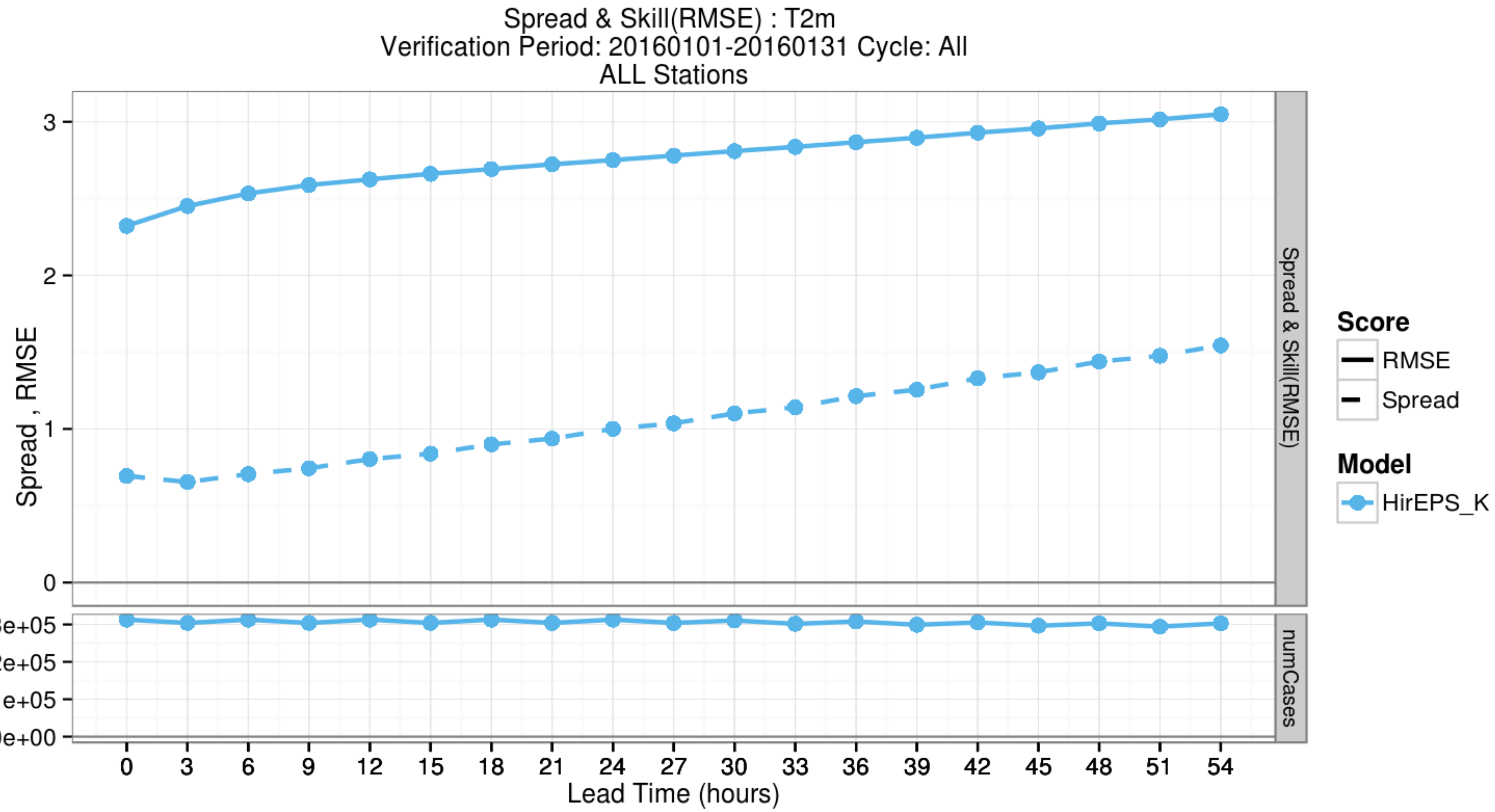
Issues in the verification of EPS

Åke Johansson
SMHI

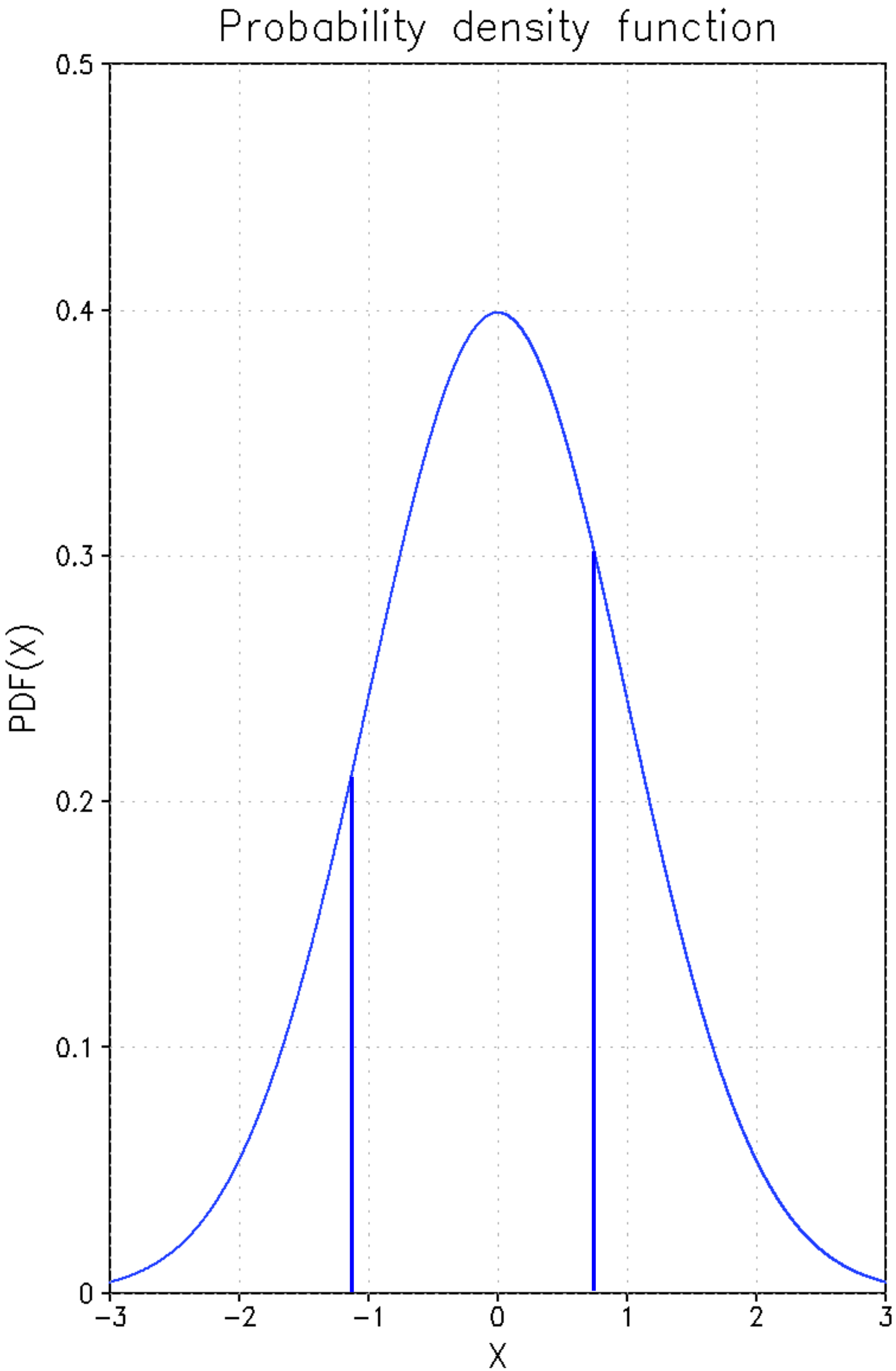
Spread-Skill T2M from HIRLAM_K for JAN 2016



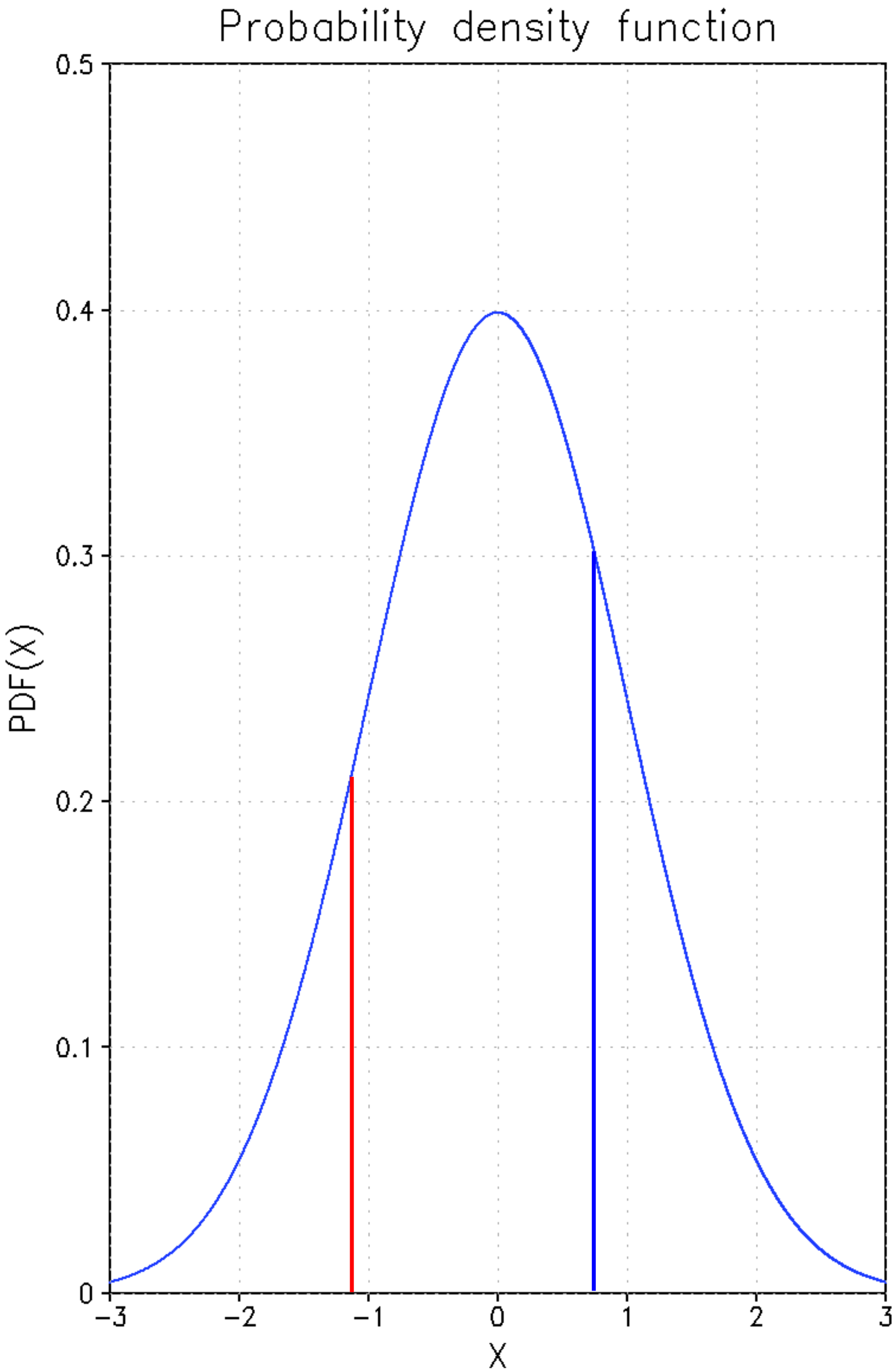
CU

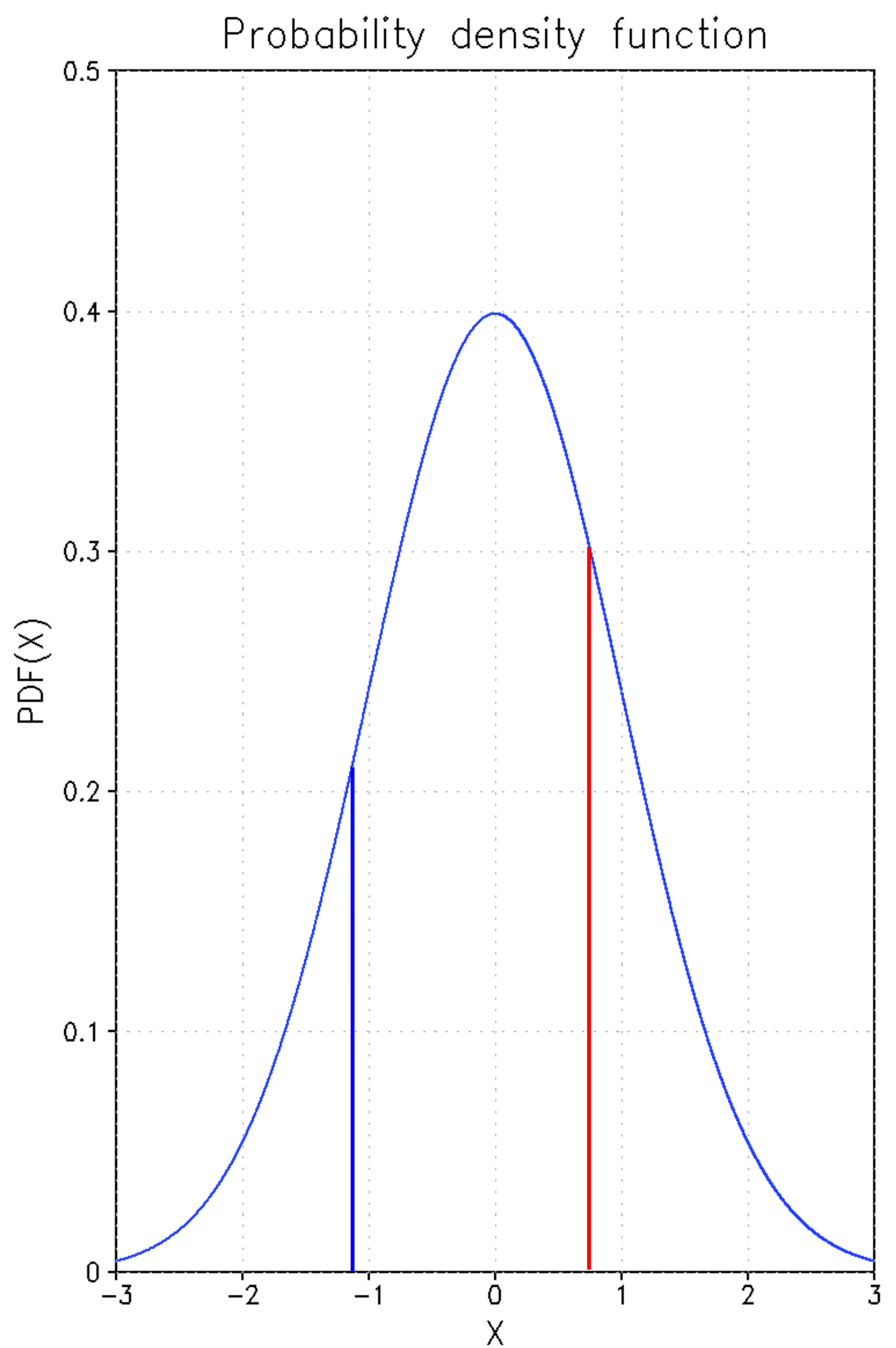
The Commonly Used

Spread-Skill Relationship



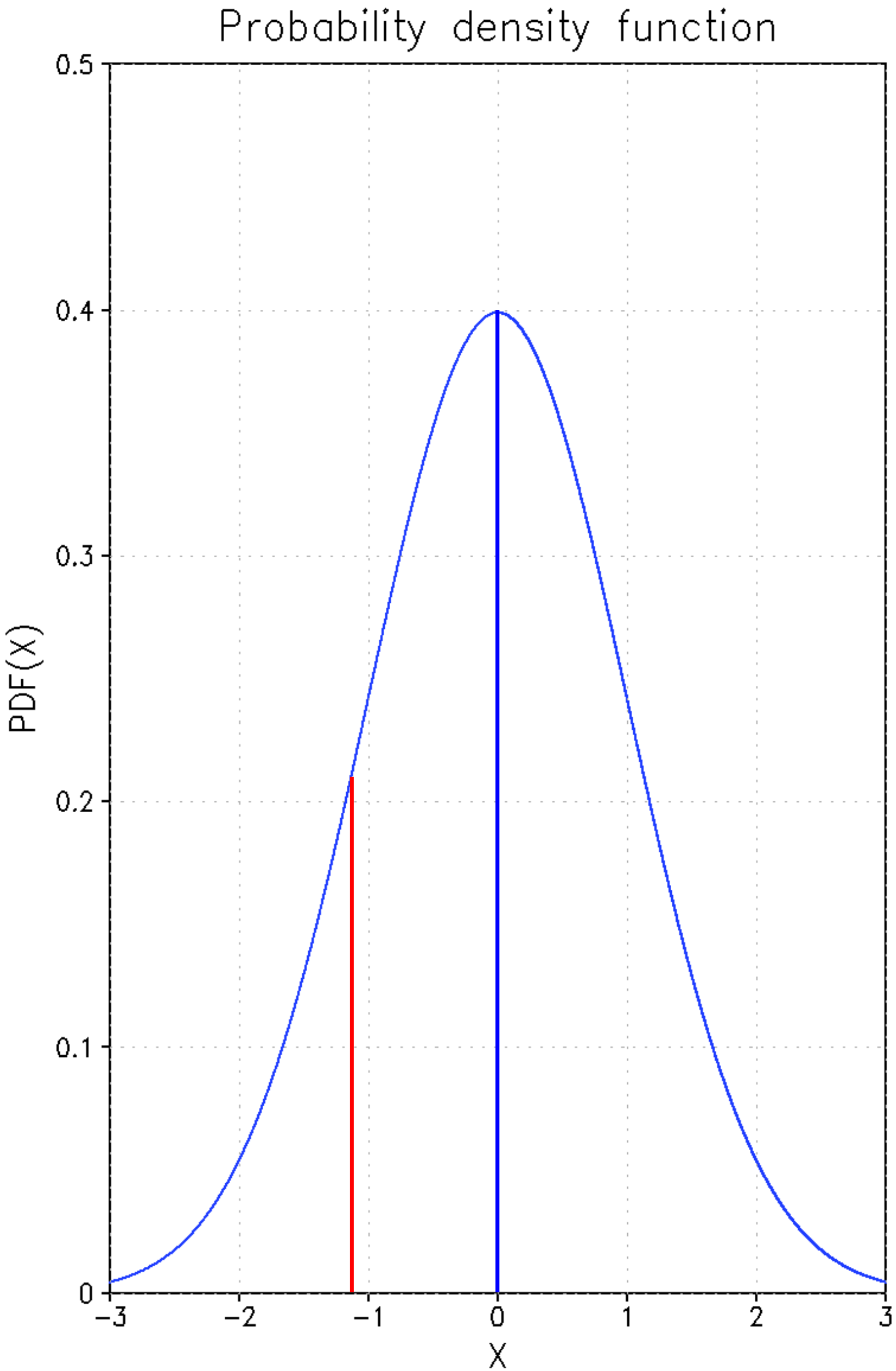
Probability density function
0.5
0.4
0.3
0.2
0.1
0
PDF(X)
−3
−2
−1
0
1
2
3
X


Probability density function
0.5
0.4
0.3
0.2
0.1
0
PDF(X)
−3
−2
−1
0
1
2
3
X


Probability density function
PDF(X)
X
0.5
0.4
0.3
0.2
0.1
0
−3
−2
−1
0
1
2
3

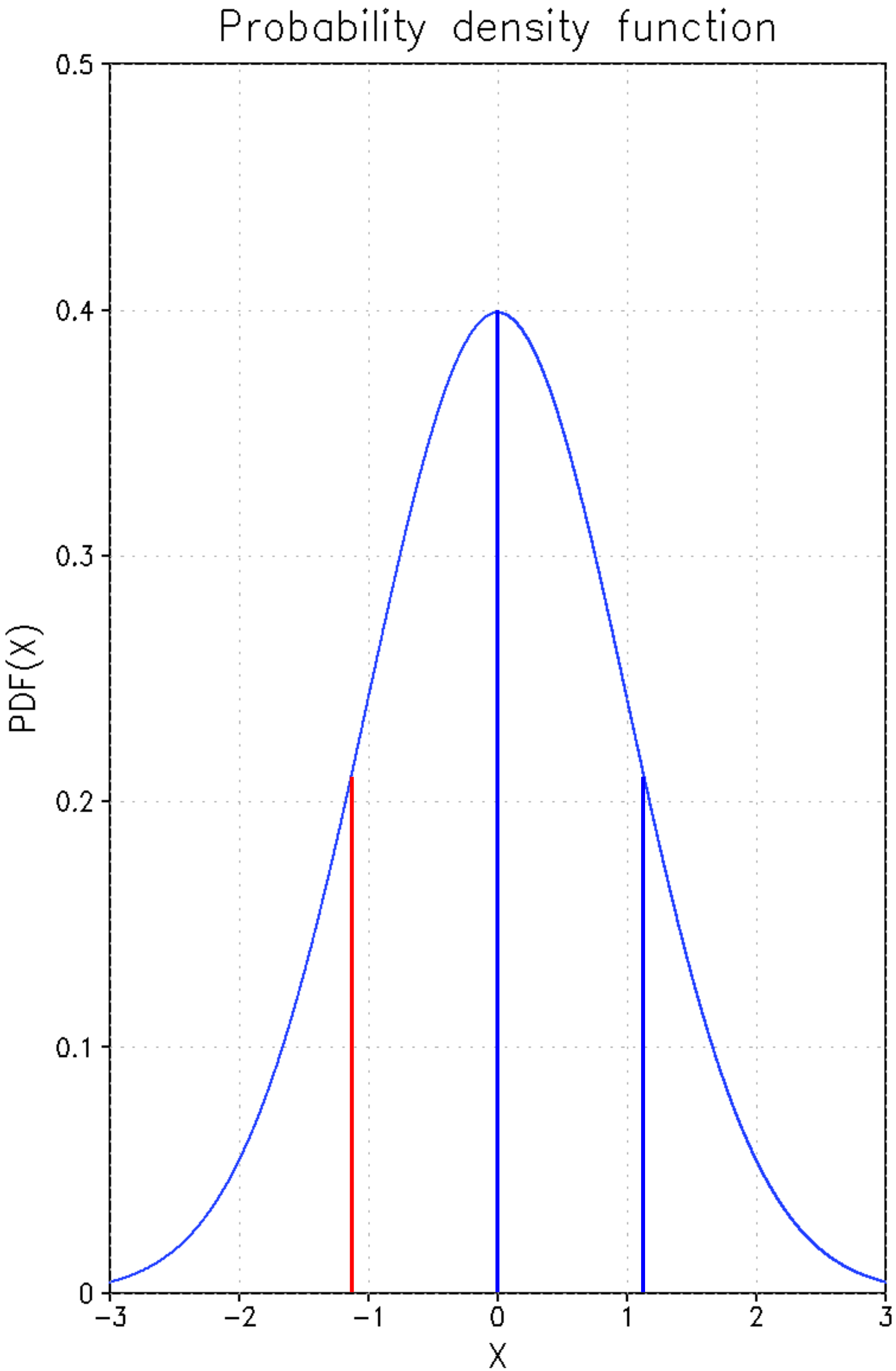
$$\boxed{E\left[RMSE_M\right]}$$

$$\sqrt{\overrightarrow{\overline{\left[o-\vec{f}\,\right]^2}}}$$



$$\boxed{E\left[RMSE_M\right]}$$

$$\sqrt{\overline{\overrightarrow{\left[o-\overrightarrow{f}\right]^2}}}$$



$$\boxed{E\left[Std.\,Dev._E\right]}$$

$$\sqrt{\overline{\overrightarrow{\left[f-\overrightarrow{f}\right]^2}}}$$

CU SKILL – SPREAD RELATION

$$E\left[\mathrm{RMSE}_{\mathrm{M}}\right] = E\left[\mathrm{Std.\,Dev.}_{\mathrm{E}}\right]$$

CU SKILL – SPREAD RELATION

$$E\left[RMSE_{M}\right] = E\left[Std.\,Dev._{E}\right]$$

$$E\left[MSE_{M}\right] = E\left[VAR_{E}\right]$$

CU Skill-Spread relation

is only valid for an

Ideal ensemble

4 reasons why

it is inappropriate for

Real World EPS

CU SKILL – SPREAD RELATION

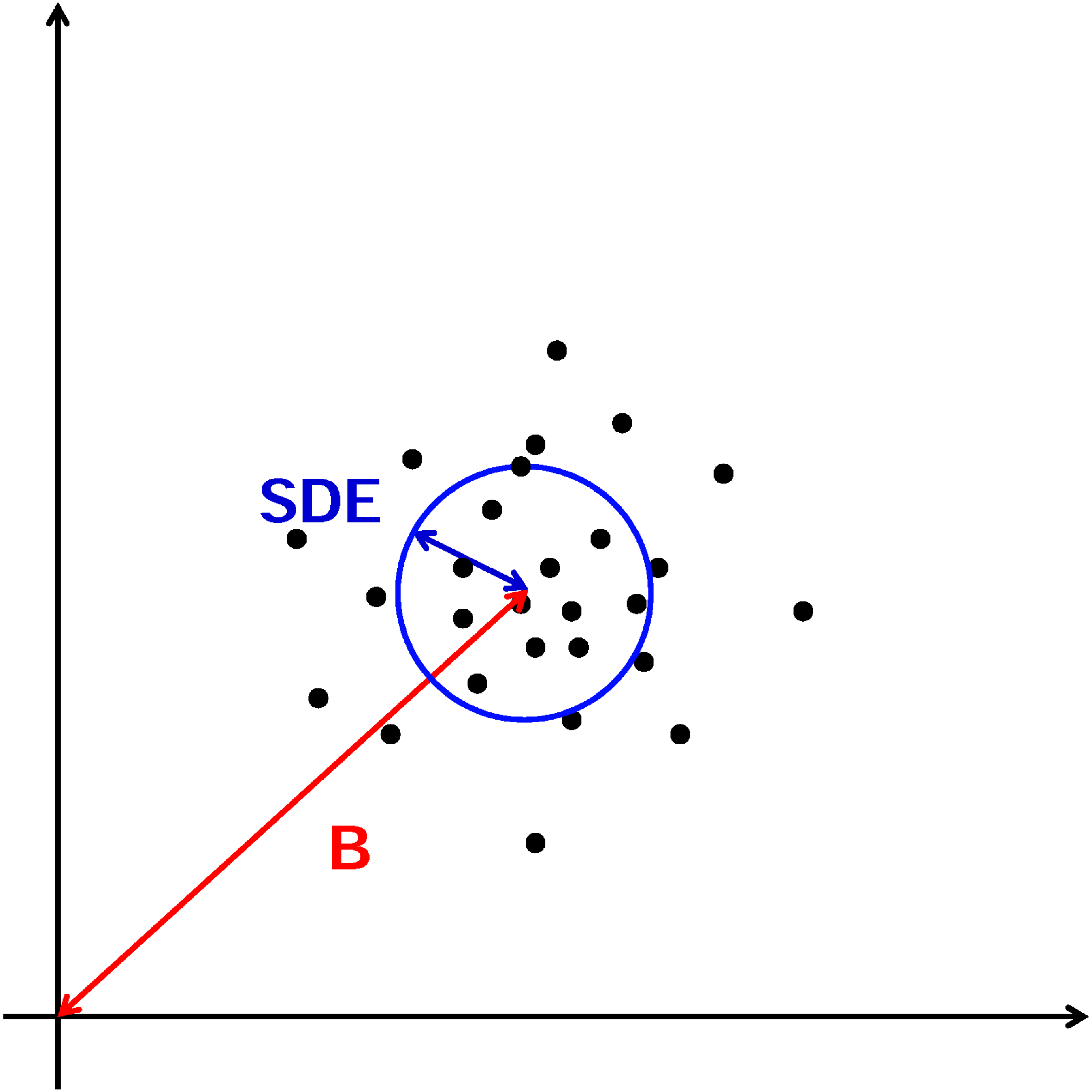
$$\boxed{E\left[MSE_M\right]} = E\left[VARE\;\right]$$

1

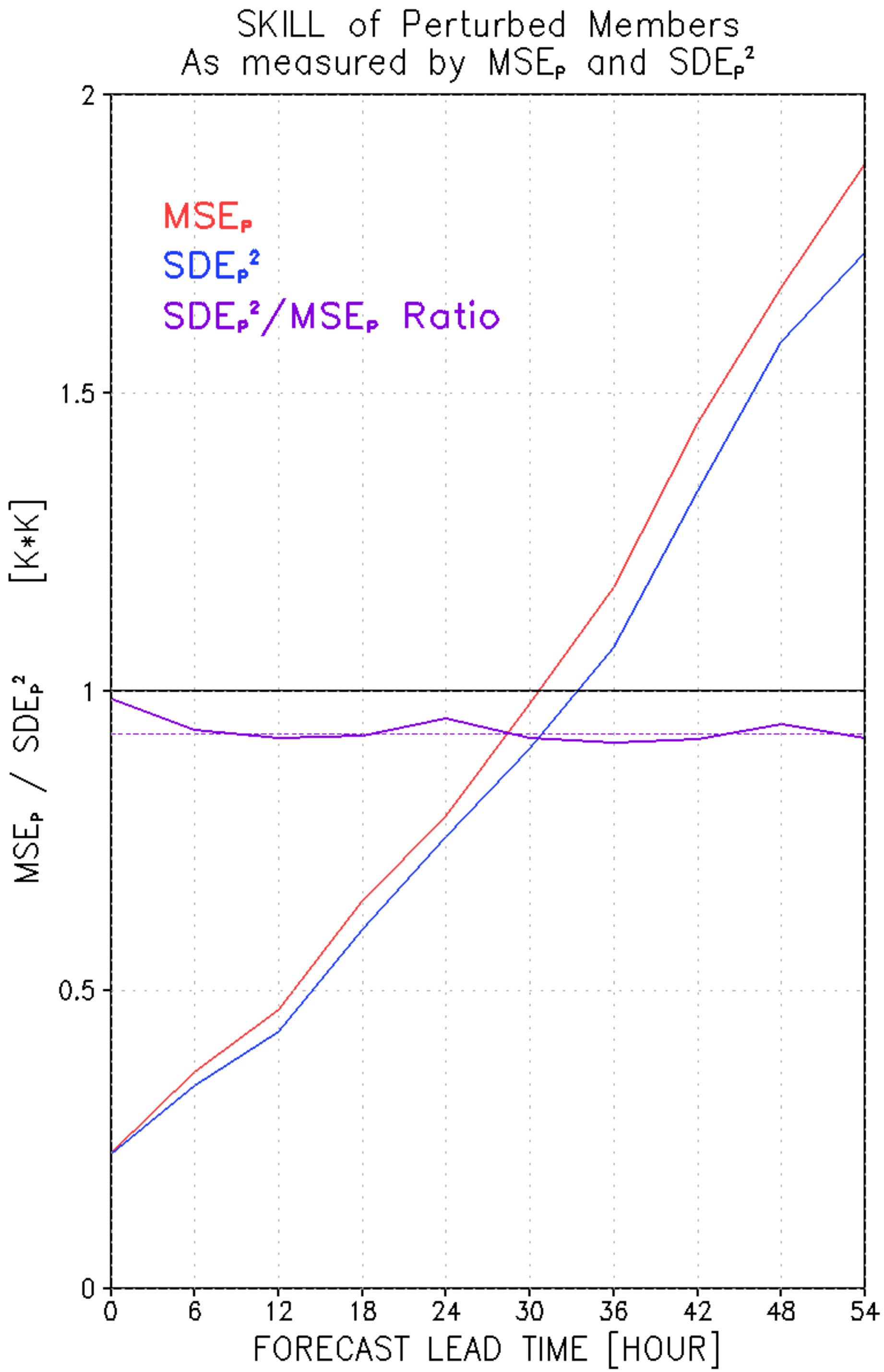
Real World Models

usually have a

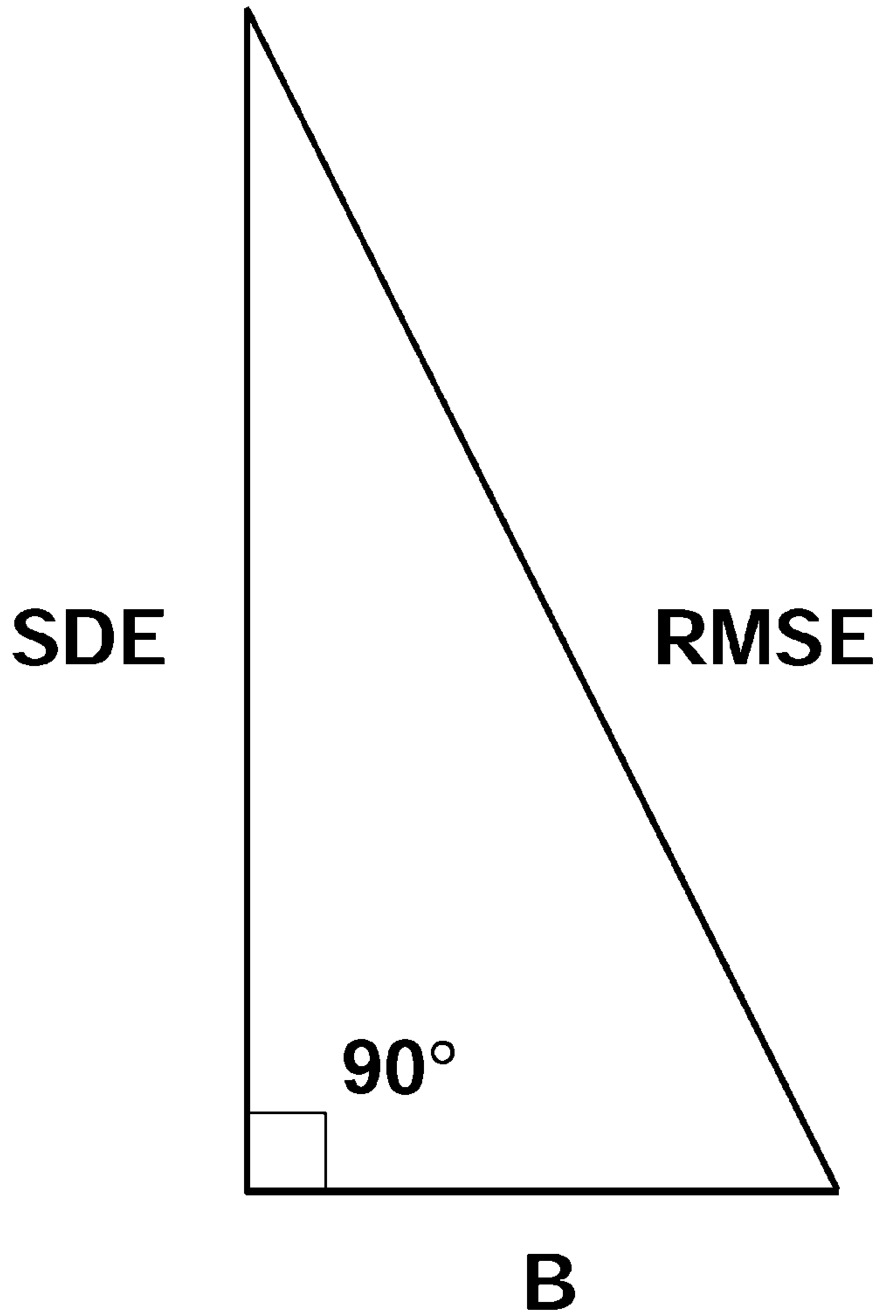
BIAS



SDE
B



SKILL of Perturbed Members
As measured by MSEP and SDEP²
MSEP
SDEP²
SDEP²/MSEP Ratio
MSEP / SDEP² [K*K]
FORECAST LEAD TIME [HOUR]
0
6
12
18
24
30
36
42
48
54
0
0.5
1
1.5
2


SDE
RMSE
90°
B

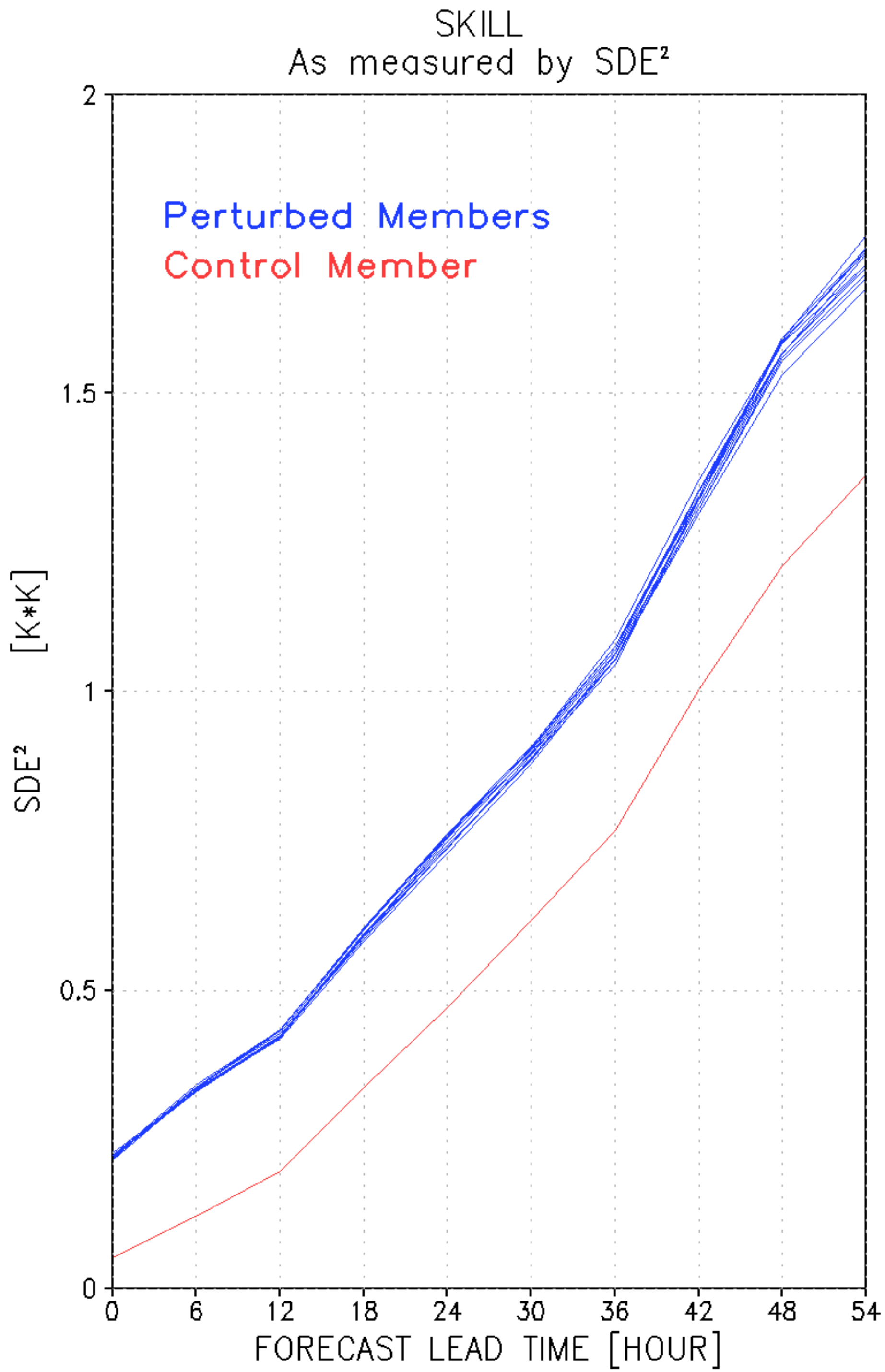
$$E\left[MSE_{M}\right] = E\left[VARE\right]$$

$$\Downarrow$$

$$E\left[SDE_{M}^{2}\right] = E\left[VARE\right]$$

2

Real World Models usually have members that are not statistically equal


SKILL
As measured by SDE²
Perturbed Members
Control Member
SDE² [K*K]
FORECAST LEAD TIME [HOUR]
0
0.5
1
1.5
2
0
6
12
18
24
30
36
42
48
54

$$E\left[SDE_M{}^2\right] = E\left[VARE\right]$$

$$\Downarrow$$

$$E\left[SDE_{MP}{}^2\right] = E\left[VARE_P\right]$$

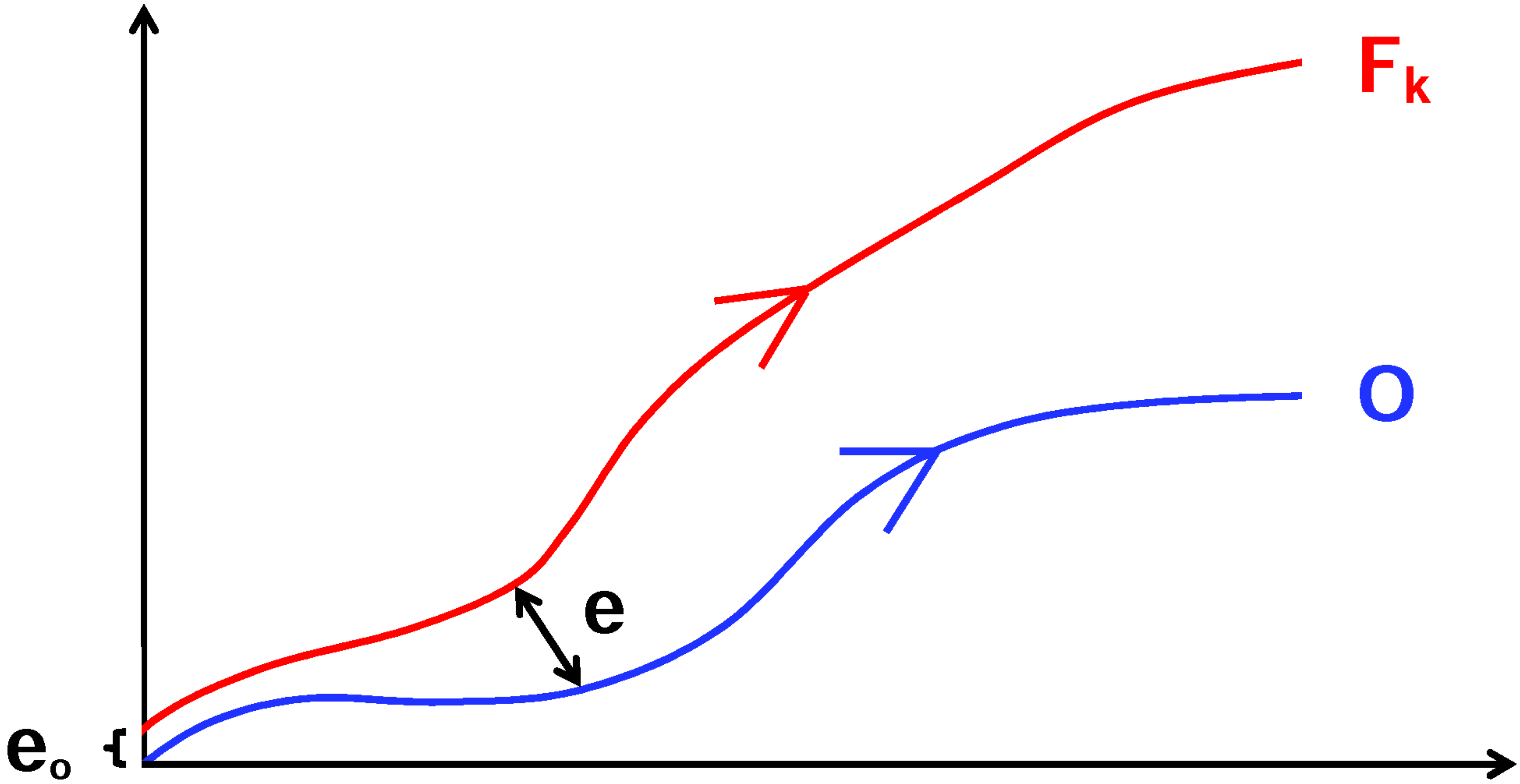
3

Using U-Statistics for estimating

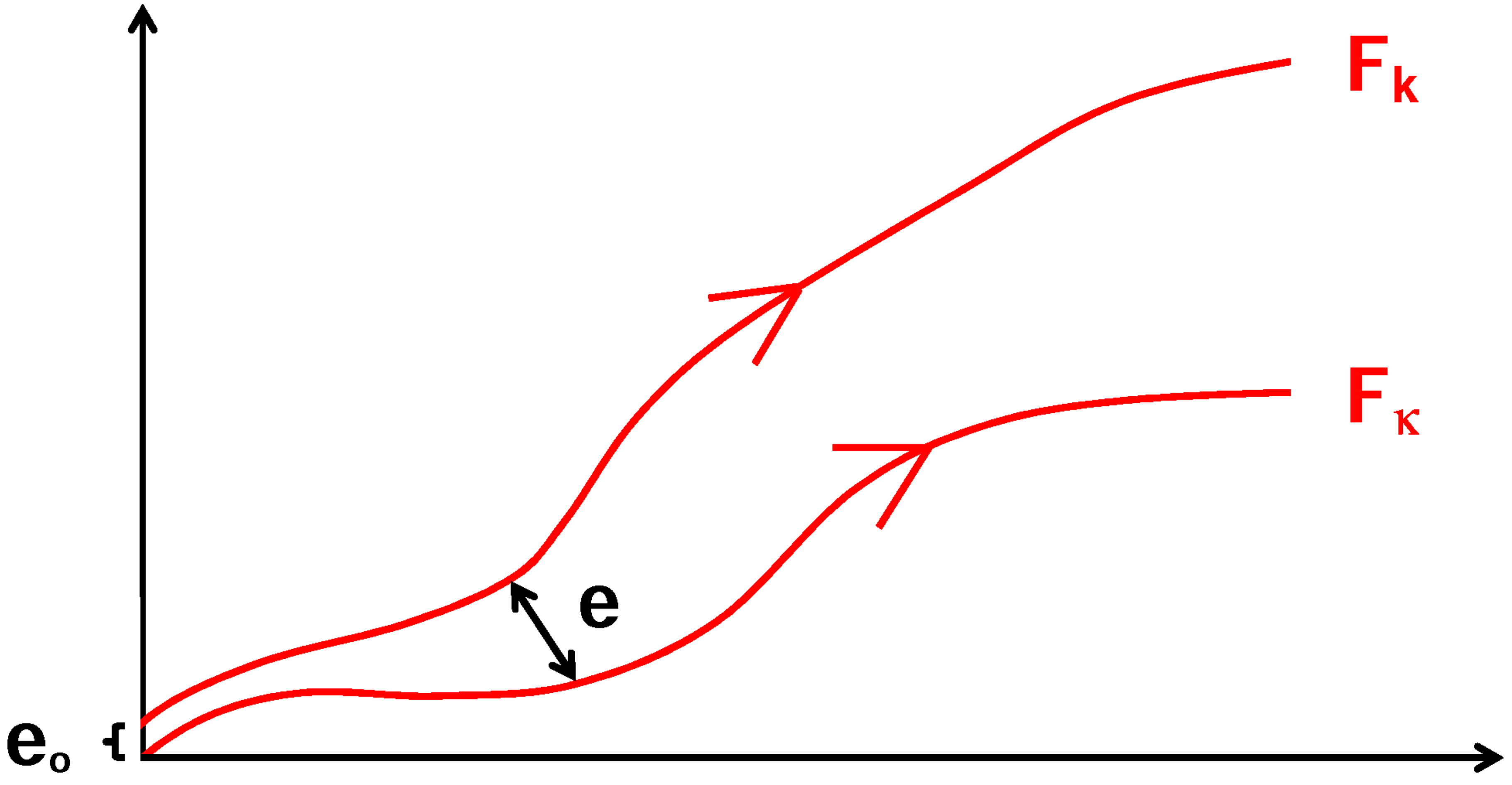
Skill and Spread

gives more realistic estimates

SKILL



SPREAD



SPREAD

$$SPRE_P = \sqrt{\frac{2}{K \cdot (K-1)} \sum_{k=1}^{K-1} \sum_{\kappa=k+1}^{K} \left[f(t_o, t_f, k) - f(t_o, t_f, \kappa) \right]^2}$$

SPREAD

$$\mathrm{SPRE}_P = \sqrt{\frac{2}{K \cdot (K-1)} \sum_{k=1}^{K-1} \sum_{\kappa=k+1}^{K} \left[f(t_o, t_f, k) - f(t_o, t_f, \kappa) \right]^2}$$

U-statistic

SPREAD

$$\mathrm{SPRE}_P = \sqrt{\frac{2}{K\cdot(K-1)}\sum_{k=1}^{K-1}\sum_{\kappa=k+1}^{K}\left[f(t_o, t_f, k) - f(t_o, t_f, \kappa)\right]^2}$$

U-statistic

A U-statistic is defined as the average –

across all combinatorial selections of the given size from the full set of observations –

of the basic estimator applied to the sub-samples.

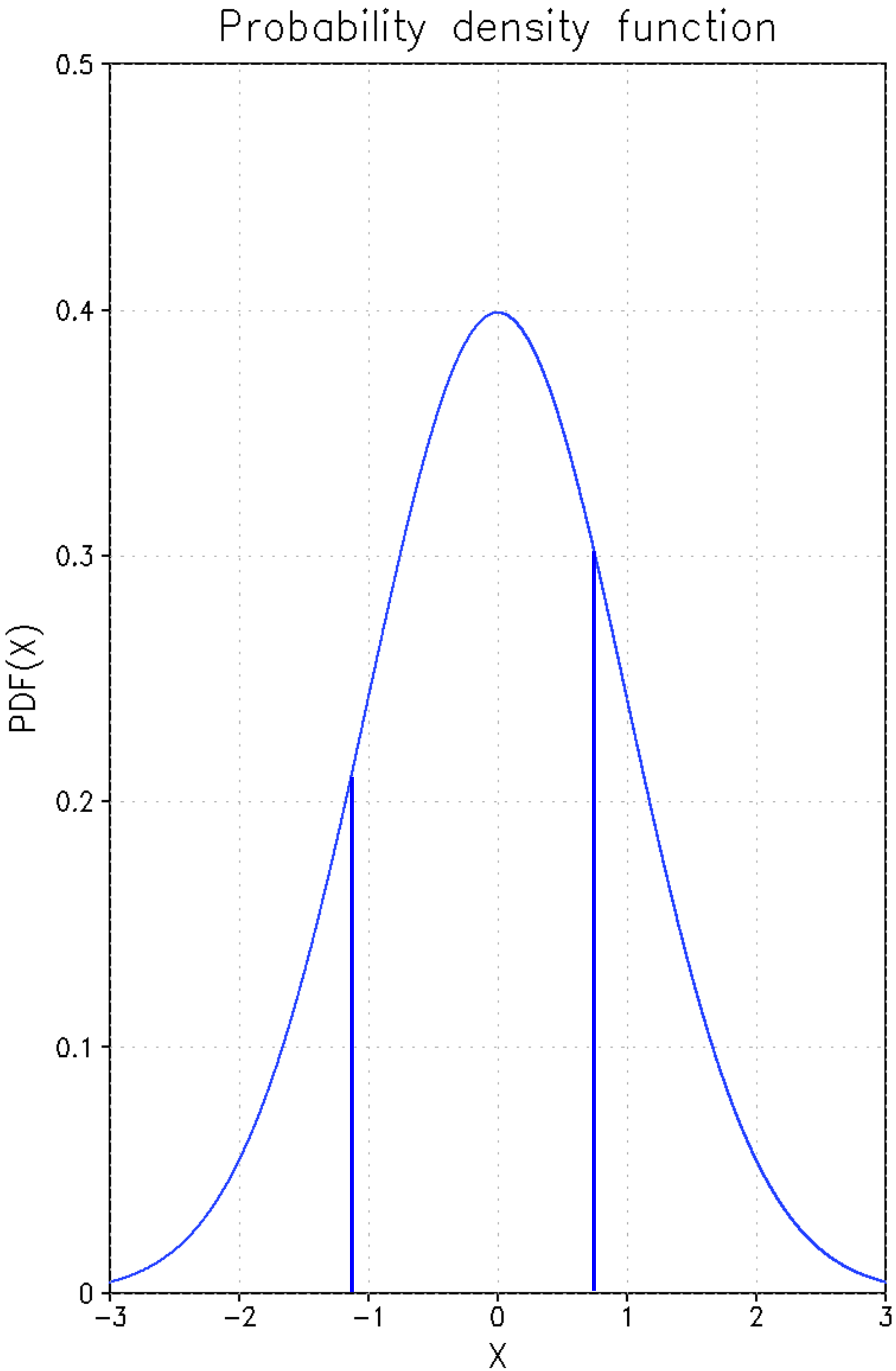
U-statistics, where the letter U stands for unbiased, arise naturally in producing **minimum-variance unbiased estimators**.

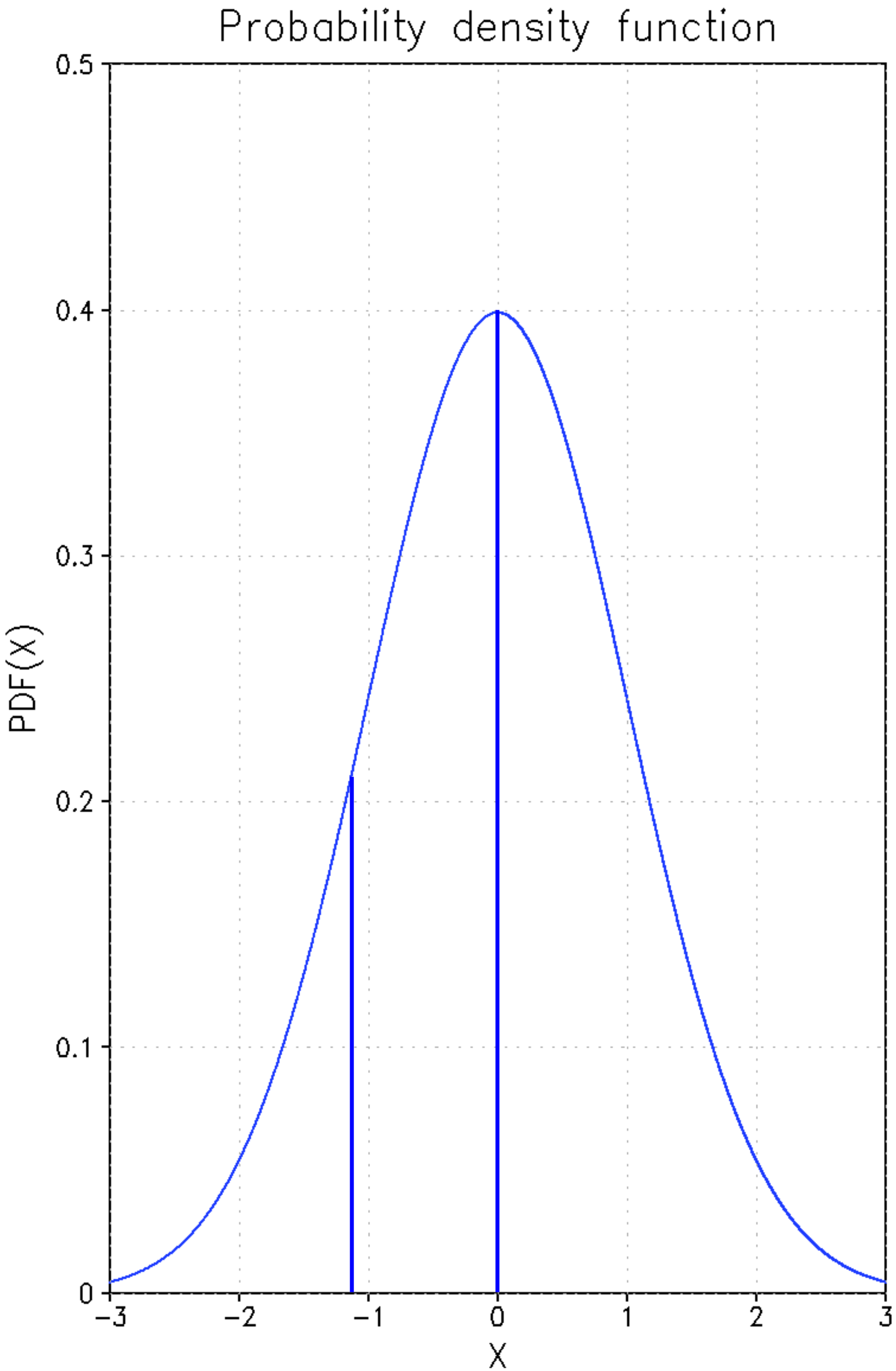
SPREAD

$$\mathrm{SPRE}_P = \overline{\frac{2}{K\cdot(K-1)}\sum_{k=1}^{K-1}\sum_{\kappa=k+1}^{K}\left[f(t_o, t_f, k) - f(t_o, t_f, \kappa)\right]^2}$$

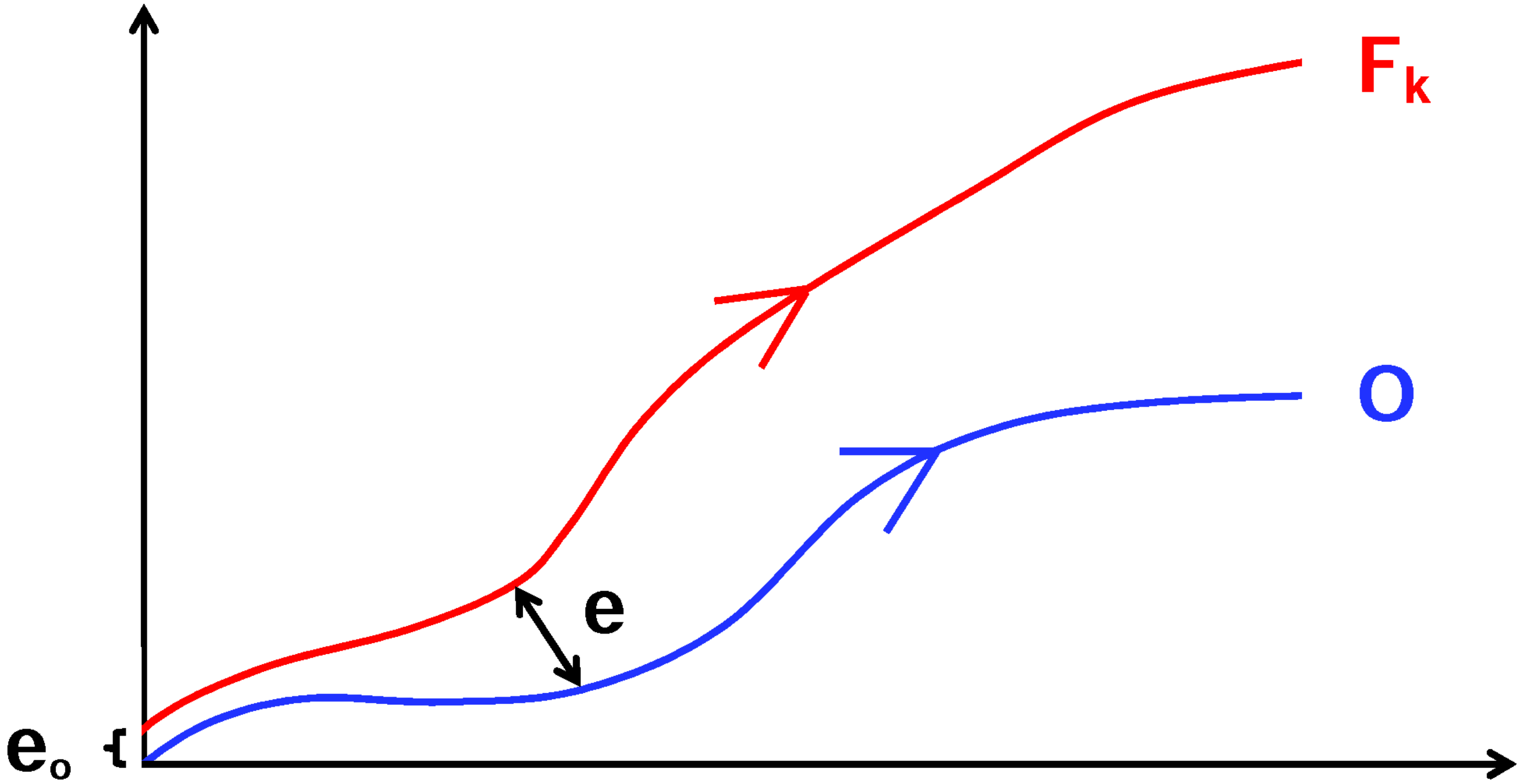
$$\mathrm{VARE}_P = \overline{\overrightarrow{\left[f(t_o, t_f, k) - \overrightarrow{f}(t_o, t_f)\right]^2}}$$

$$\mathrm{SPRE}_P = 2\frac{K}{K-1}\mathrm{VARE}_P \quad \Leftrightarrow \quad E\left[\mathrm{SPRE}_P\right] = 2E\left[\mathrm{VARE}_P\right]$$


Probability density function
PDF(X)
X
0.5
0.4
0.3
0.2
0.1
0
−3
−2
−1
0
1
2
3


Probability density function
0.5
0.4
0.3
0.2
0.1
0
PDF(X)
−3
−2
−1
0
1
2
3
X

SKILL



SKILL

$$SDE_P^{\,2} = MSE_P - ME_P^{\,2}$$

U-statistic

U.UI SKILL – SPREAD CONDITION

$$E\left[SDE_P^{\,2}\right] = E\left[SPRE_P\right]$$

U.UI SKILL – SPREAD CONDITION

$$E\left[SDE_P^{\ 2}\right] = E\left[SPRE_P\right]$$

$$\frac{1}{2}E\left[SDE_P^{\ 2}\right] = E\left[VARE_P\right]$$

CU → U.UI

$$E\left[MSE_M\right] = E\left[VARE\right]$$

$$\Downarrow$$

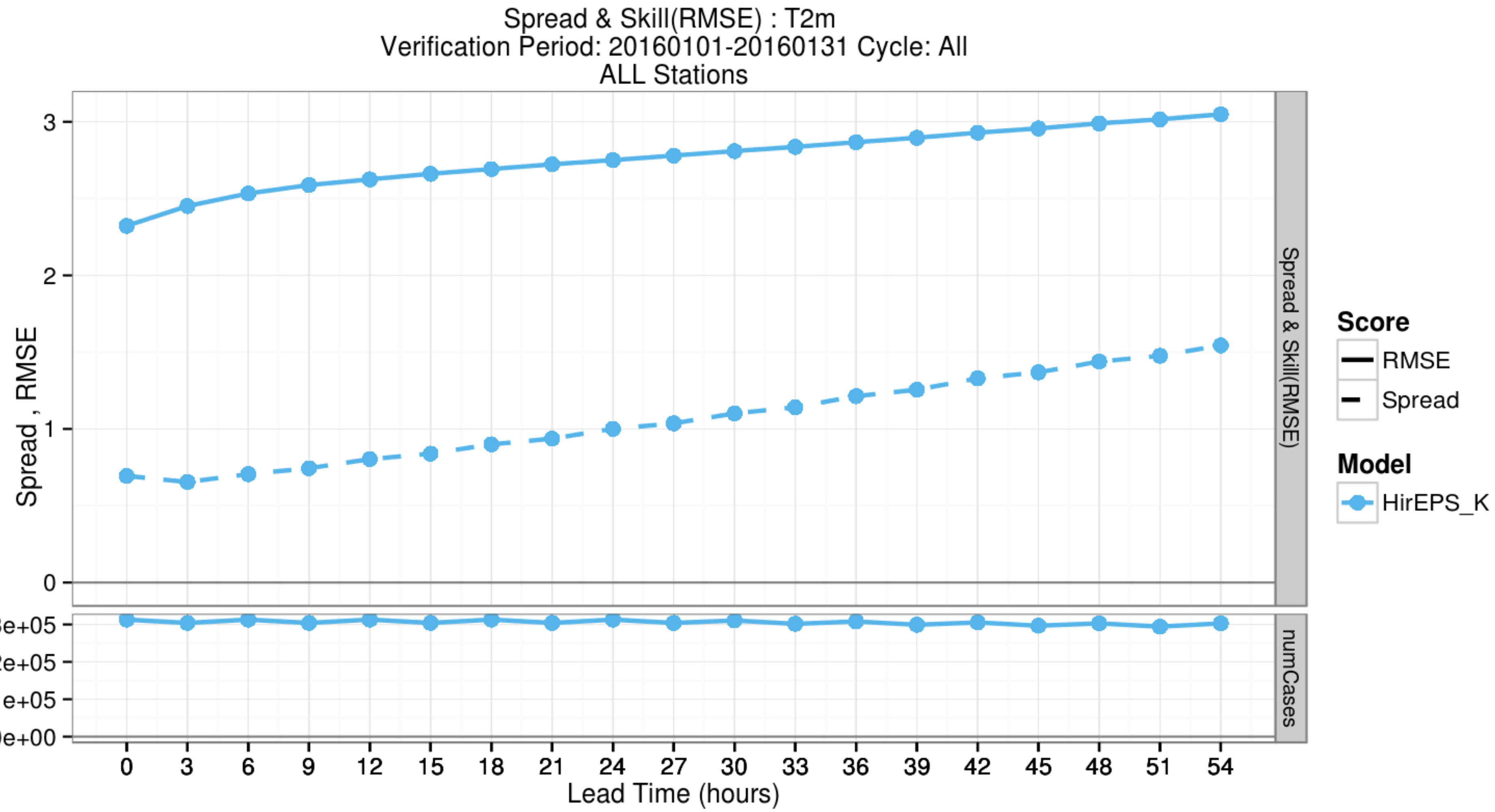
$$E\left[SDE_M{}^2\right] = E\left[VARE\right] \qquad \textbf{U}$$

$$E\left[SDE_{MP}{}^2\right] = E\left[VARE_P\right] \qquad \textbf{I}$$

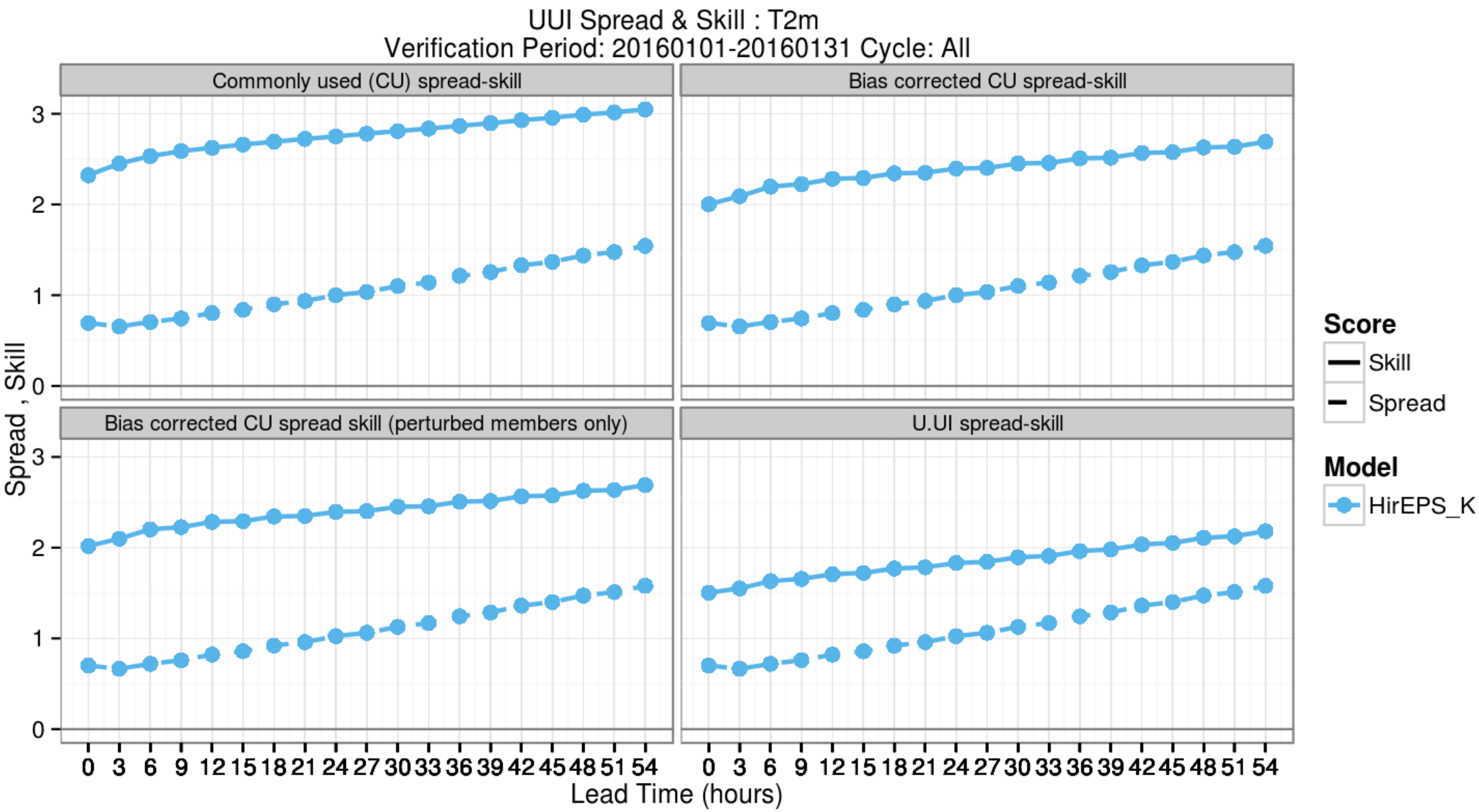
$$\Downarrow$$

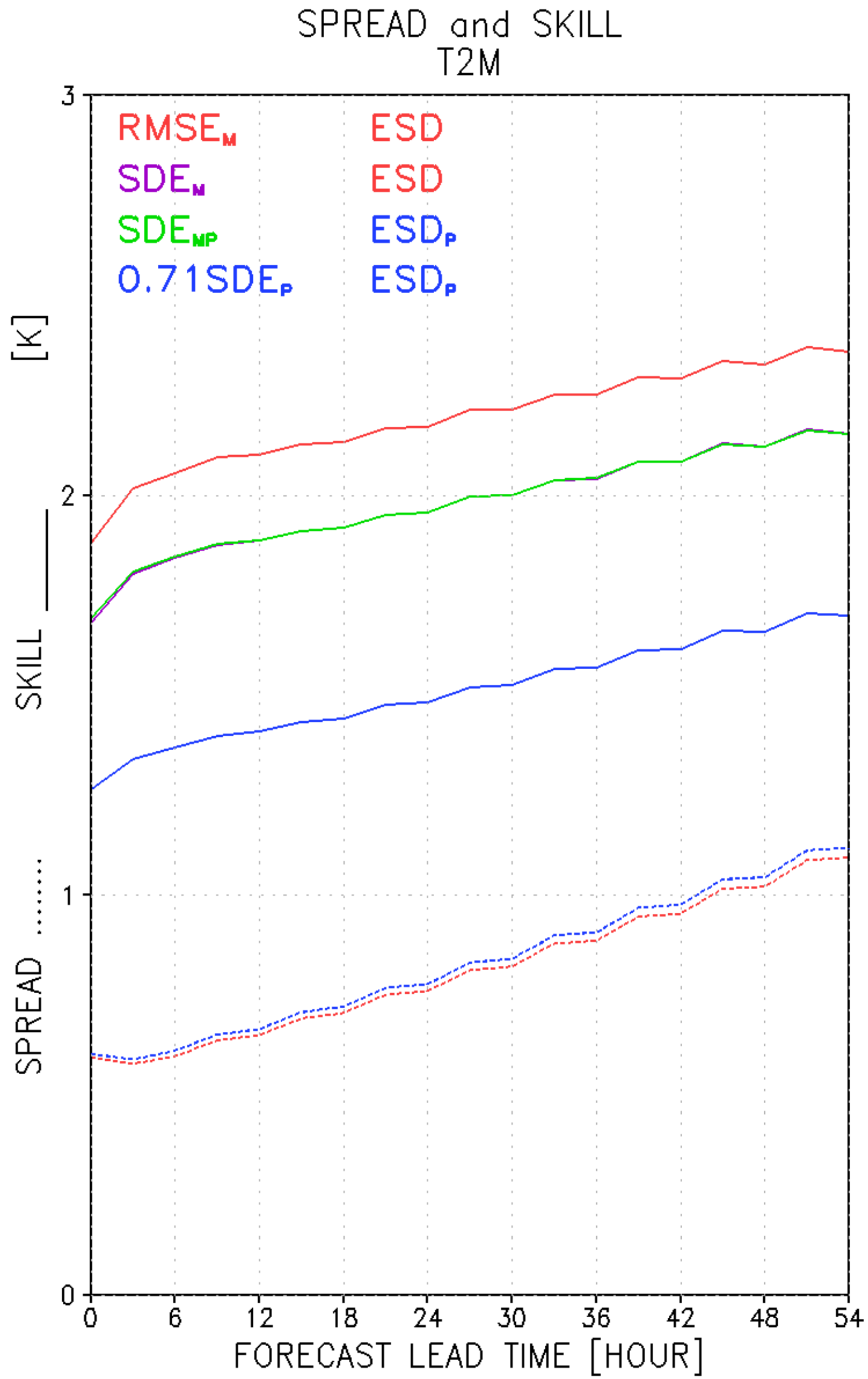
$$\frac{1}{2}E\left[SDE_P{}^2\right] = E\left[VARE_P\right] \qquad \textbf{U}$$

Example for T2M from HIRLAM_K for JAN 2016



Example for T2M from HIRLAM_K for JAN 2016





SPREAD and SKILL
T2M
RMSE_M ESD
SDE_M ESD
SDE_MP ESD_P
0.71SDE_P ESD_P
SPREAD .... SKILL ____ [K]
FORECAST LEAD TIME [HOUR]

Why U-Statistics

Theoretical justification - Spread

$$E[F] = a$$

$$E\left[(F - a)^2\right] = E\left[F'^2\right]$$

Theoretical justification - Spread

$$E\left[F\right] = a$$

$$E\left[\left(F - a\right)^2\right] = E\left[F'^2\right]$$

$$E\left[\left(F_i - F_j\right)^2\right] = E\left[\left(a + F_i' - a - F_j'\right)^2\right] =$$

$$= E\left[\left(F_i' - F_j'\right)^2\right] =$$

$$= 2 \cdot E\left[F'^2\right]$$

Theoretical justification - Spread

$$E\left[(F-a)^2\right] = E\left[F'^2\right]$$

$$E\left[(F_i - F_j)^2\right] = 2 \cdot E\left[F'^2\right]$$

Theoretical justification - Skill

$$E[O] = 0 \quad , \quad E[F] = a$$

$$E\left[(O - a)^2\right] = E\left[O'^2\right] + a^2$$

Theoretical justification - Skill

$$E[O] = 0 \qquad , \qquad E[F] = a$$

$$E\left[(O-a)^2\right] = E\left[O'^2\right] + a^2$$

$$E\left[(F-O)^2\right] = E\left[\left(E(F) + F' - E(O) - O'\right)^2\right] =$$

$$= E\left[(a + F' - O')^2\right] =$$

$$= E\left[F'^2\right] + E\left[O'^2\right] + a^2$$

Theoretical justification - Skill

$$E\left[(O-a)^2\right] = E\left[F'^2\right] + a^2$$

$$E\left[(F-a)^2\right] = 2\cdot\left[E\left[F'^2\right] + \frac{1}{2}a^2\right]$$

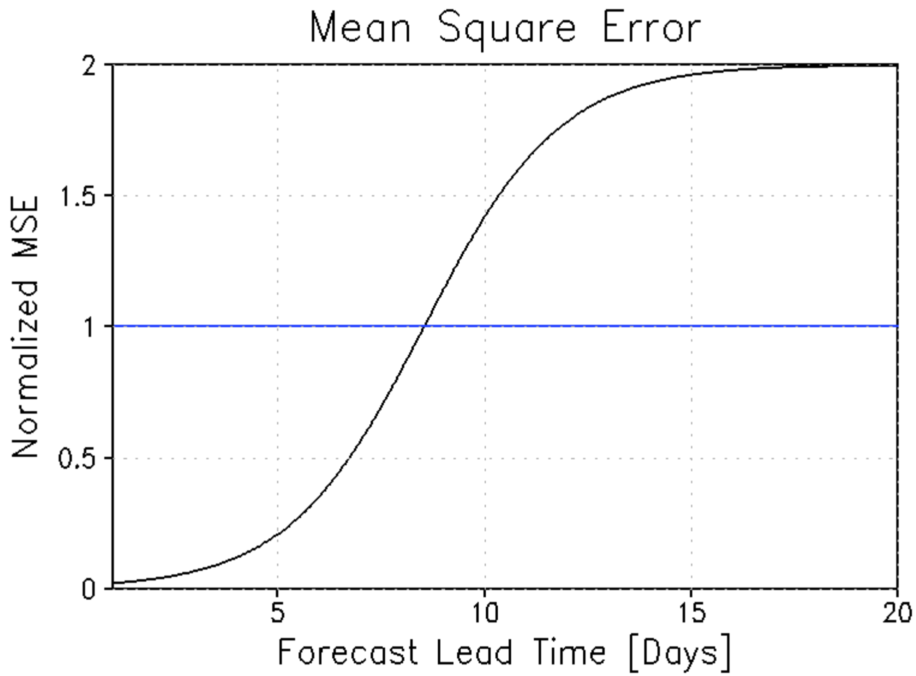
Spread	Skill
$E\left[F'^2\right]$	$E\left[F'^2\right] + a^2$
$2 \cdot E\left[F'^2\right]$	$2 \cdot \left[E\left[F'^2\right] + \frac{1}{2} a^2\right]$

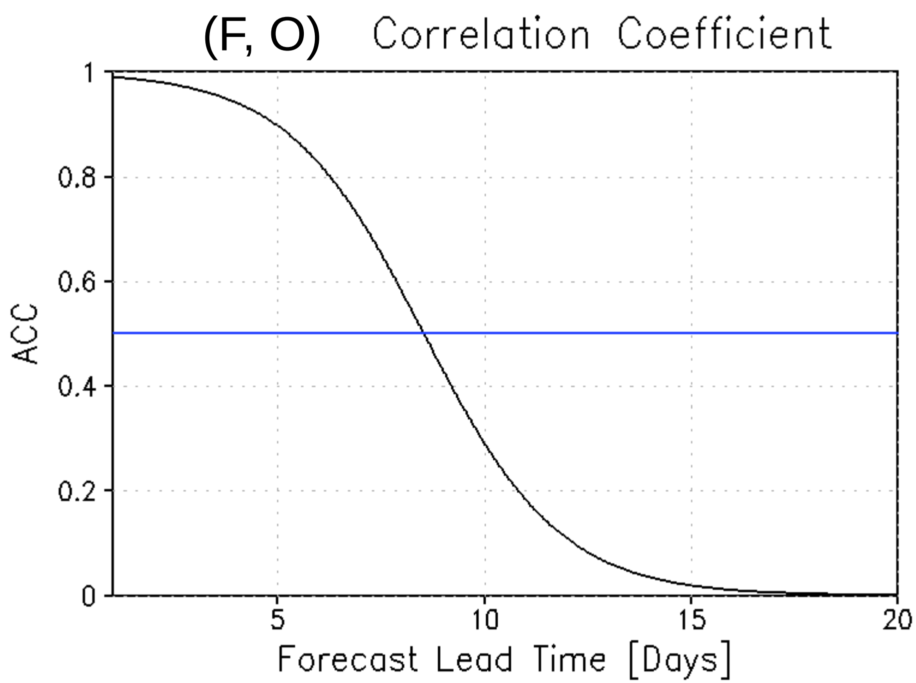
4

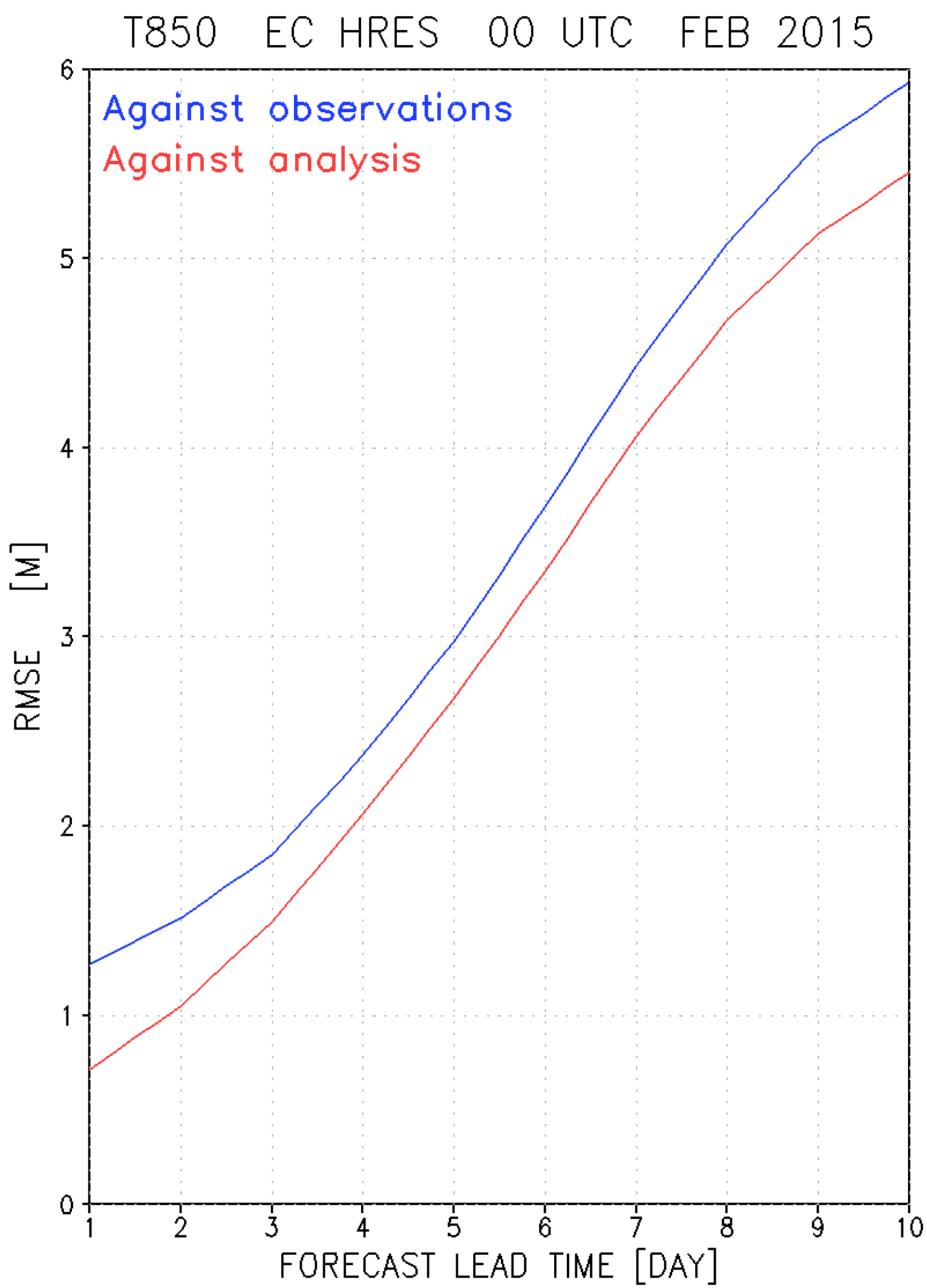
Real World Models

are usually not verified against

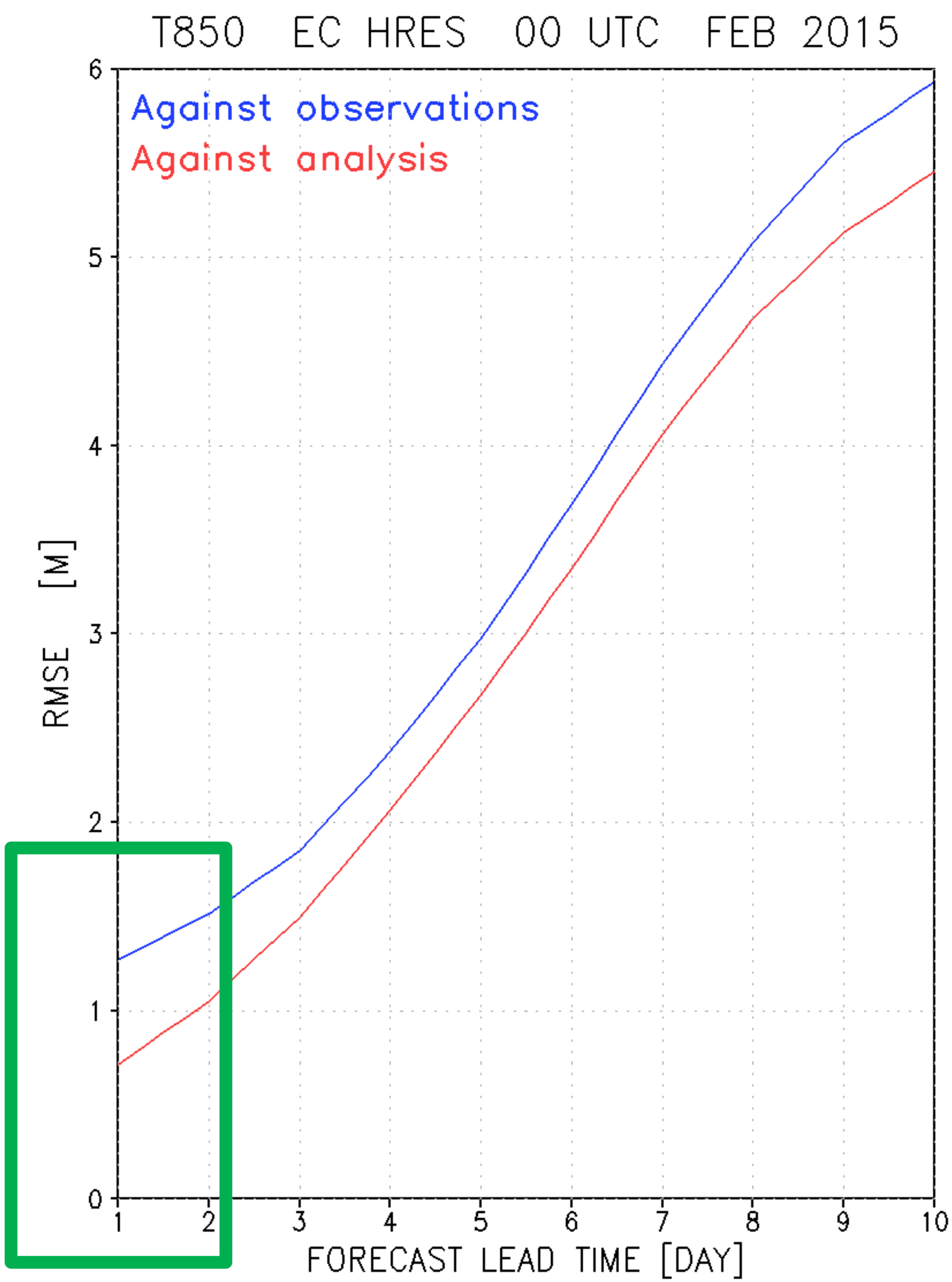
the ”Truth”

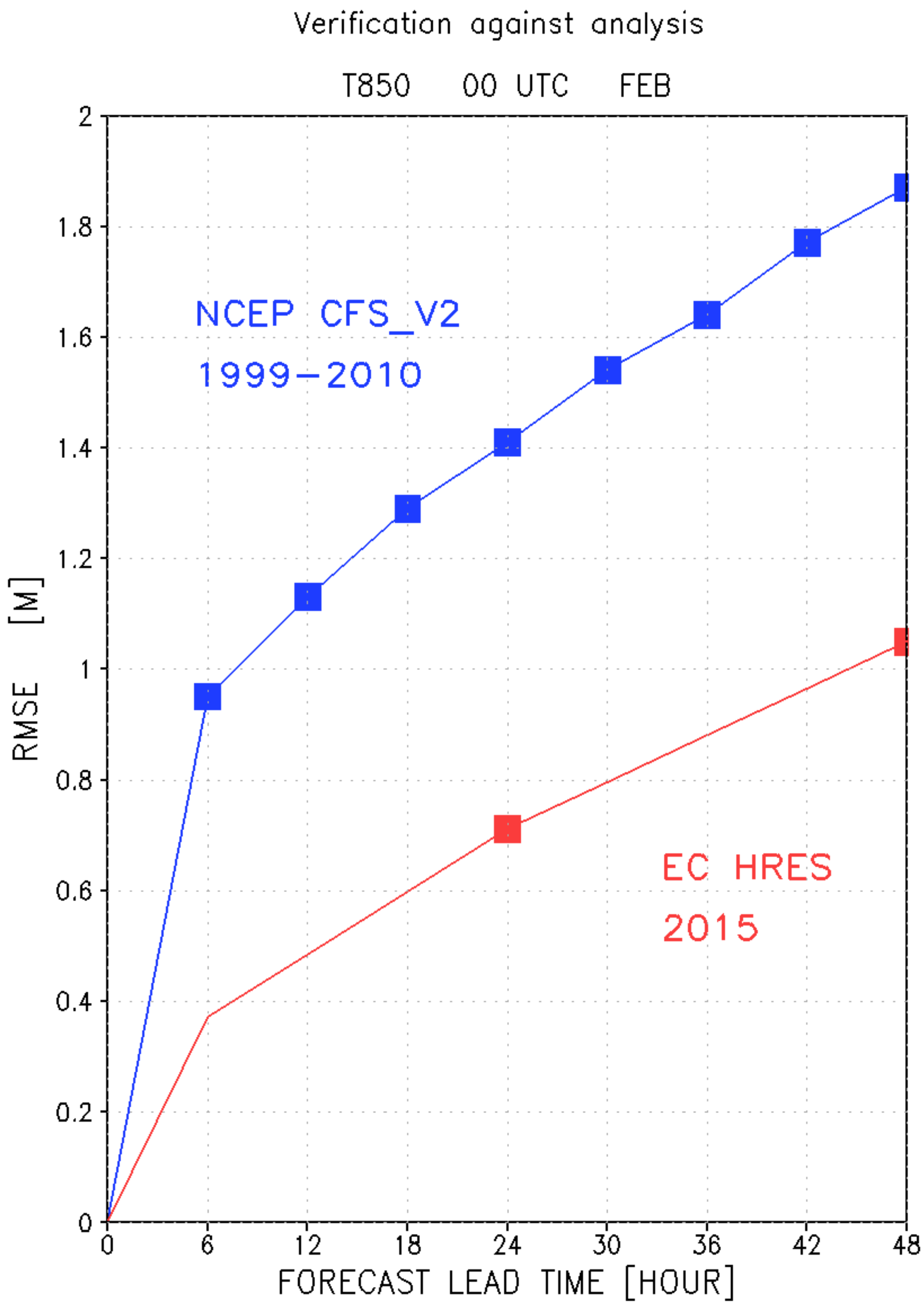

Mean Square Error
Normalized MSE
0
0.5
1
1.5
2
5
10
15
20
Forecast Lead Time [Days]

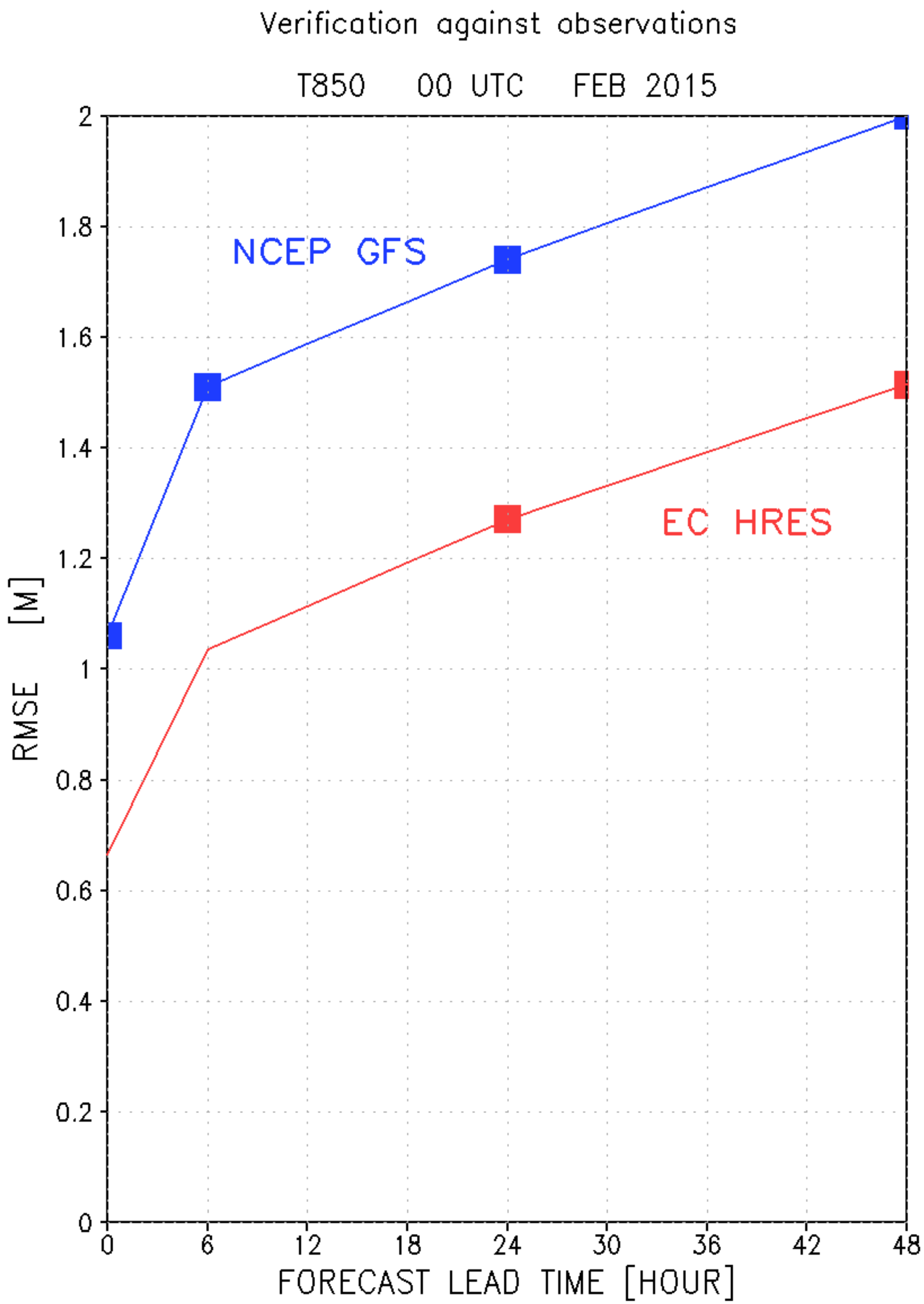

(F, O) Correlation Coefficient
ACC
0
0.2
0.4
0.6
0.8
1
5
10
15
20
Forecast Lead Time [Days]



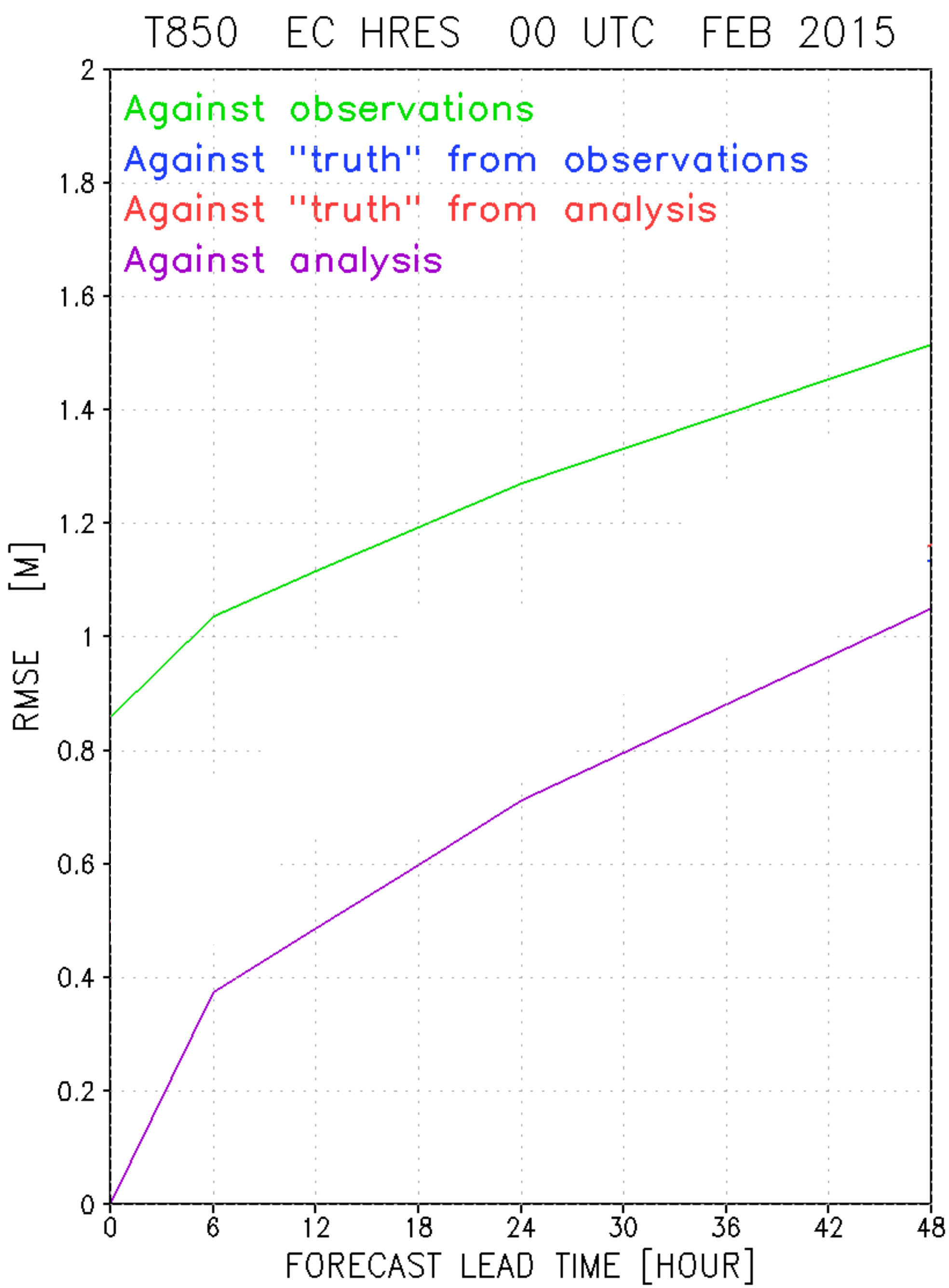
T850 EC HRES 00 UTC FEB 2015
Against observations
Against analysis
RMSE [M]
FORECAST LEAD TIME [DAY]
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10


T850 EC HRES 00 UTC FEB 2015
Against observations
Against analysis
RMSE [M]
FORECAST LEAD TIME [DAY]
0
1
2
3
4
5
6
1
2
3
4
5
6
7
8
9
10


Verification against analysis
T850 00 UTC FEB
NCEP CFS_V2
1999-2010
EC HRES
2015
RMSE [M]
FORECAST LEAD TIME [HOUR]
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0
6
12
18
24
30
36
42
48

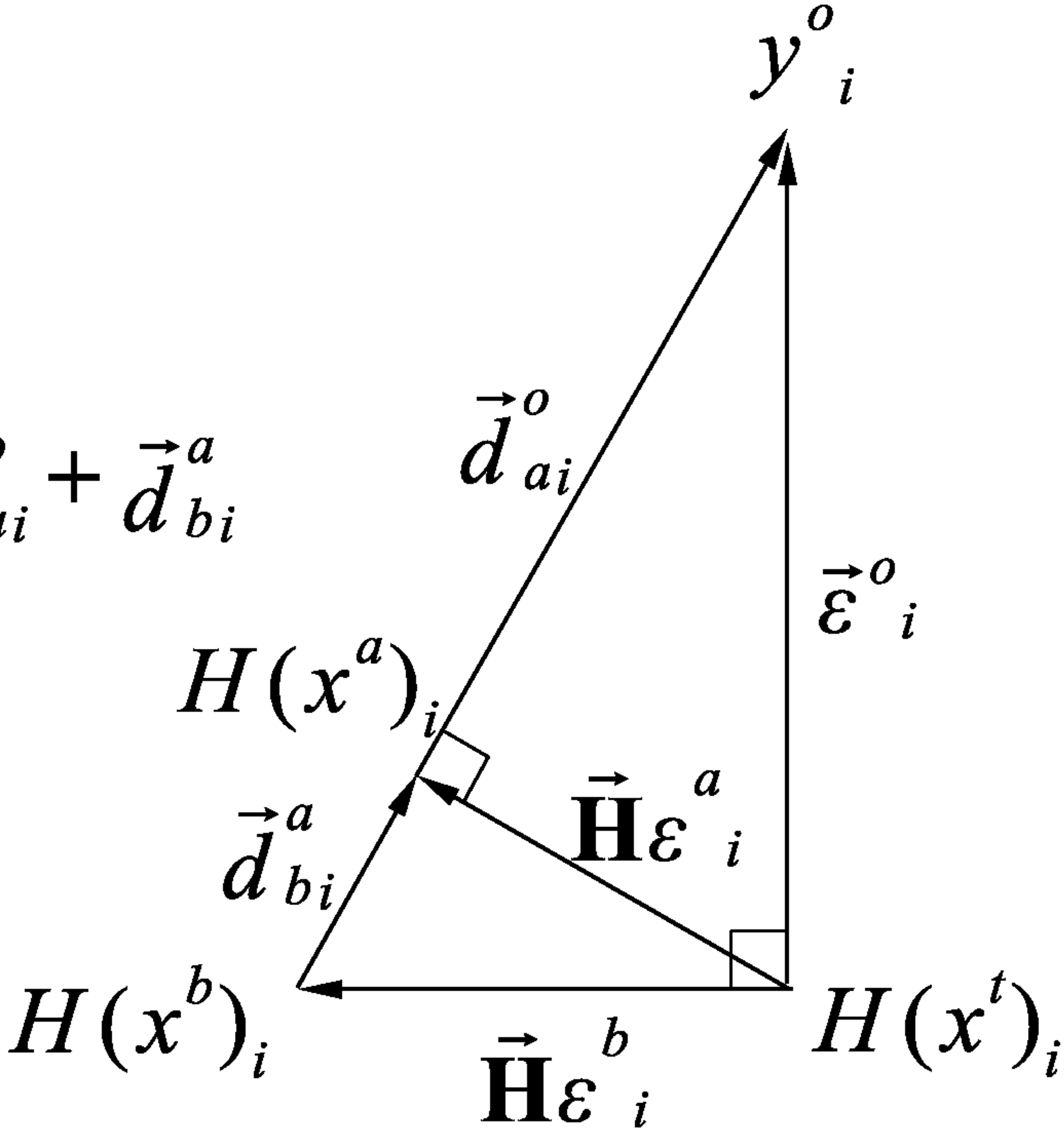


Verification against observations
T850 00 UTC FEB 2015
NCEP GFS
EC HRES
RMSE [M]
FORECAST LEAD TIME [HOUR]
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0
6
12
18
24
30
36
42
48

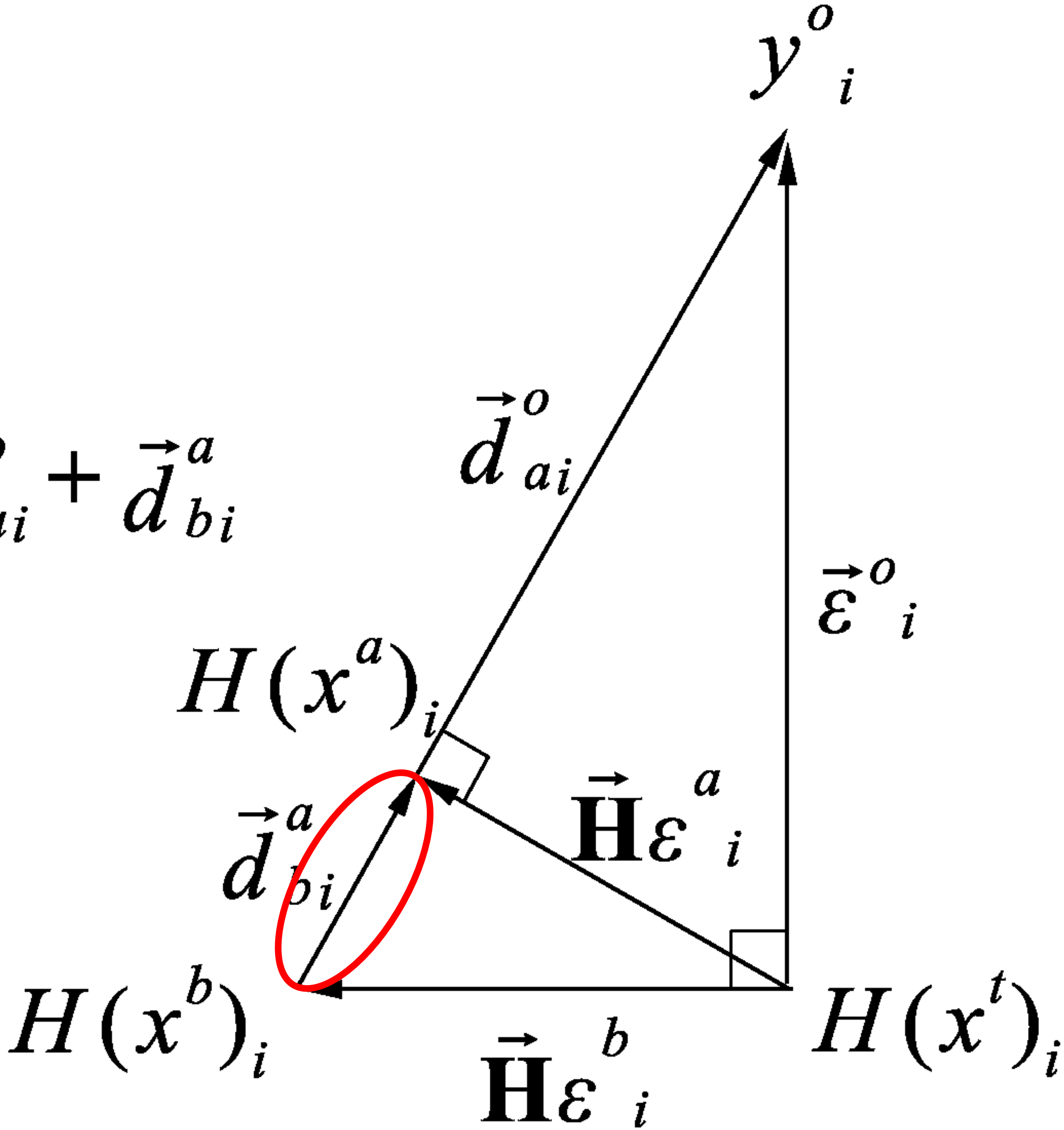


T850 EC HRES 00 UTC FEB 2015
Against observations
Against "truth" from observations
Against "truth" from analysis
Against analysis
RMSE [M]
FORECAST LEAD TIME [HOUR]
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0
6
12
18
24
30
36
42
48

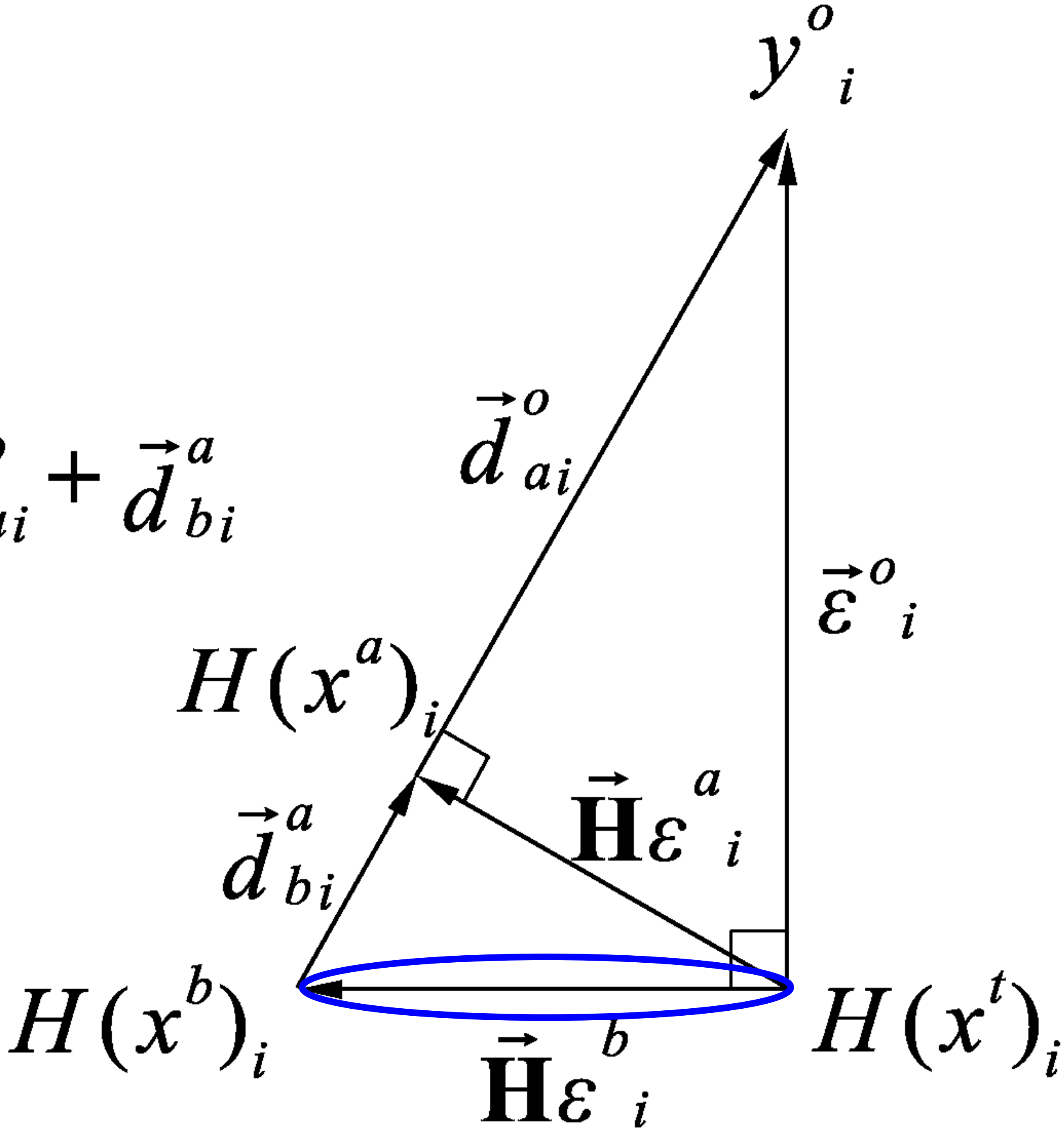
$$\vec{d}^{\,o}_{bi} = \vec{d}^{\,o}_{ai} + \vec{d}^{\,a}_{bi}$$



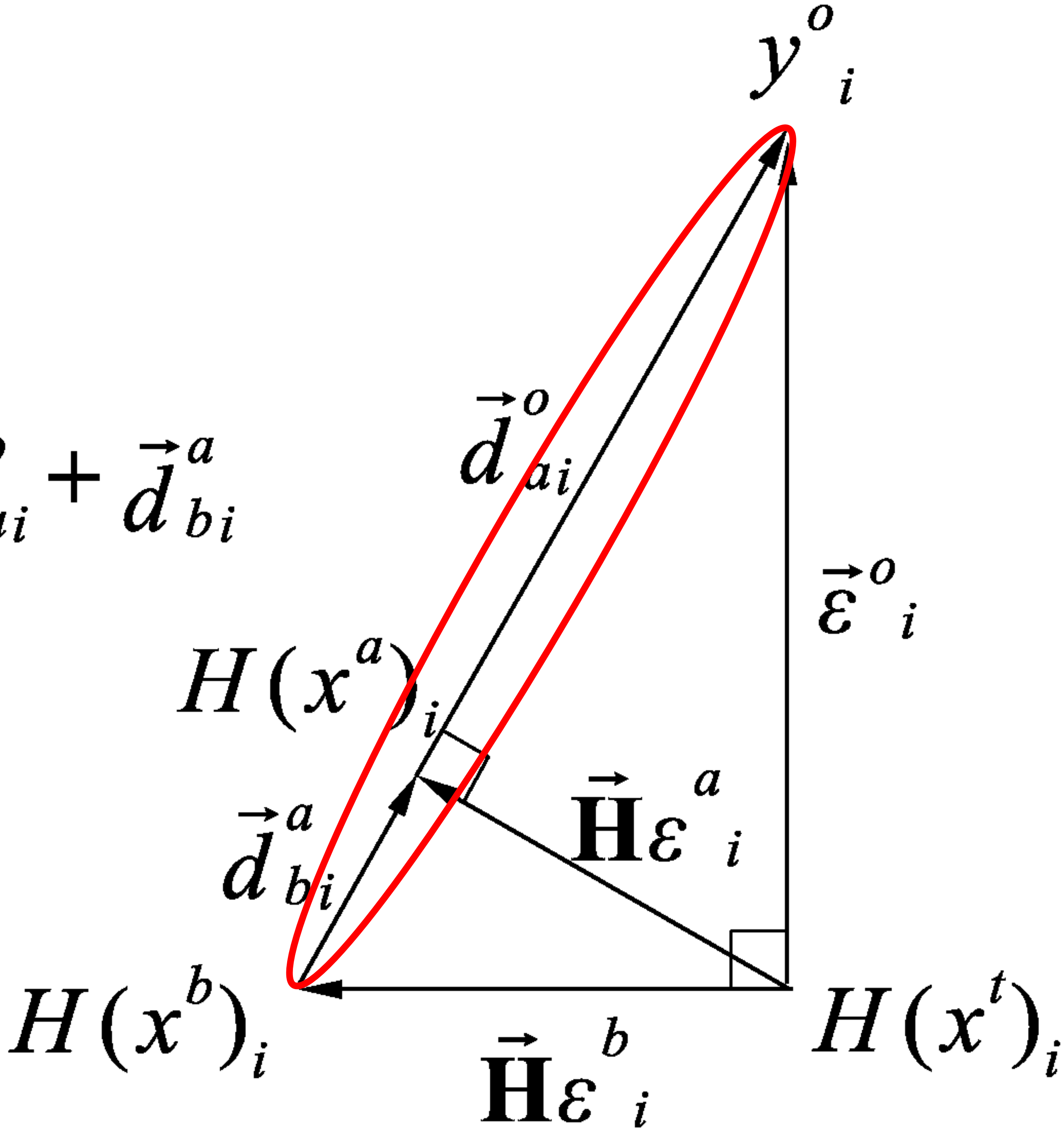
$$\vec{d}^{\,o}_{bi} = \vec{d}^{\,o}_{ai} + \vec{d}^{\,a}_{bi}$$

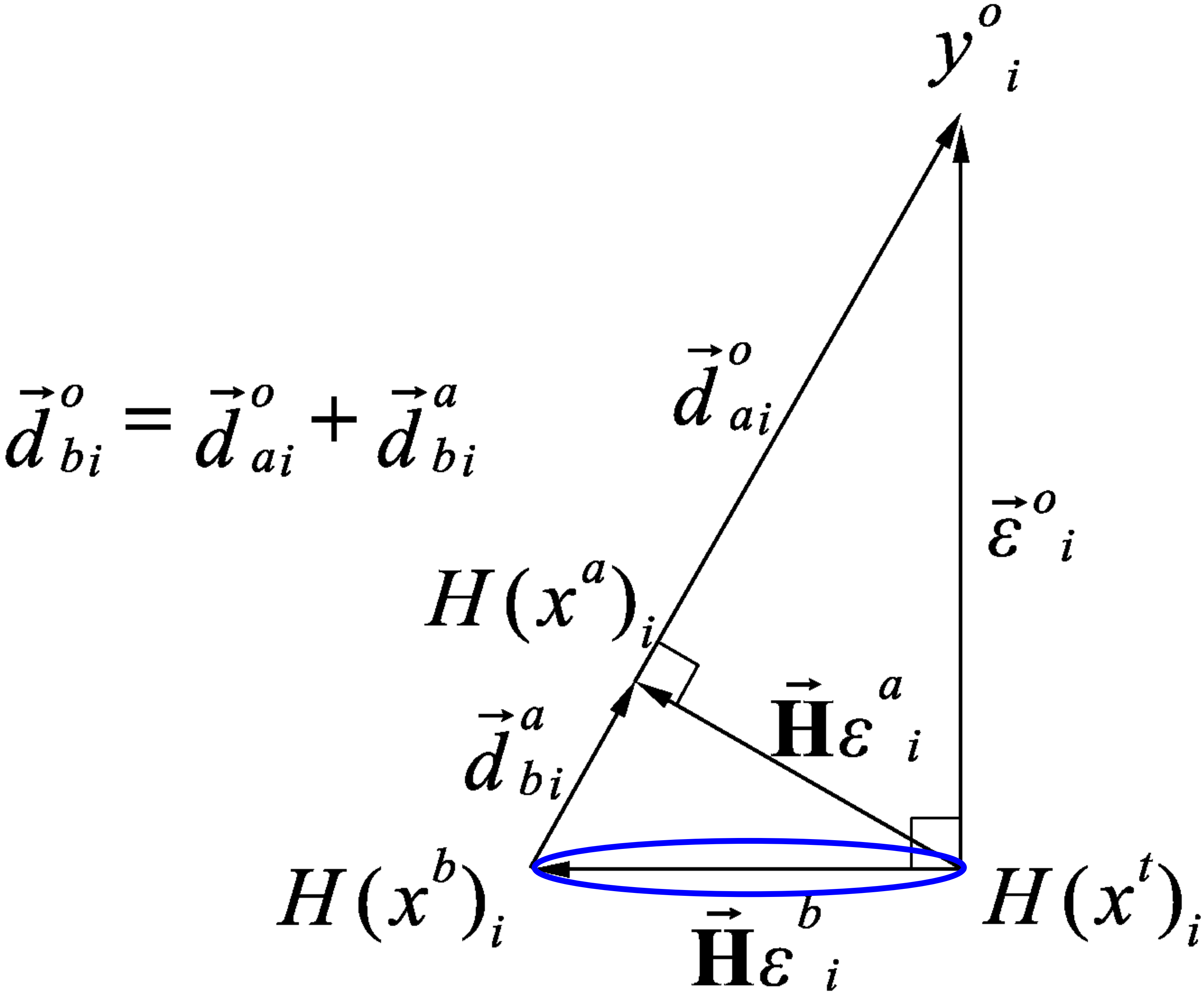


$$\vec{d}^{\,o}_{bi} = \vec{d}^{\,o}_{ai} + \vec{d}^{\,a}_{bi}$$

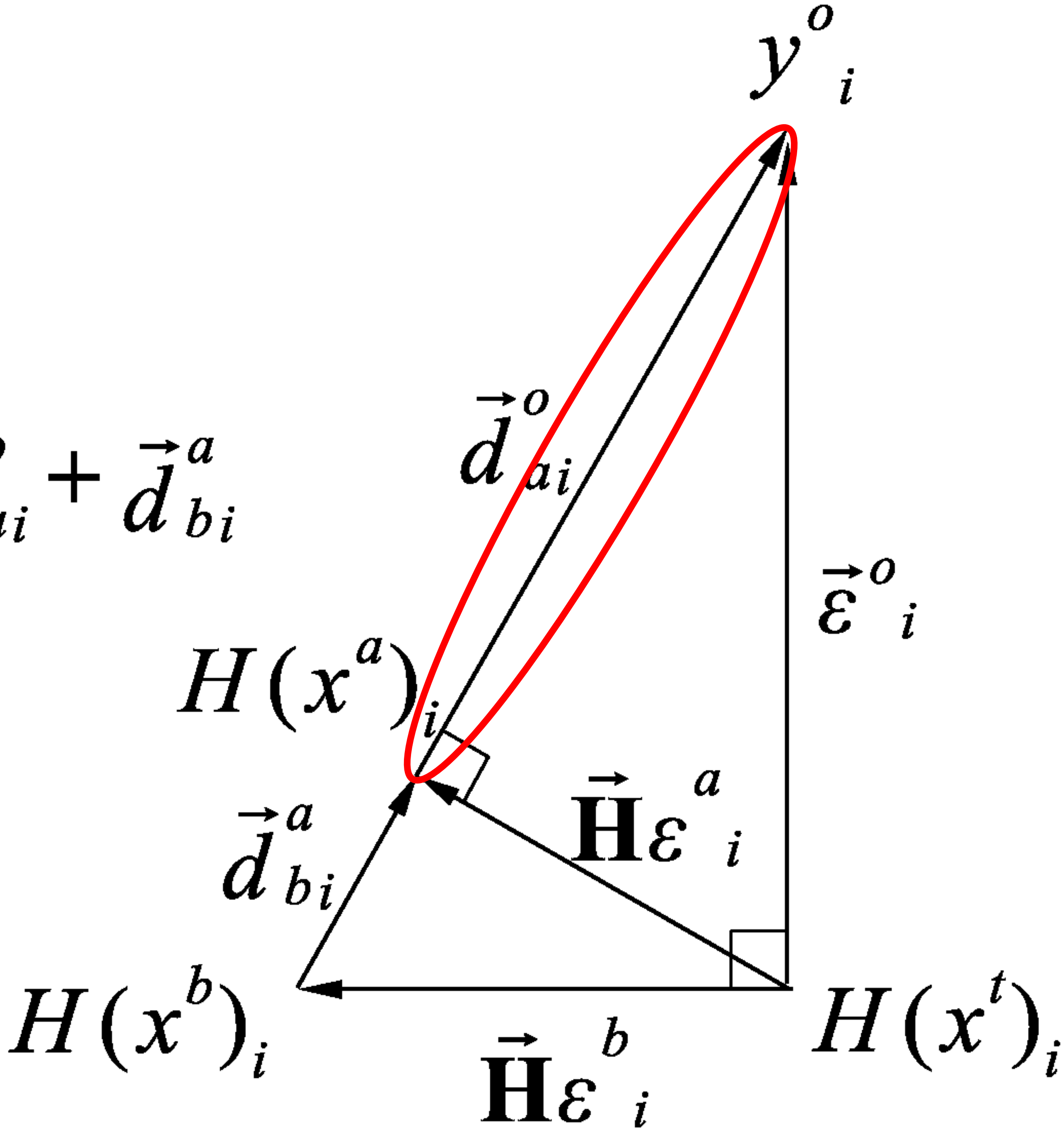


$$\vec{d}^{\,o}_{bi} = \vec{d}^{\,o}_{ai} + \vec{d}^{\,a}_{bi}$$

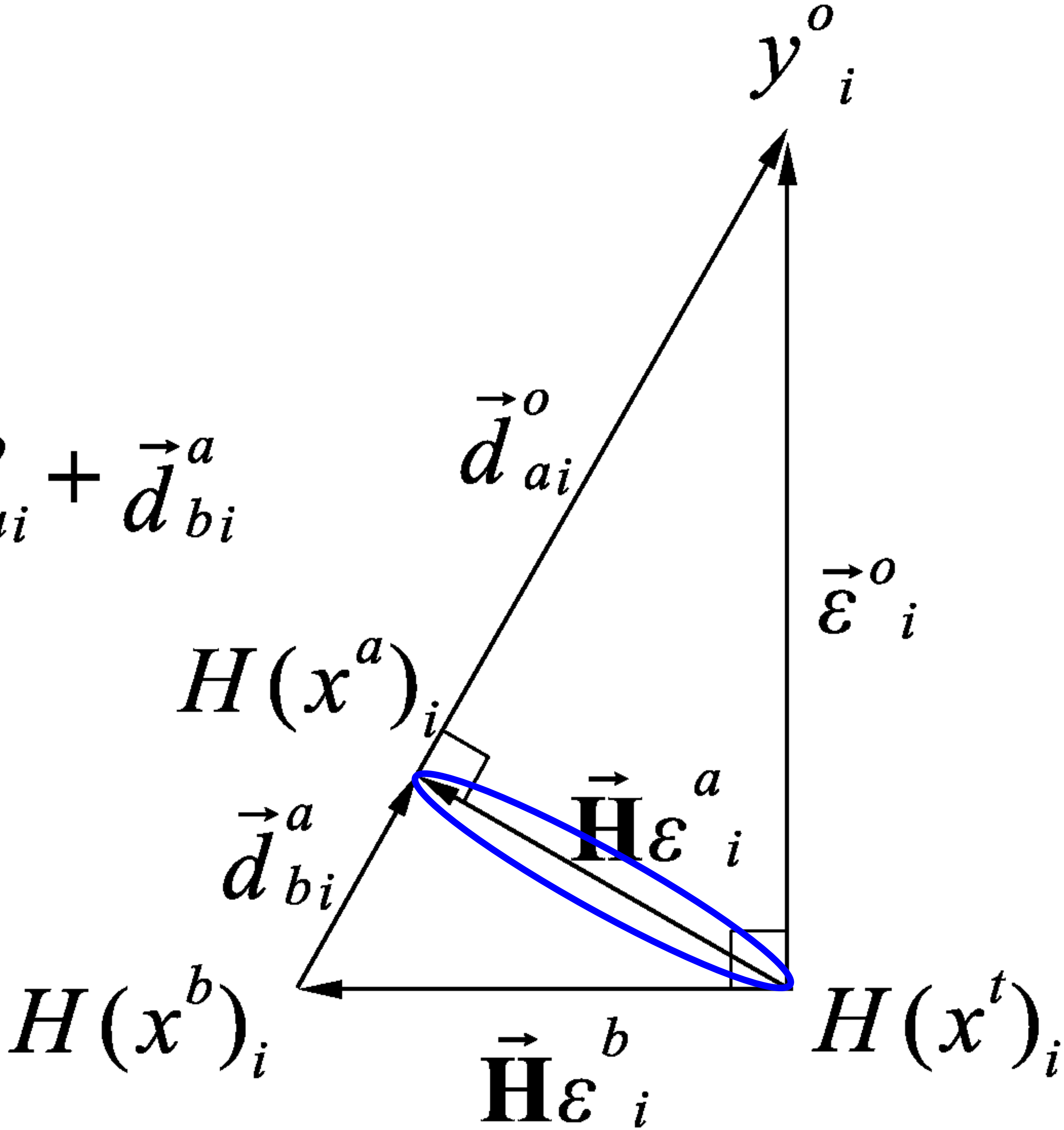


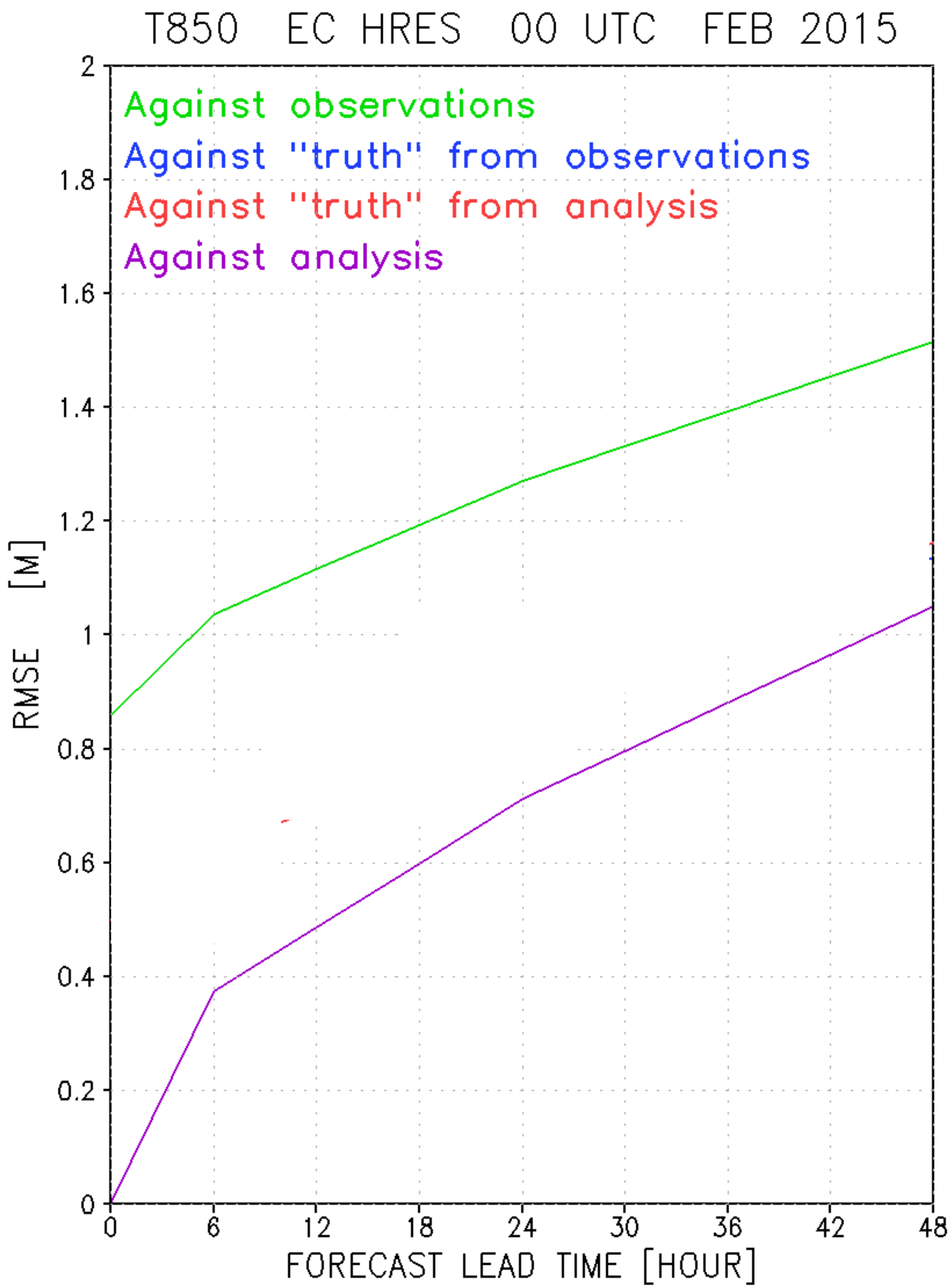

$\vec{d}^{\,o}_{bi} = \vec{d}^{\,o}_{ai} + \vec{d}^{\,a}_{bi}$
y^o_i
$\vec{d}^{\,o}_{ai}$
$\vec{\varepsilon}^{\,o}_i$
$H(x^a)_i$
$\vec{\mathbf{H}}\varepsilon^a_i$
$\vec{d}^{\,a}_{bi}$
$H(x^b)_i$
$H(x^t)_i$
$\vec{\mathbf{H}}\varepsilon^b_i$

$$\vec{d}^{\,o}_{bi} = \vec{d}^{\,o}_{ai} + \vec{d}^{\,a}_{bi}$$

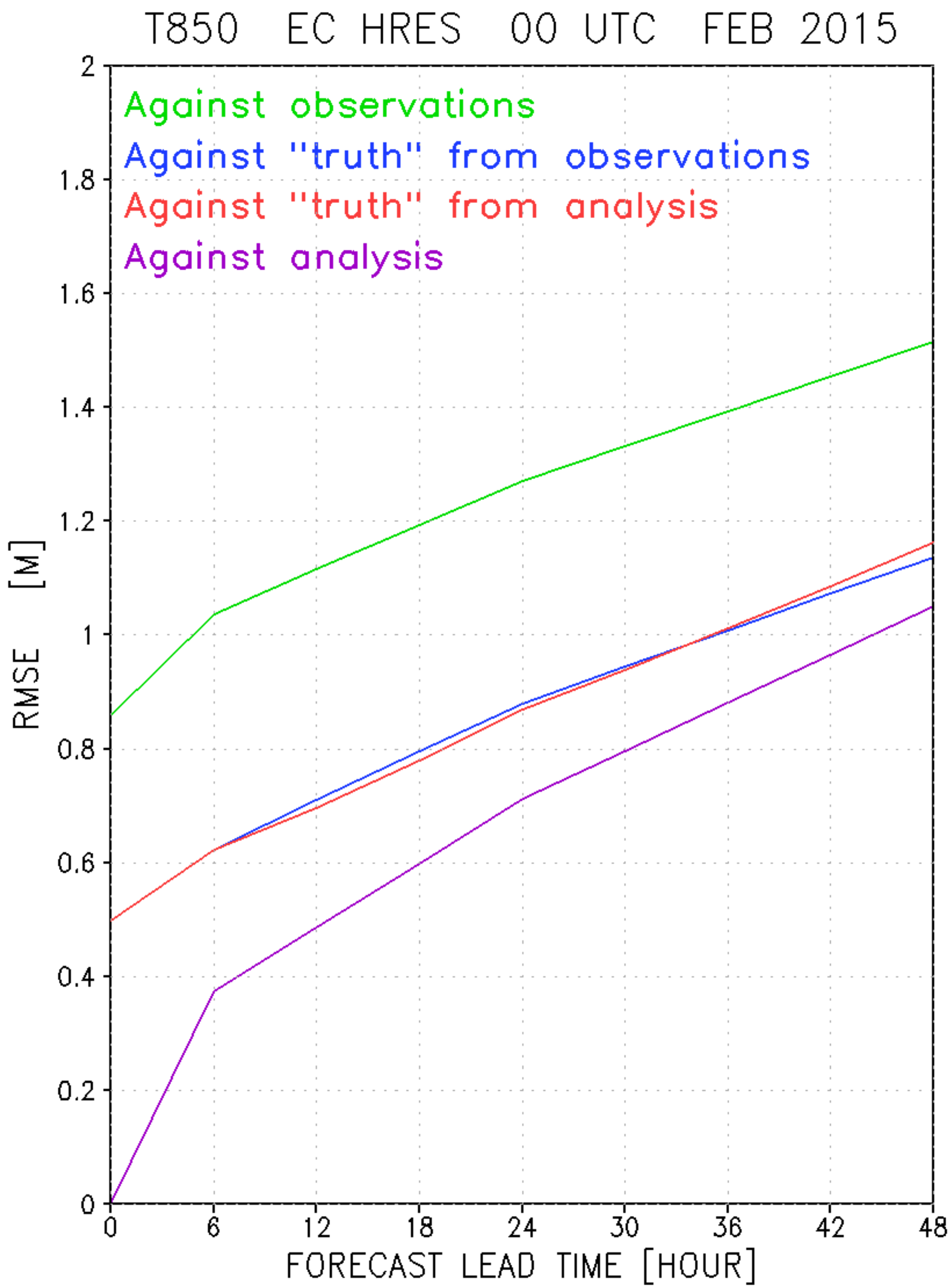


$$\vec{d}^{\,o}_{bi} = \vec{d}^{\,o}_{ai} + \vec{d}^{\,a}_{bi}$$

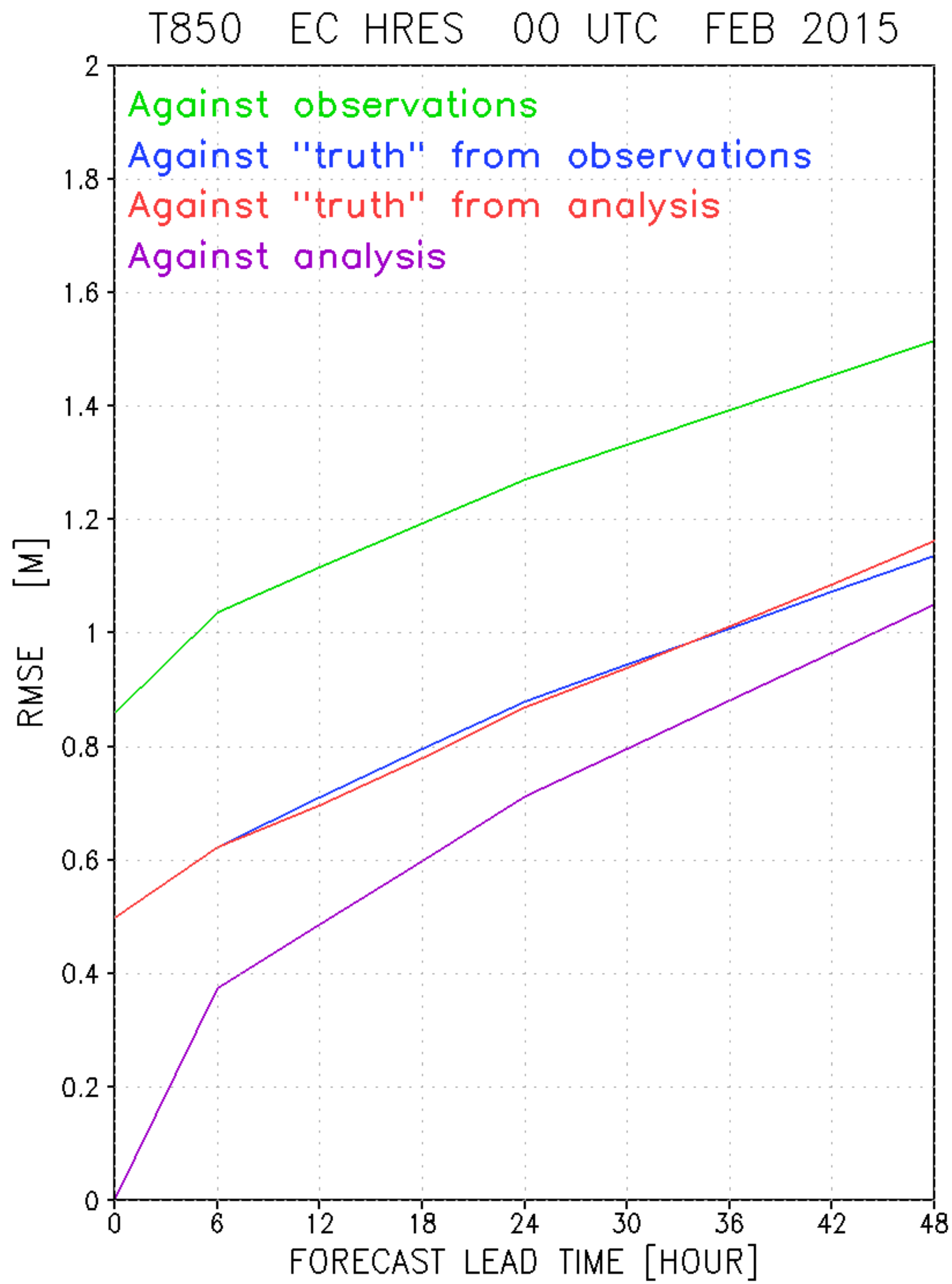




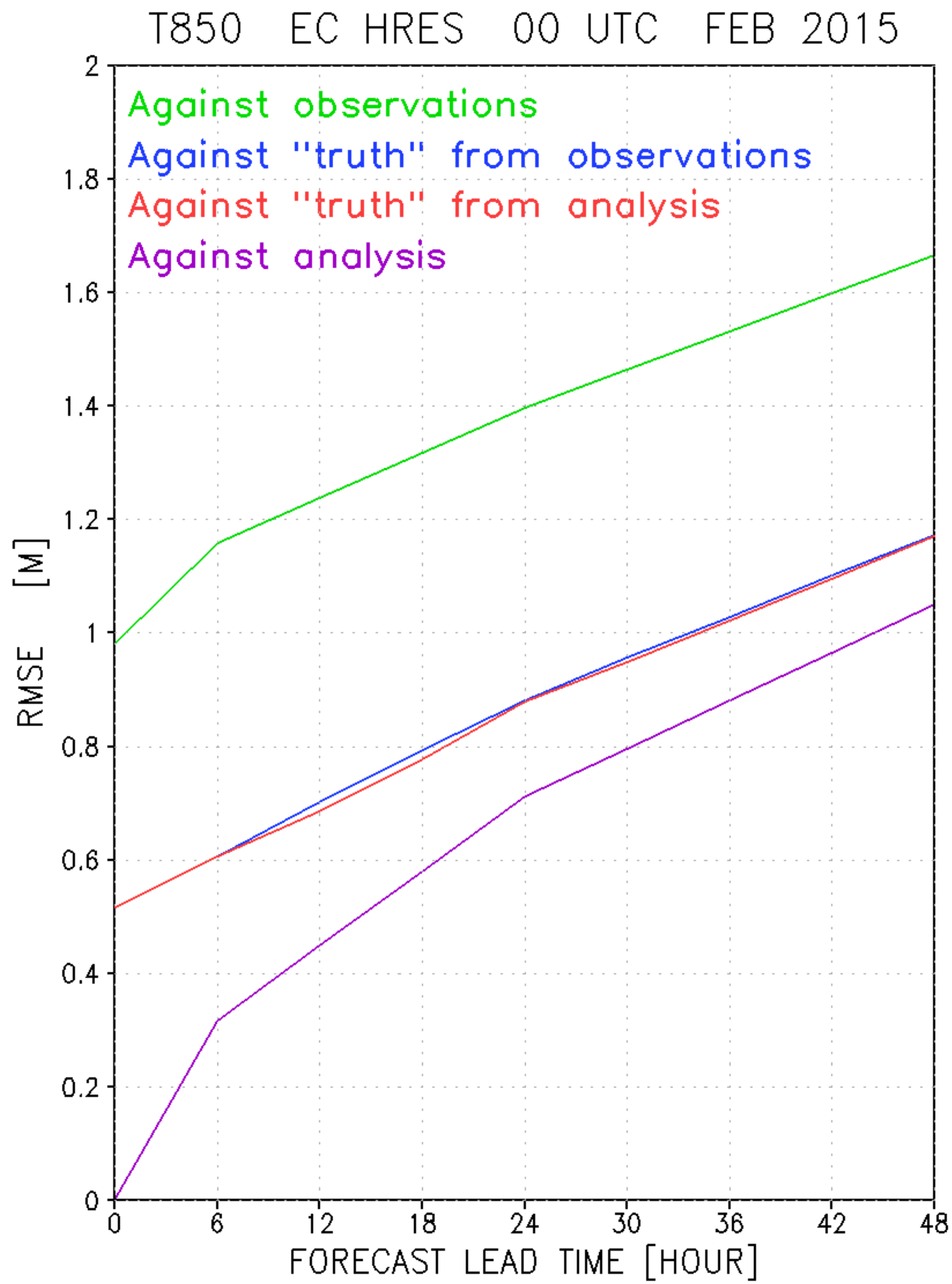
T850 EC HRES 00 UTC FEB 2015
Against observations
Against "truth" from observations
Against "truth" from analysis
Against analysis
RMSE [M]
FORECAST LEAD TIME [HOUR]
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0
6
12
18
24
30
36
42
48



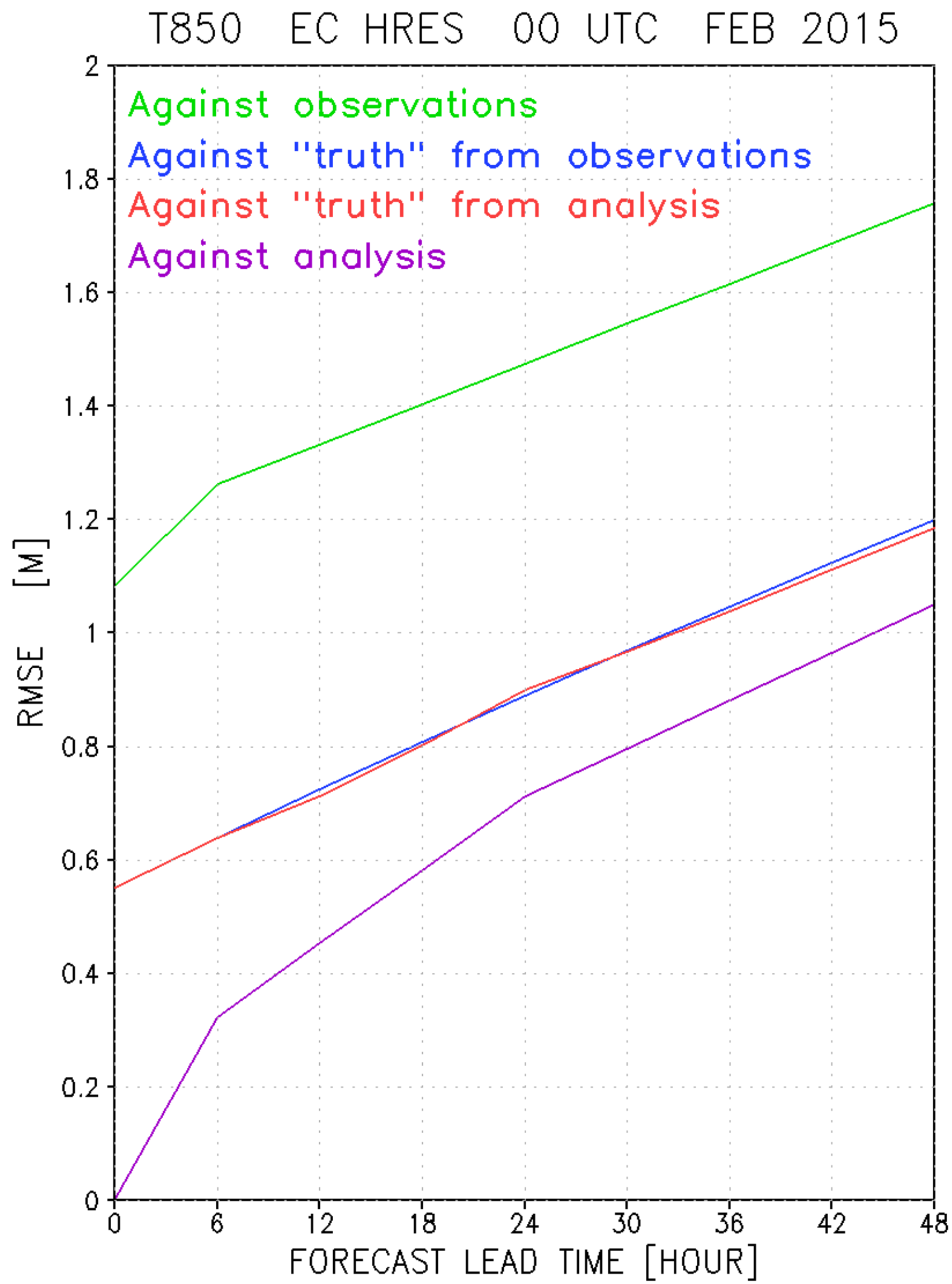
T850 EC HRES 00 UTC FEB 2015
Against observations
Against "truth" from observations
Against "truth" from analysis
Against analysis
RMSE [M]
FORECAST LEAD TIME [HOUR]
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0
6
12
18
24
30
36
42
48



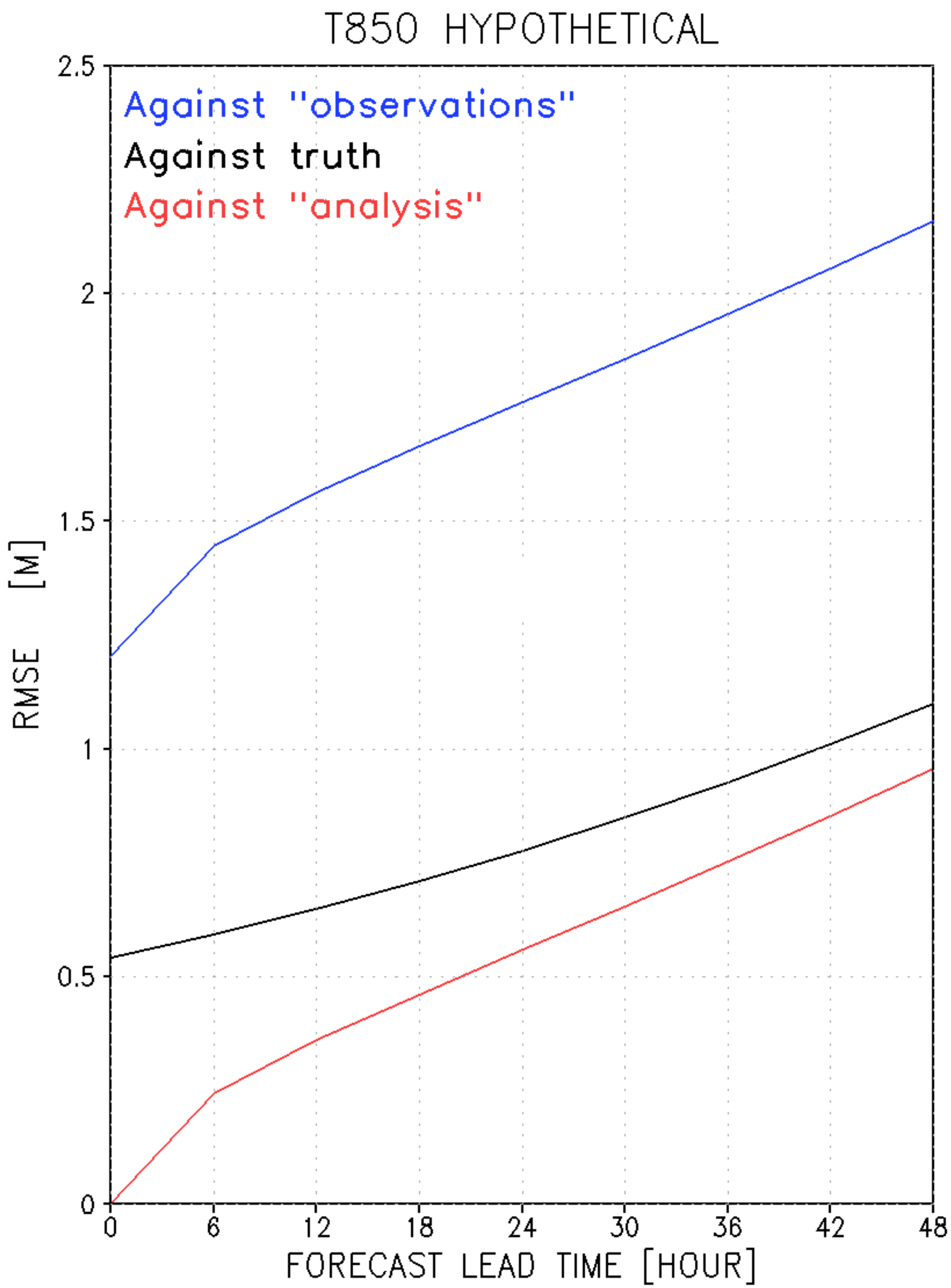
Inflation = 1.0



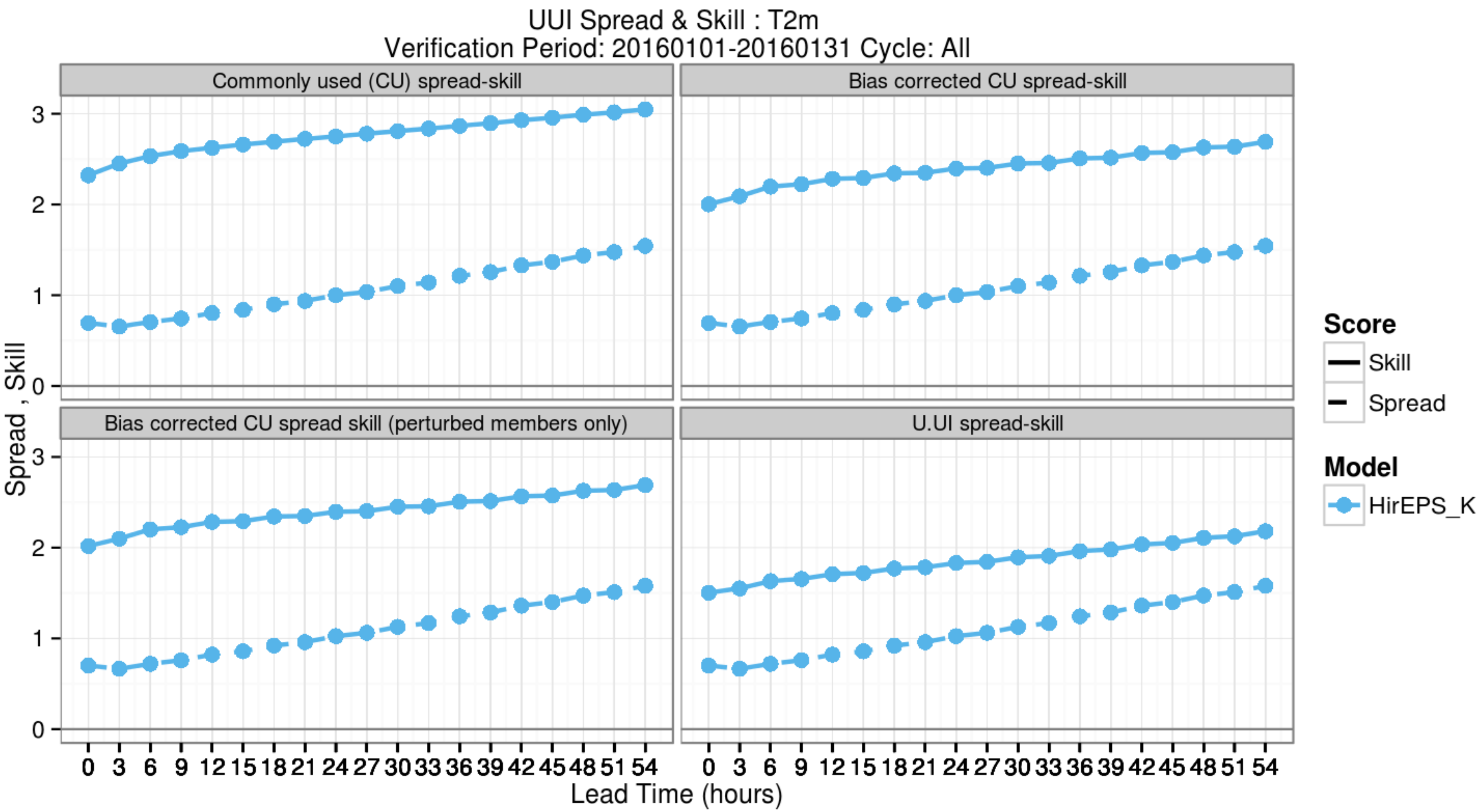
Inflation = 1.1

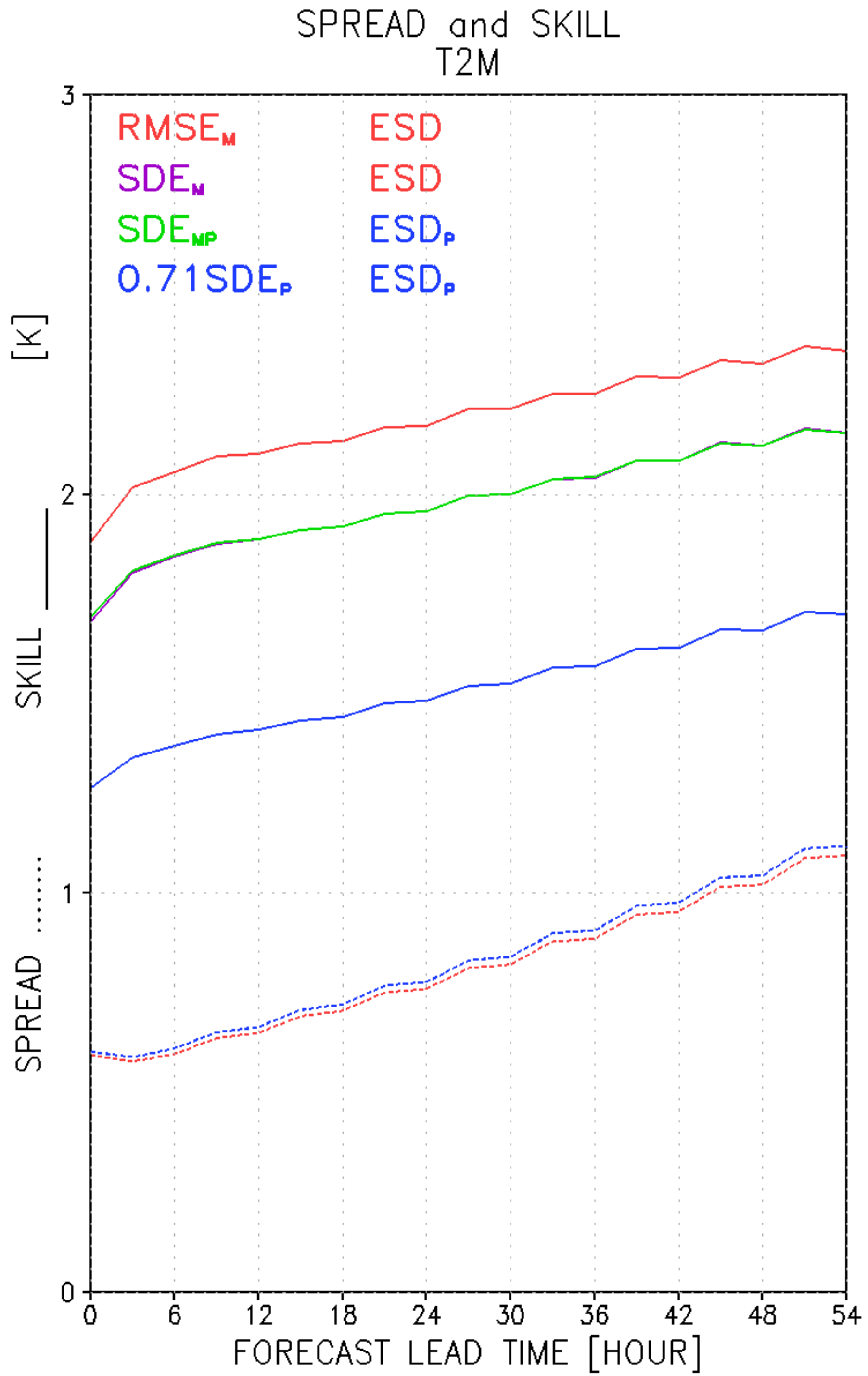


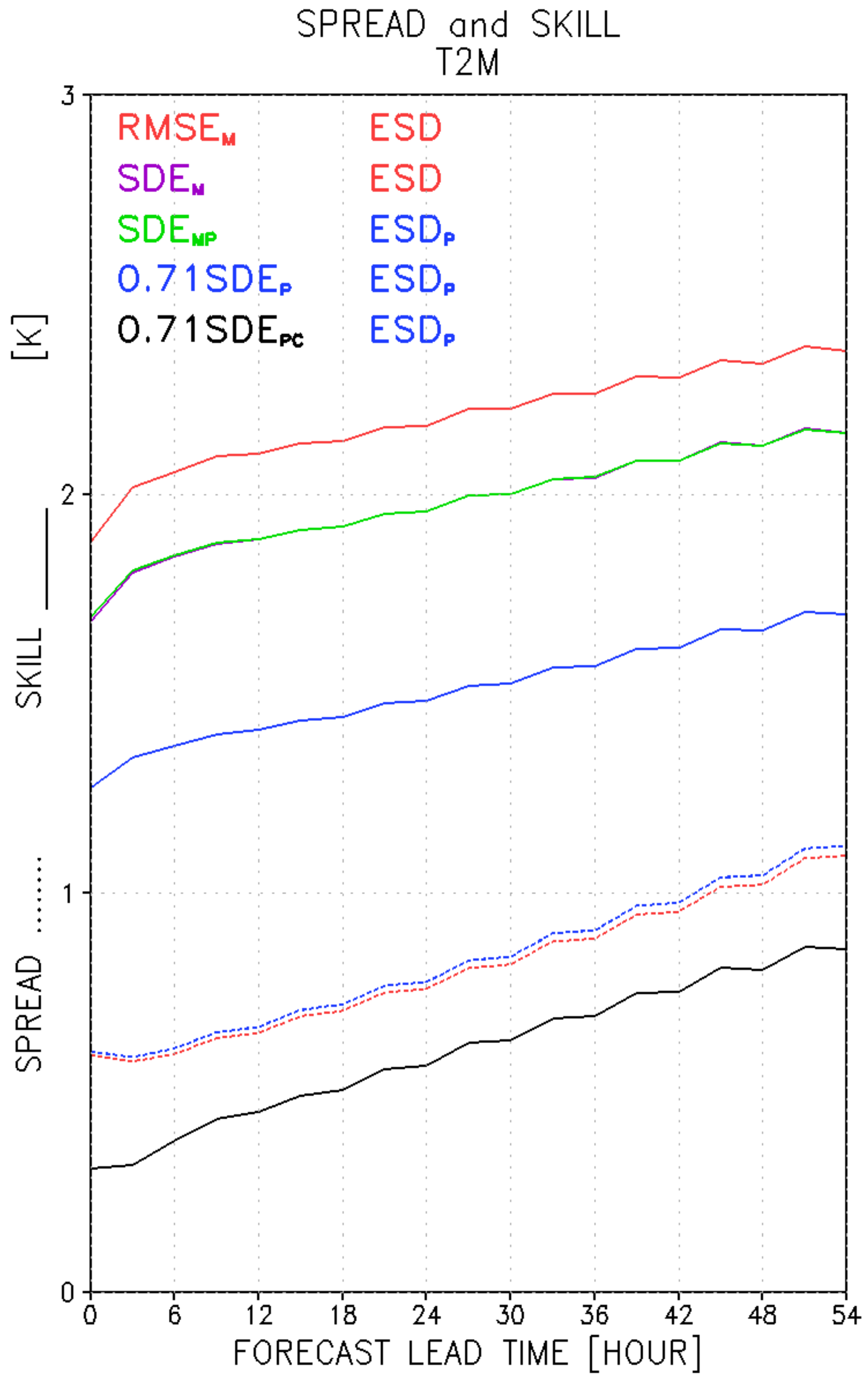
Inflation = 1.16


T850 HYPOTHETICAL
Against "observations"
Against truth
Against "analysis"
RMSE [M]
FORECAST LEAD TIME [HOUR]
2.5
2
1.5
1
0.5
0
6
12
18
24
30
36
42
48

Example for T2M from HIRLAM_K for JAN 2016




SPREAD and SKILL
T2M
RMSE_M ESD
SDE_M ESD
SDE_MP ESD_P
0.71SDE_P ESD_P
[K]
SKILL ——
SPREAD ....
FORECAST LEAD TIME [HOUR]
0
1
2
3
0 6 12 18 24 30 36 42 48 54



SPREAD and SKILL
T2M
RMSE_M ESD
SDE_M ESD
SDE_MP ESD_P
0.71SDE_P ESD_P
0.71SDE_PC ESD_P
[K]
SKILL ——
SPREAD ....
0
1
2
3
0
6
12
18
24
30
36
42
48
54
FORECAST LEAD TIME [HOUR]

Conclusions

Over-dispersiveness – not under-dispersiveness – seems to be the problem with EPS

Therefore one can wonder whether the following really is necessary

- **Parameter perturbations**
- **Stochastic physics**
- **Stochastic kinetic energy backscatter**