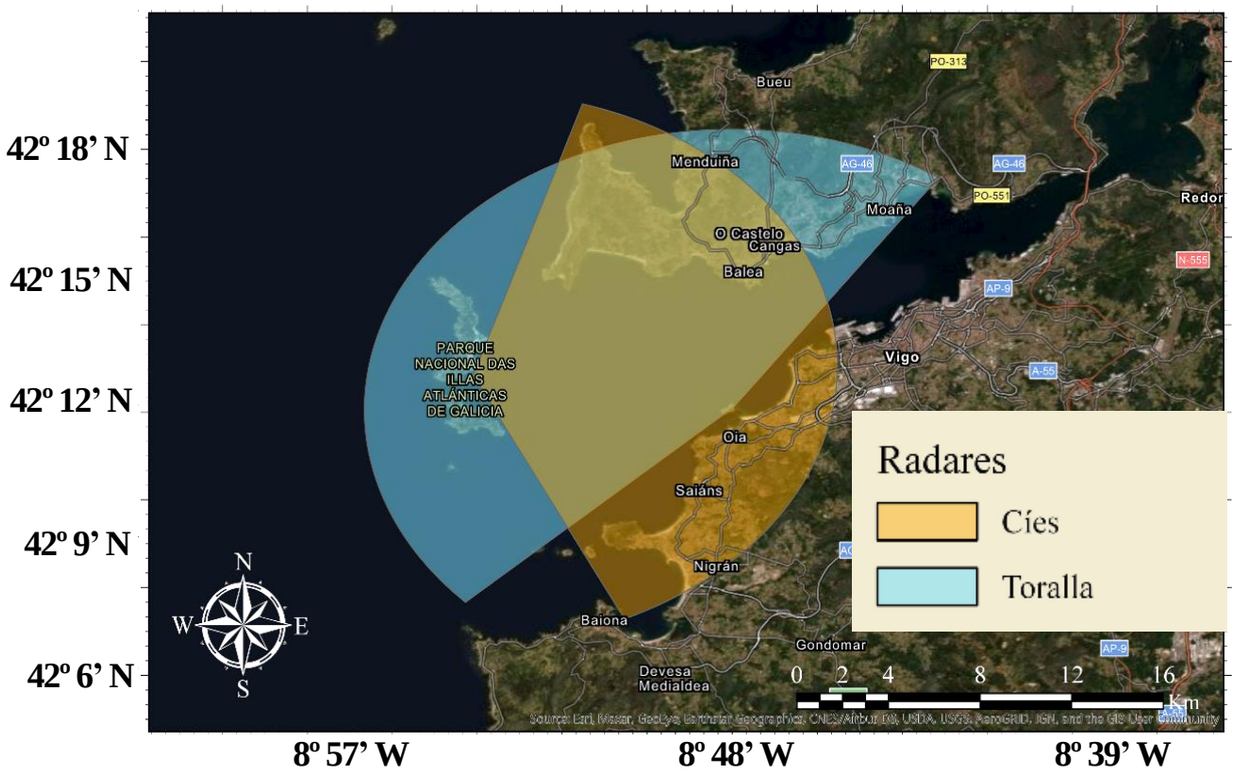
CONTEXT & STUDY AREA

Interreg España - Portugal
Fondo Europeo de Desarrollo Regional
Fundo Europeu de Desenvolvimento Regional
UNIÓN EUROPEA
UNIÃO EUROPEIA

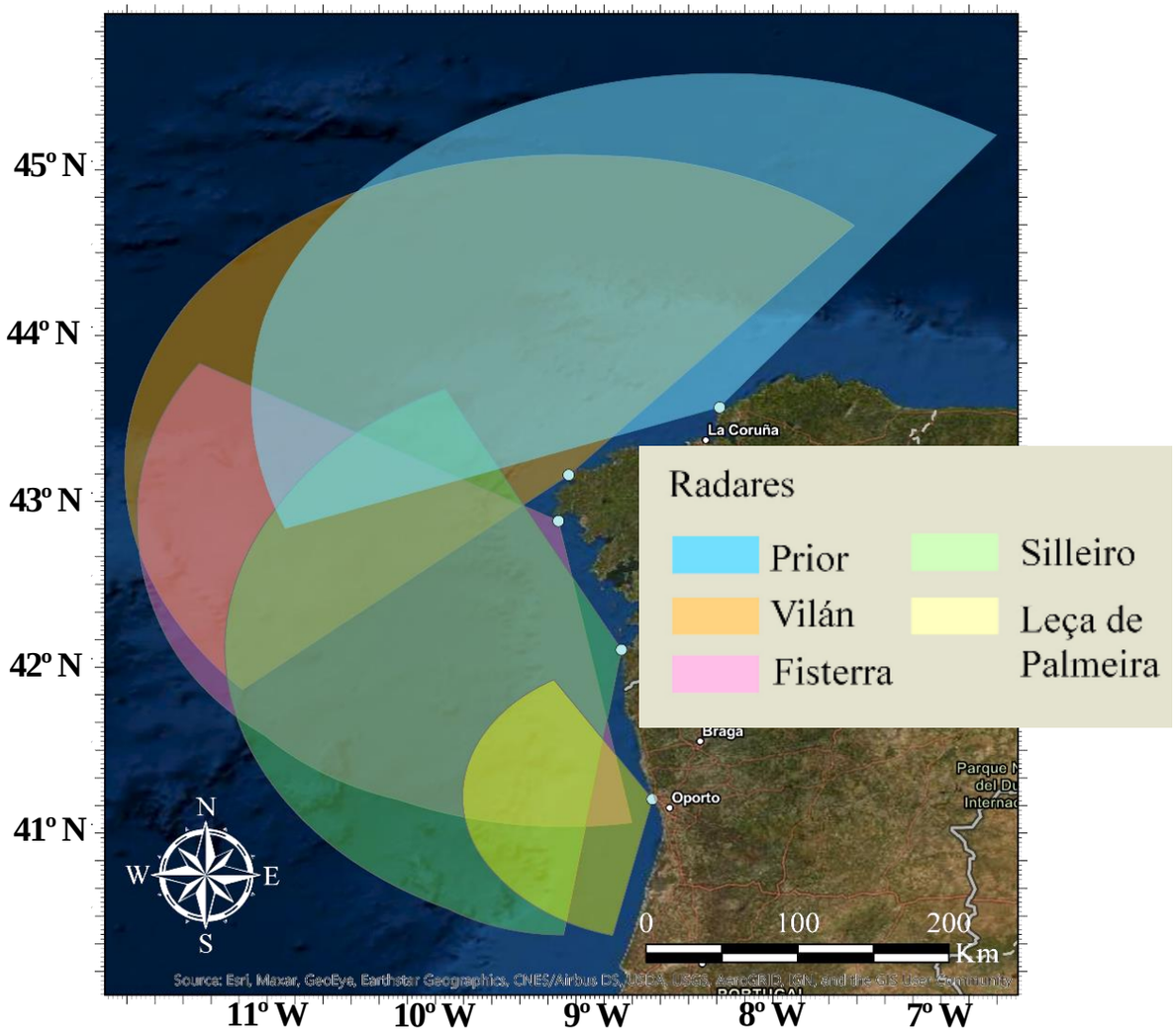
RADAR ON RAIA

MAR TECH

HF RAIA Vigo



HF RAIA NW Iberian



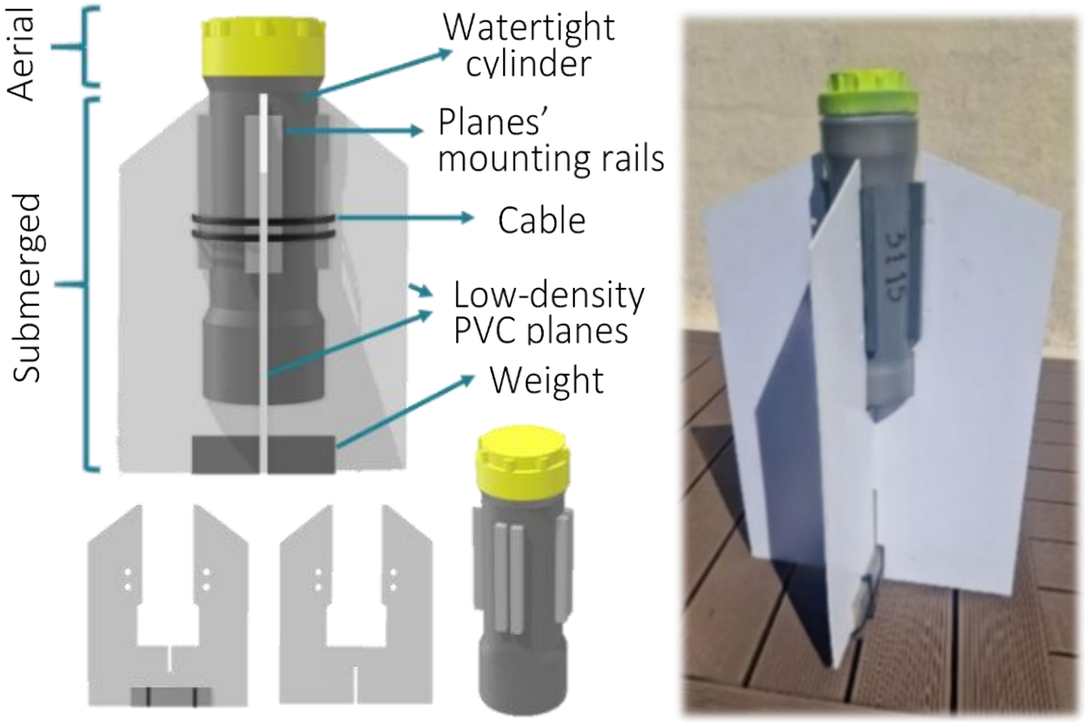
For further information please check: Adrián Martínez-Fernández et al., MARTECH 2021: SPOT and GPRS drifting buoys for HF Radar calibration

Model type II regression for lagrangian validation of HF Radar velocities in the NW Iberian Peninsula

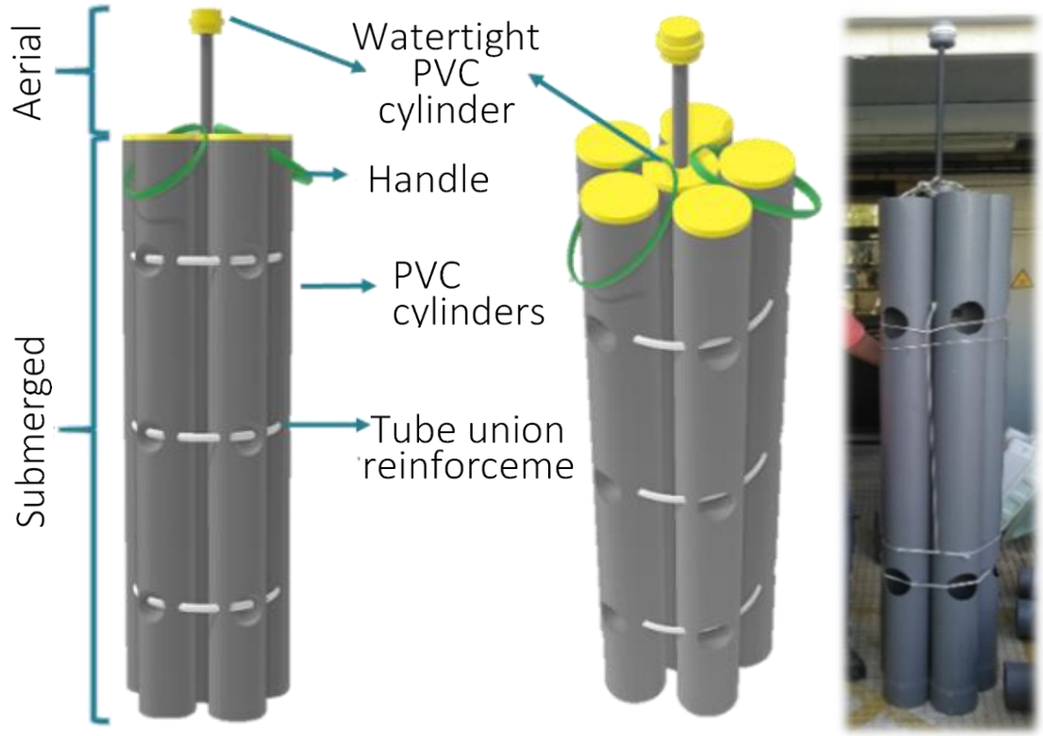
Adrián Martínez-Fernández et al.

LOW COST BUOYS

GPRS



SPOT



For further information please check: Adrián Martínez-Fernández et al., MARTECH 2021: SPOT and GPRS drifting buoys for HF Radar calibration

DATASETS FOR VALIDATION

HFR Data

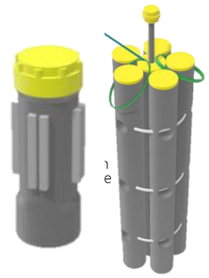
Buoys Data

Filtered by quality

Matched in time and space data

Filtering by percentile 80 of velocity's SD

I. VFLG (Vector flag – quality parameter)



I. Data with same date (tolerance of 5 seconds).

II. Data with velocities over 120 cm/s.

III. Outliers (data more than 2 standard deviations away from the mean) in latitude and longitude using 900 second moving windows.

DATASETS FOR VALIDATION

HFR Data

Buoys Data

↓

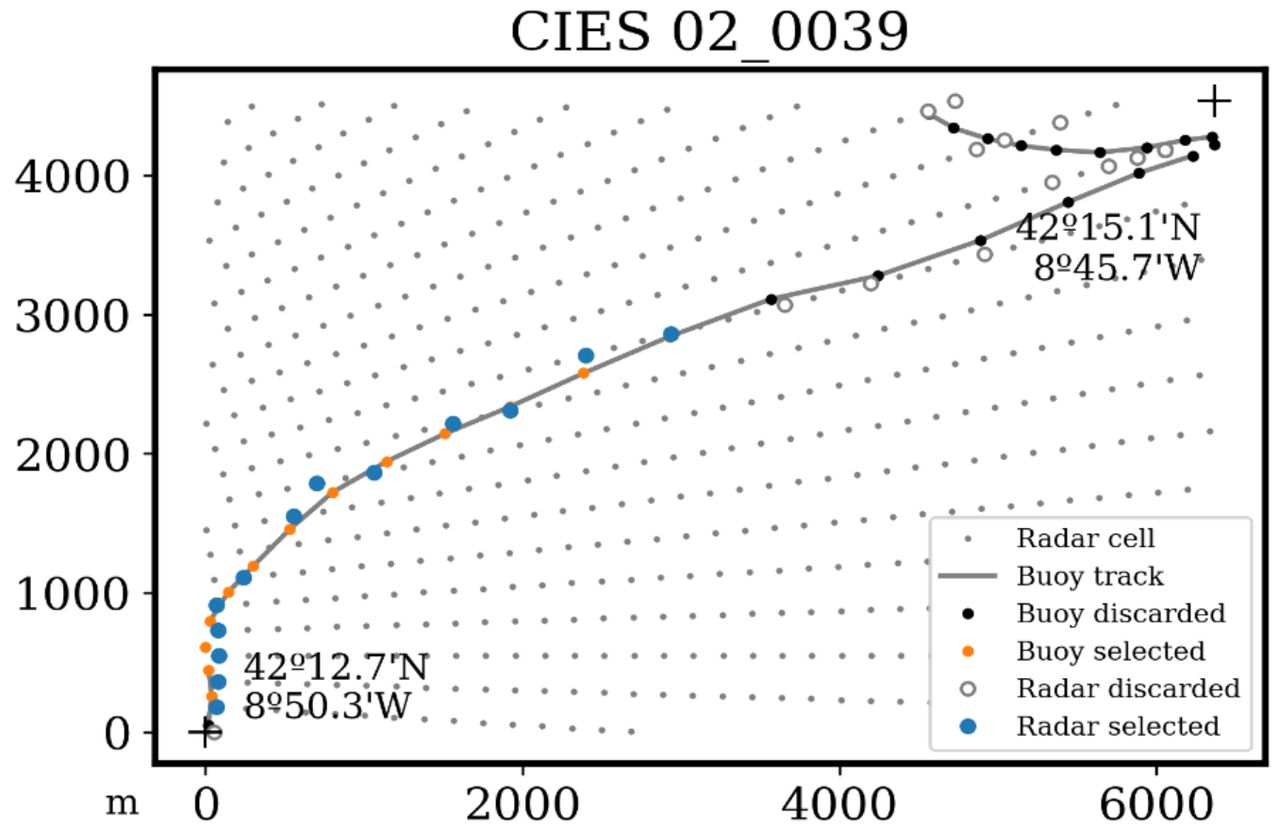
Filtered by quality

↓

Matched in time and space data

↓

Filtering by percentile 80 of velocity's SD



Buoy traces (grey line), radar positions (grey dots), buoy data (orange dots) matching good radar data (blue dots), and buoy data (black dots) not matching data because they do not exist or are bad (VFLG > 0, black circles).

DATASETS FOR VALIDATION

HFR Data

Buoys Data

Filtered by quality

Matched in time and space data

Filtering by percentile 80 of velocity's SD

Filtering by percentile 80 of velocity's SD

Filtered In

↑ **Homogeneity**

↑ **Quality**

All data
=
Filtered in
+
Filtered Out

DATA COMPARISONS

Traditionally

Regression Model I

$$Y + \in_Y = intercept + slope \cdot X$$

Y: Radar, X: Buoy

Only contains error in predictor variable (X)

Both HFR and buoys contains errors

Proposed

Regression Model II

$$Y + \in_Y = intercept + slope \cdot (X + \in_X)$$

Y: Radar, X: Buoy

Contains error in both predictor and predicted variables (X & Y)

RESULTS

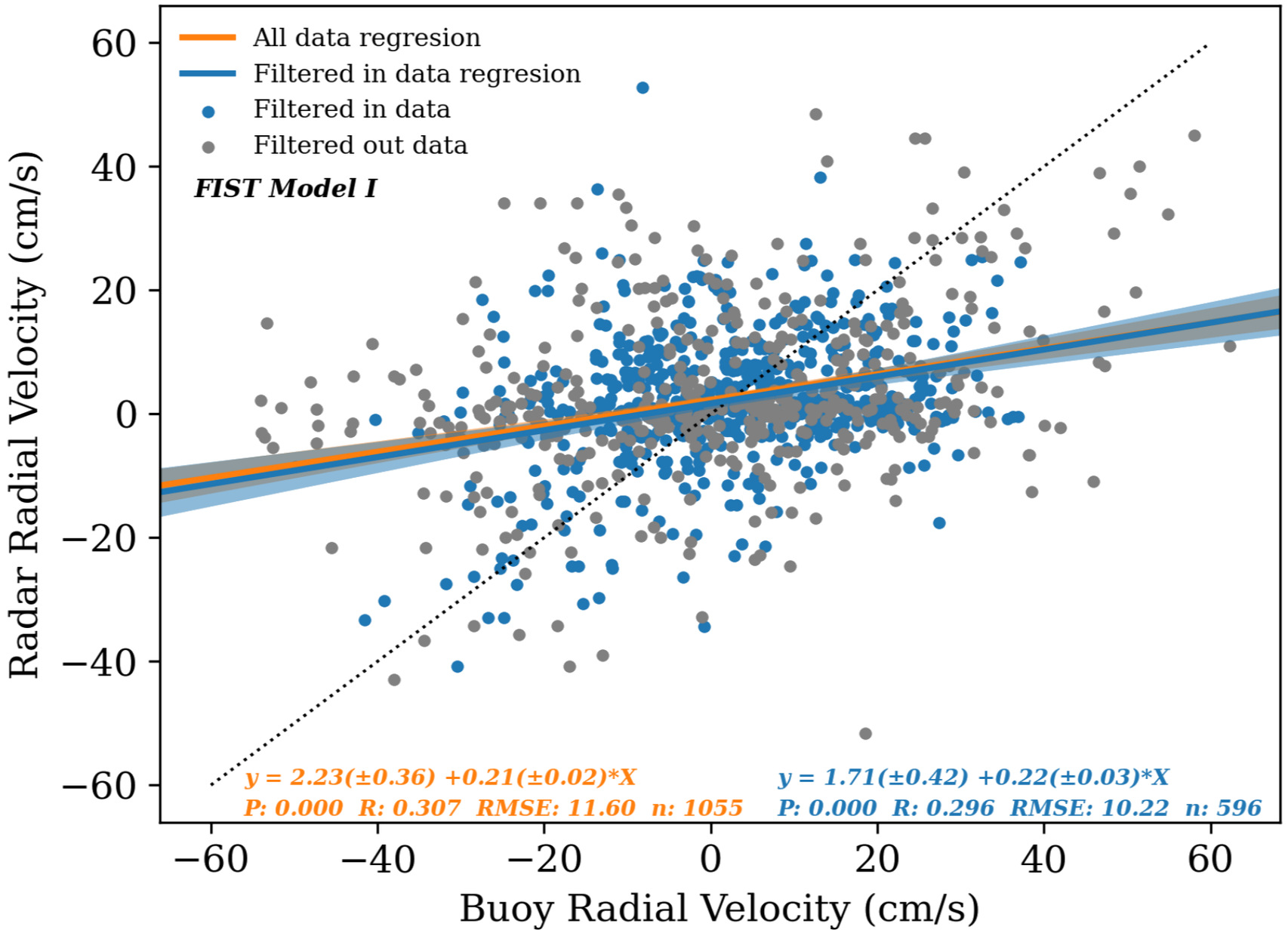
Regression Model I
+
Filtered data

Regression Model II
+
Filtered data

Regression Model I
+
Unfiltered data

Regression Model II
+
Unfiltered data

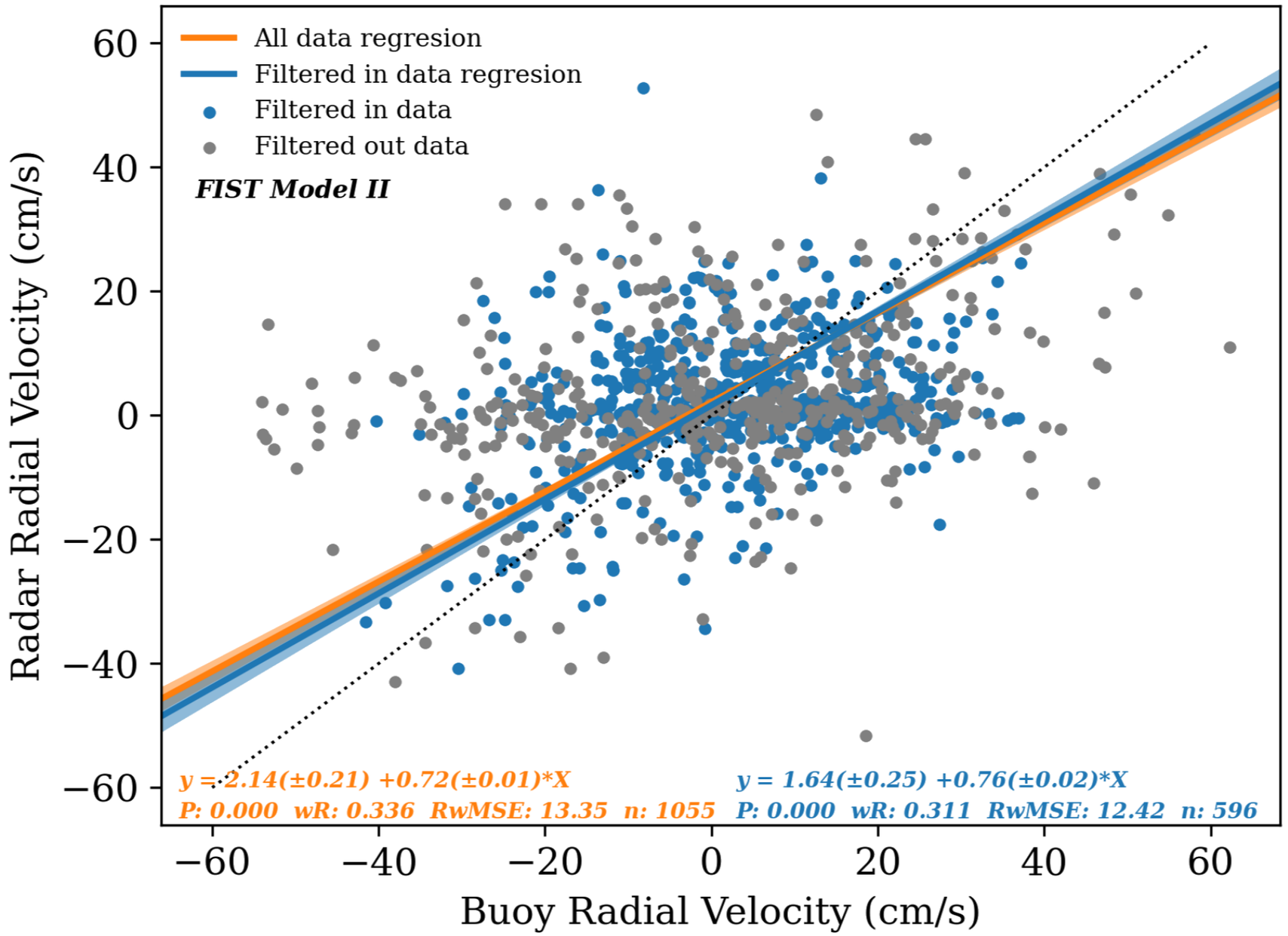
REGRESION MODEL I



Model type II regression for lagrangian validation of HF Radar velocities in the NW Iberian Peninsula

Adrián Martínez-Fernández et al.

REGRESION MODEL II



Model type II regression for lagrangian validation of HF Radar velocities in the NW Iberian Peninsula

Adrián Martínez-Fernández et al.

COMPARING MODELS

$$\widehat{Y} = a + bx$$

$$R$$

$$Var(\widehat{Y})$$

$$RMSE \ / \ RMSD$$

COMPARING MODELS

$$RMSE^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\hat{Y}_i - Y_i\right)^2 \qquad (1)$$

$$RMSD^2 = \frac{1}{n} \cdot \sum_{i=1}^{n} \left(\left(\hat{Y}_i - E\left(\hat{Y}\right)\right) - \left(Y_i - E(Y)\right) \right)^2 \qquad (2)$$

$$RMSD^2 = RMSE^2 - \left(E\left(\hat{Y}\right) - E(Y)\right)^2 \qquad (3)$$

Linear regressions of 1 variable

$$E(Y) = E\left(\hat{Y}\right)$$

$$RMSD = RMSE$$

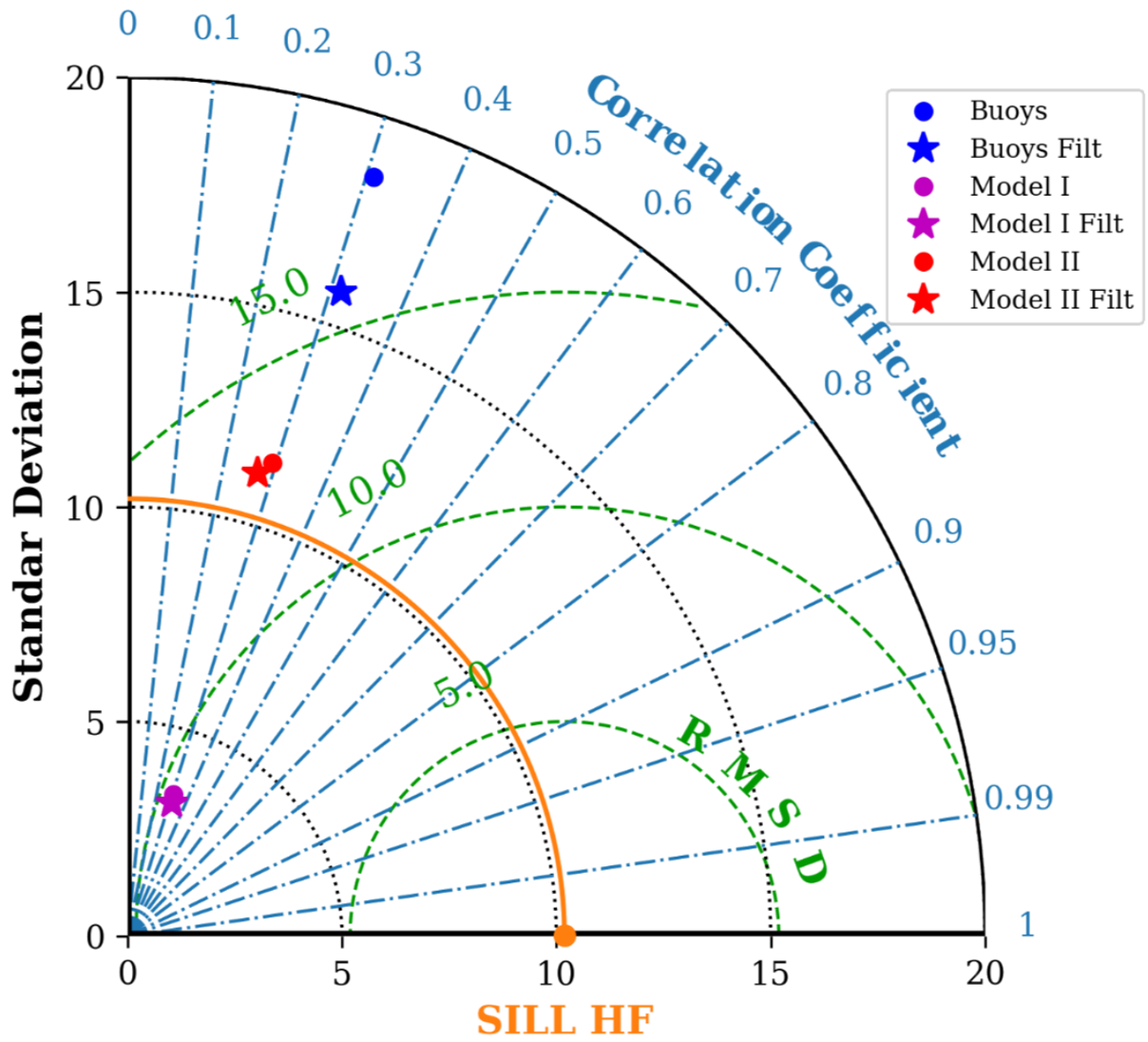
$$RMSD^2 = Var(\hat{Y}) + Var(Y) - 2 \cdot SD(\hat{Y}) \cdot SD(Y) \cdot R_{Y\hat{Y}} \quad (4)$$

$$\downarrow RMSD \quad \uparrow Performance$$

COMPARING MODELS

$$RMSD^2 = Var(\hat{Y}) + Var(Y) - 2 \cdot SD(\hat{Y}) \cdot SD(Y) \cdot R_{Y\hat{Y}}$$

(4)



COMPARING MODELS

Given by original data

$$RMSD^2 = Var(\hat{Y}) + Var(Y) - 2 \cdot SD(\hat{Y}) \cdot SD(Y) \cdot R_{Y\hat{Y}} \quad (4)$$

Can be modified

$$SD(\hat{Y}) = \sqrt{Var(\hat{Y})}$$

Regressions of 1 Variable $R_{Y\hat{Y}} = R_{XY}$
(R_{XY} given by original data)

$$\downarrow RMSD \quad \uparrow Performance$$

NEW METRIC: RWSD

Root Weighted Square Difference

$$RWSD^2 = \left(SD(\hat{Y}) - SD(Y)\right)^2 \cdot w_V + RMSD^2 \cdot w_R \quad (5)$$

$$w_V + w_R = 1$$

NEW METRIC: RWSD

Root Weighted Square Difference expanded

$$RWSD^2 = Var(\hat{Y}) + Var(Y) - 2 \cdot SD(\hat{Y}) \cdot SD(Y) \cdot (w_V + R_{Y\hat{Y}} \cdot w_R)$$

(6)

NEW METRIC: RWSD

Root Weighted Square Difference expanded

$$RWSD^2 = Var(\hat{Y}) + Var(Y) - 2 \cdot SD(\hat{Y}) \cdot SD(Y) \cdot (w_V + R_{Y\hat{Y}} \cdot w_R) \quad (6)$$

With minimum in function of $Var(\hat{Y})$

$$min\{RWSD^2\} = Var(Y) \cdot (1 - (w_V + R_{Y\hat{Y}} \cdot w_R)^2) \quad (7)$$

Produced when:

$$SD(\hat{Y}) = SD(Y) \cdot (w_V + R_{Y\hat{Y}} \cdot w_R) \quad (8)$$

NEW METRIC: RWSD

Root Weighted Square Difference expanded

$$RWSD^2 = Var(\hat{Y}) + Var(Y) - 2 \cdot SD(\hat{Y}) \cdot SD(Y) \cdot (0.75 + R_{Y\hat{Y}} \cdot 0.25)$$

(9)

$$w_V = 0.75$$

$$w_R = 0.25$$

RESULTS

Regression models chosen per antenna according to the criterion of lowest RMSD (eq. 4)

	Regression	Slope	Intercept	R^2	N	RMSD	RWSD	SD Diff
PRIO	MI_f	0.198 (±0.027)	+2.169 (±0.501)	0.11	447	10.60	8.36	7.484
VILA	MII	0.047 (±0.007)	+0.482 (±0.168)	0.11	311	2.58	1.86	1.540
FIST	MI_f	0.217(±0.029)	+1.714 (±0.423)	0.09	596	10.22	8.28	7.526
SILL	MI_f	0.212 (±0.021)	+0.578 (±0.334)	0.11	873	9.83	7.81	7.015
LPRO	MII	0.427 (±0.016)	-1.218 (±0.344)	0.21	202	9.67	4.95	1.282
GALI_vx	MI_f	0.383 (±0.022)	-1.503 (±0.380)	0.29	730	10.24	7.07	5.630
GALI_vy	MI	0.304 (±0.019)	-0.553 (±0.399)	0.18	1116	12.50	9.27	7.912
CIES	MI_f	0.708 (±0.069)	+2.856 (±0.524)	0.18	493	11.47	8.54	7.32
TORA	MI	0.120 (±0.020)	+1.506 (±0.140)	0.07	461	2.71	2.23	2.045
VIGO_vx	MII_f	0.925 (±0.053)	-0.793 (±0.286)	0.73	162	2.78	1.43	6.16
VIGO_vy	MI_f	0.000 (±0.059)	-1.625 (±0.382)	0.00	176	4.42	4.12	5.010

MI: Modelo I
MI_f: Modelo I filtrado
MII: Modelo II
MII_f: Modelo II

Model type II regression for lagrangian validation of HF Radar velocities in the NW Iberian Peninsula

Adrián Martínez-Fernández et al.

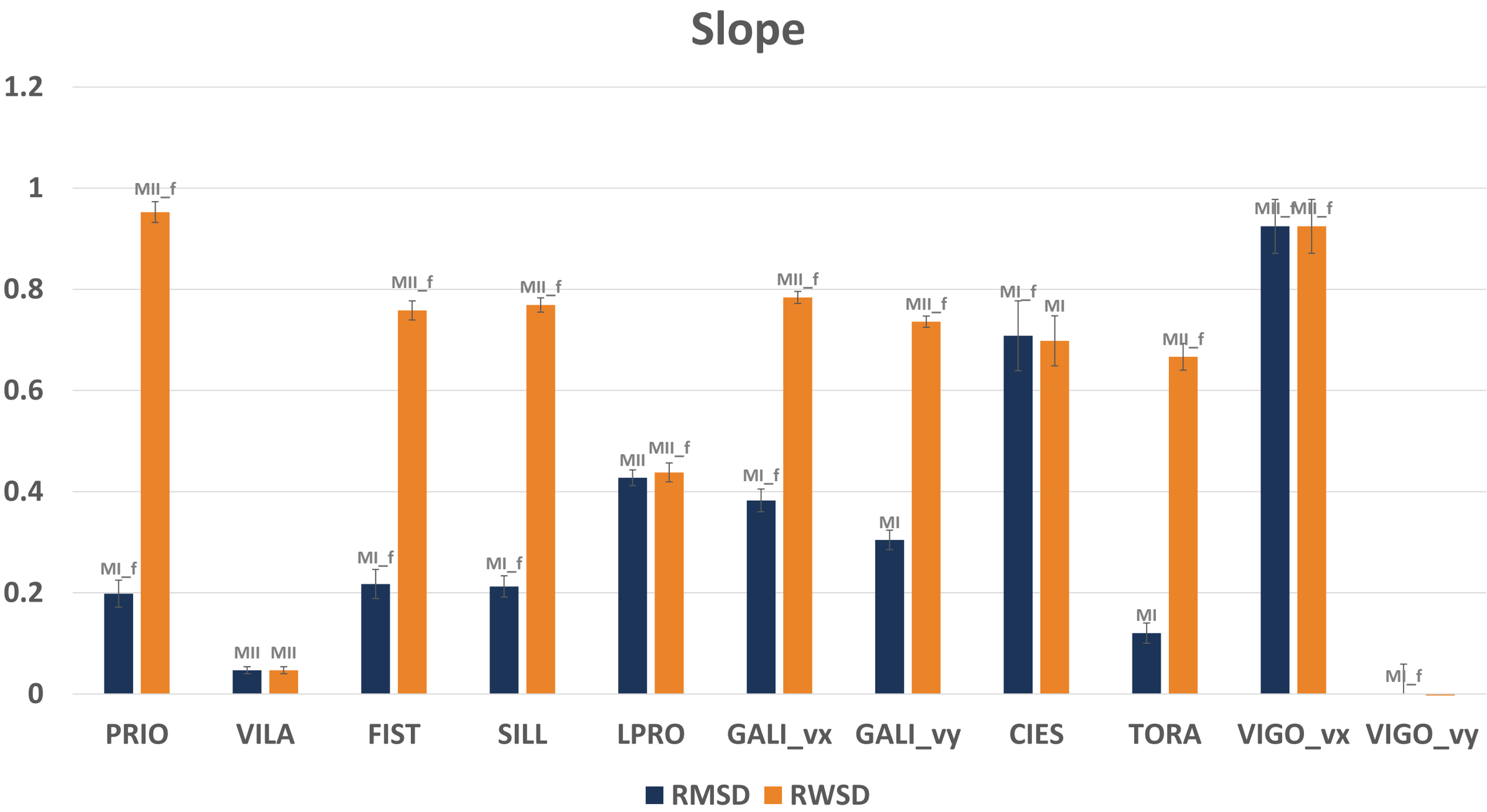
RESULTS

Regression models chosen per antenna according to the criterion of lowest RWSD (eq. 9)

	Regression	Slope	Intercept	R^2	N	RMSD	RWSD	SD Diff
PRIO	MII_f	0.953 (±0.021)	+10.296 (±0.262)	0.19	447	15.96	8.01	0.956
VILA	MII	0.047 (±0.007)	+0.482 (±0.168)	0.11	311	2.58	1.86	1.540
FIST	MII_f	0.758 (±0.019)	+1.638 (±0.245)	0.10	596	12.42	6.21	0.140
SILL	MII_f	0.769 (±0.014)	-0.896 (±0.178)	0.07	873	13.43	6.87	1.688
LPRO	MII_f	0.438 (±0.019)	-1.078 (±0.382)	0.20	106	9.69	4.91	1.073
GALI_vx	MII_f	0.784 (±0.012)	-0.106 (±0.163)	0.28	730	11.45	5.75	0.683
GALI_vy	MII_f	0.736 (±0.011)	+3.045 (±0.191)	0.12	719	14.99	7.50	0.480
CIES	MI	0.698 (±0.049)	+2.442 (±0.427)	0.21	772	11.60	8.45	9.46
TORA	MII_f	0.666 (±0.264)	+0.494 (±0.124)	0.20	315	3.43	1.72	0.184
VIGO_vx	MII_f	0.925 (±0.053)	-0.793 (±0.286)	0.73	162	2.78	1.43	4.91
VIGO_vx	MII_f	-0.072 (±0.039)	-0.102 (±0.230)	0.01	176	4.42	4.12	3.991

MI: Modelo I
MI_f: Modelo I filtrado
MII: Modelo II
MII_f: Modelo II

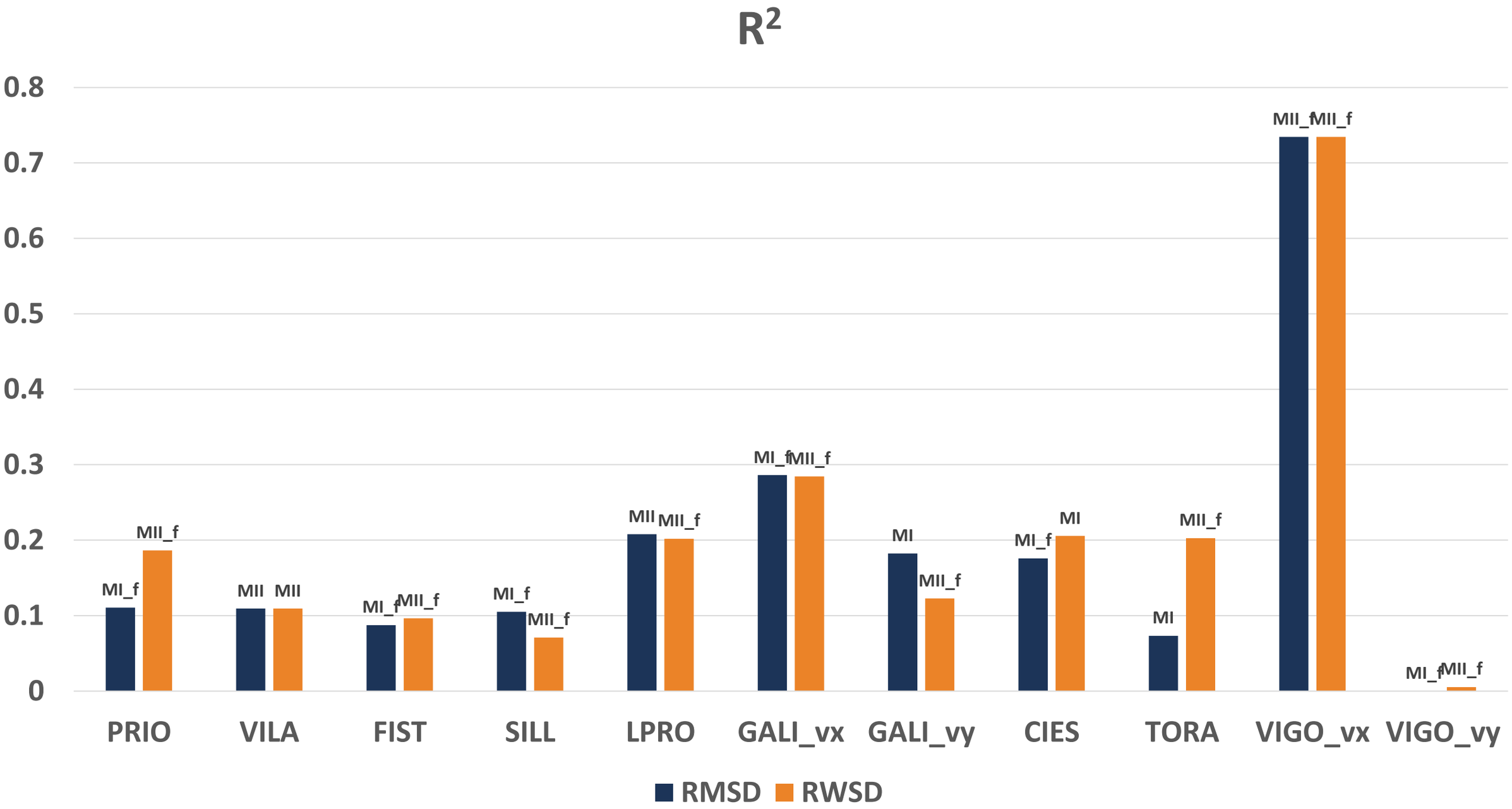
RESULTS



Model type II regression for lagrangian validation of HF Radar velocities in the NW Iberian Peninsula

Adrián Martínez-Fernández et al.

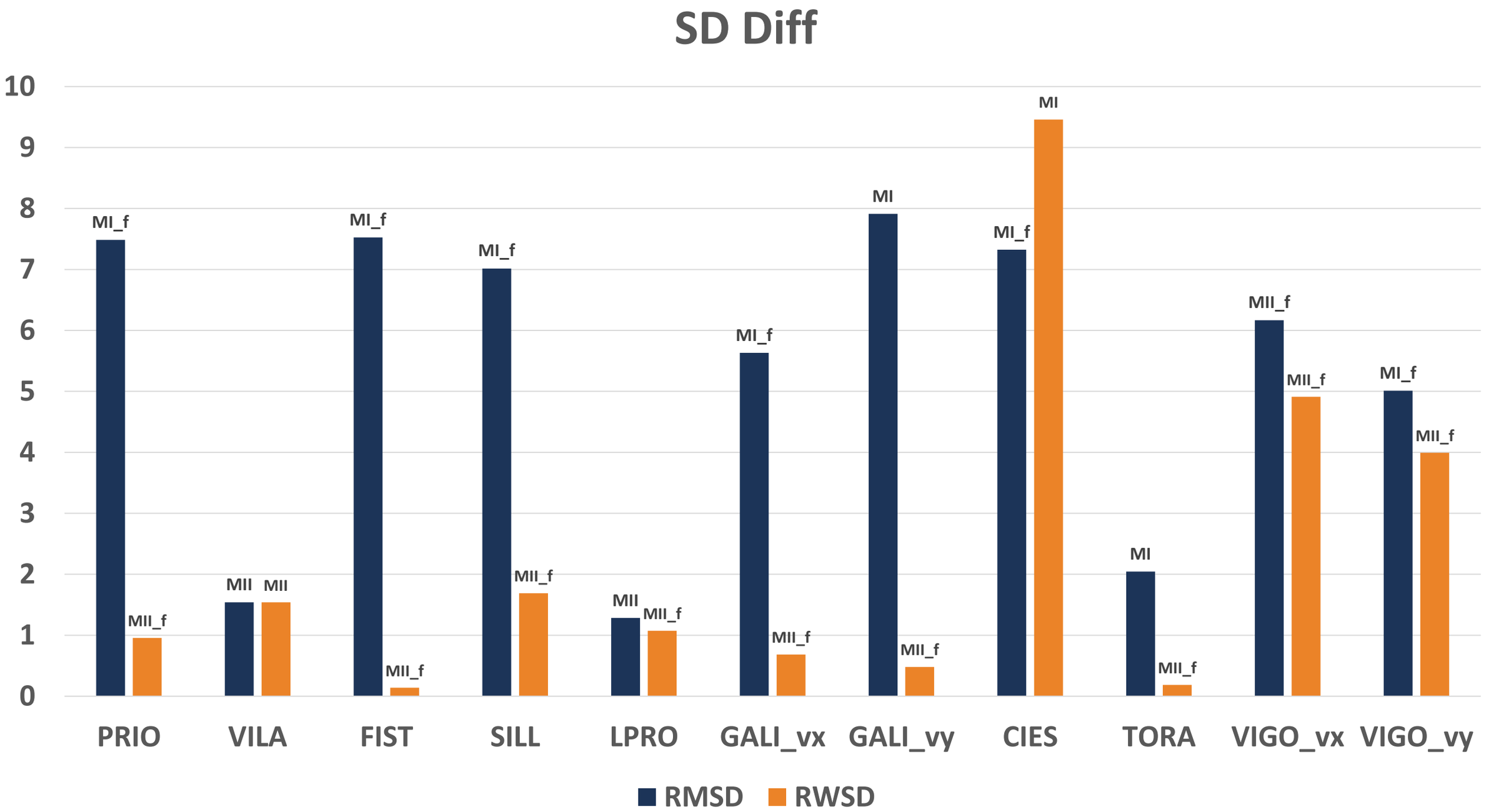
RESULTS



Model type II regression for lagrangian validation of HF Radar velocities in the NW Iberian Peninsula

Adrián Martínez-Fernández et al.

RESULTS



Model type II regression for lagrangian validation of HF Radar velocities in the NW Iberian Peninsula

Adrián Martínez-Fernández et al.

CONCLUSIONS

1

Regression model type II is proposed to perform the lagrangrian validations of HFR since both radar and buoys contains errors and regression model type I only contains errors in the predicted variable.

2

A new metric, RWSD, has been proposed to asses the comparisons of the skill of the linear models used in lagrangian validation of HF Radars. This new metric uses weights to assure a SD of the estimated data similar to the SD of the original data.

3

Regression model type II showed better results than regression model type I: Less difference between $SD(\hat{Y})$ and $SD(Y)$ and higher values of $Slope$

CONCLUSIONS

1 Regression model type II is proposed to perform the lagrangrian validations of HFR since both radar and buoys contains errors and regression model type I only contains errors in the predicted variable.

2 **A new metric, RWSD, has been proposed to asses the comparisons of the skill of the linear models used in lagrangian validation of HF Radars. This new metric uses weights to assure a SD of the estimated data similar to the SD of the original data.**

3 Regression model type II showed better results than regression model type I: Less difference between $SD(\hat{Y})$ and $SD(Y)$ and higher values of $Slope$

CONCLUSIONS

1 Regression model type II is proposed to perform the lagrangrian validations of HFR since both radar and bouys contains errors and regression model type I only contains errors in the predicted variable.

2 A new metric, RWSD, has been proposed to asses the comparisons of the skill of the linear models used in lagrangian validation of HF Radars. This new metric uses weights to assure a SD of the estimated data similar to the SD of the original data.

3 **Regression model type II showed better results than regression model type I: Less difference between $SD(\hat{Y})$ and $SD(Y)$ and higher values of $Slope$**

Thank you so much for your attention

Model type II regression for lagrangian validation of HF Radar velocities in the NW Iberian Peninsula

Adrián Martínez-Fernández, Waldo Redondo-Caride, Fernando Alonso-Pérez, Silvia Piedracoba, Pablo Lorente, Silvia Allen-Perkins, Pedro Montero, Garbiñe Ayensa, Silvia Torres-López, Adrián Fernández-Baladrón, Ramiro A. Varela, Antón Velo, Miguel Gilcoto