ARMY RESEARCH LABORATORY

Revisiting the Analysis of the High Resolution Wind Model Using MADONA Field Data

by Chatt C. Williamson, Dennis M. Garvey, and Sam S. Chang

ARL-TR-3497 **April 2005**

NOTICES

Disclaimers

Army Research Laboratory
Adelphi, MD 20783-1145

ARL-TR-3497 April 2005

Revisiting the Analysis of the High Resolution Wind Model Using MADONA Field Data

Chatt C. Williamson, Dennis M. Garvey, and Sam S. Chang
Computational and Information Sciences Directorate, ARL

REPORT DOCUMENTATION PAGE					Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE *(DD-MM-YYYY)* April 2005	**2. REPORT TYPE** Summary				**3. DATES COVERED (From - To)**
4. TITLE AND SUBTITLE Revisiting the Analysis of the High Resolution Wind Model Using MADONA Field Data					**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Chatt C. Williamson, Dennis M. Garvey, and Sam S. Chang					**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory Sensors & Electron Devices Directorate (ATTN: AMSRD-ARL-CI-EM) ccwillia@arl.army.mil Adelphi, MD 20783-1145					8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-3497
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) ARL 2800 Powder Mill Road Adelphi, MD 20783-1145					**10. SPONSOR/MONITOR'S ACRONYM(S)** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Cionco and Byers (1995) conducted an evaluation of the High Resolution Wind model (HRW) performance using the MADONA field study data (Cionco *et. al.*1999). The primary indicator for model performance used in their evaluation was the regression coefficient obtained from fitting a linear model to the observed data versus the simulated data. The regression coefficients obtained in their study were remarkably high, 0.9701 and 0.8567 for wind direction and speed respectively. In this report, we look further into this result to try to understand the significance of the large correlations indicated by these coefficients. We further investigate other statistical measures to evaluate the performance of the HRW model and compare these measures with a simple model using a homogeneous wind field.					
15. SUBJECT TERMS MADONA, HRW.					
16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT** UNCLASSIFIED	**18. NUMBER OF PAGES** 16	**19a. NAME OF RESPONSIBLE PERSON** Chatt C. Williamson
a. REPORT UNCLASSIFIED	**b. ABSTRACT** UNCLASSIFIED	**c. THIS PAGE** UNCLASSIFIED			**19b. TELEPHONE NUMBER** (*Include area code*) (301) 394-1378

Standard Form 298 (Rev. 8/98)

Contents

Figures

Tables

This page is intentionally left blank.

Introduction

A number of regression studies addressing the performance of the HRW model were performed by Cionco and Byers (1995) using data obtained during the MADONA field study. These authors evaluated the model using a simple linear regression analysis, namely y=ax+b, where x represents the measured wind direction or speed and y the corresponding model result. At least two wind fields were used to initialize the model. In this report, we will focus on the regressions performed using the M10 station to initialize the wind field of the model. Table 1 lists the 39 cases used in the evaluation of the HRW model performance and indicates the wind direction and speed used to initialize HRW. Our investigation was motivated by two factors: first, the unexpectedly high correlation values for wind direction as reported by Cionco and Byers (1995) and, secondly, graphs such as that presented in figure 1. Figure 1 is a plot of the wind direction, both modeled and observed at the various observation stations ordered by the observed wind direction measured at station M10. (The observed wind direction from the M10 station is also included for reference.) From this plot one can see that there can be significant spread between the observed and modeled wind direction. The same can also be said for measured and modeled wind speeds. It should also be noted that there are a number of cases where the observed wind direction and the corresponding model wind direction are on opposite sides of the M10 observation, the significance of which will be discussed later.

While HRW is a fairly simple model for this particular case, i.e., the MADONA evaluation, there is an even simpler model that we can compare against. This model can be termed the spatially homogeneous model (SHM). One advantage to comparing against this particular type of model is that the methodology for initializing the HRW model is to use an initial wind field that is homogeneous throughout the domain. The advantage here is we can estimate the value added to the estimation obtained by running the HRW model.

Table 1. The 39 cases used for the evaluation of the HRW models performance. The last column shows the order number for wind direction corresponding to the abscissa for Figure 1.

Case No.	Day	Month	Time	Wind @ M10		Order Number by Wind Direction
				Speed	Direction	
1	14	Sept	14:10	5.9	286	34
2			15:55	7.6	281	29
3			17:20	8	289	36
4	15	Sept	12:10	8.5	236	22
5			13:55	8.1	234	21
6			15:55	8	229	18
7			17:40	5	231	20
8	16	Sept	10:30	4	98	2
9			13:55	2.7	153	7
10			16:10	4.7	144	6
11			17:50	3.2	179	9
12	17	Sept	10:30	5.2	132	4
13			14:15	3.8	206	12
14			17:25	7.4	193	11
15	18	Sept	10:30	4.1	285	33
16			13:10	4.5	275	25
17			14:55	2.2	213	13
18			18:15	2.8	182	10
19	19	Sept	10:30	5.8	286	35
20			12:10	4.9	283	31
21			14:00	4.8	282	30
22			16:10	4.4	295	37
23			18:25	1.4	223	15
24	20	Sept	10:30	4.7	218	14
25			12:40	3.3	224	16
26			14:45	3.4	249	23
27			15:55	2.7	230	19
28	21	Sept	10:30	6.5	75	1
29			14:25	4	129	3
30			15:25	4.7	133	5
31			18:00	3.1	161	8
32	22	Sept	10:30	1.6	228	17
33			13:30	3.4	336	38
34			15:10	3.4	351	39
35	23	Sept	10:30	6.7	279	28
36			12:00	6.8	284	32
37			14:15	6.9	277	26
38			16:10	4.1	278	27
39			17:40	2.3	258	24

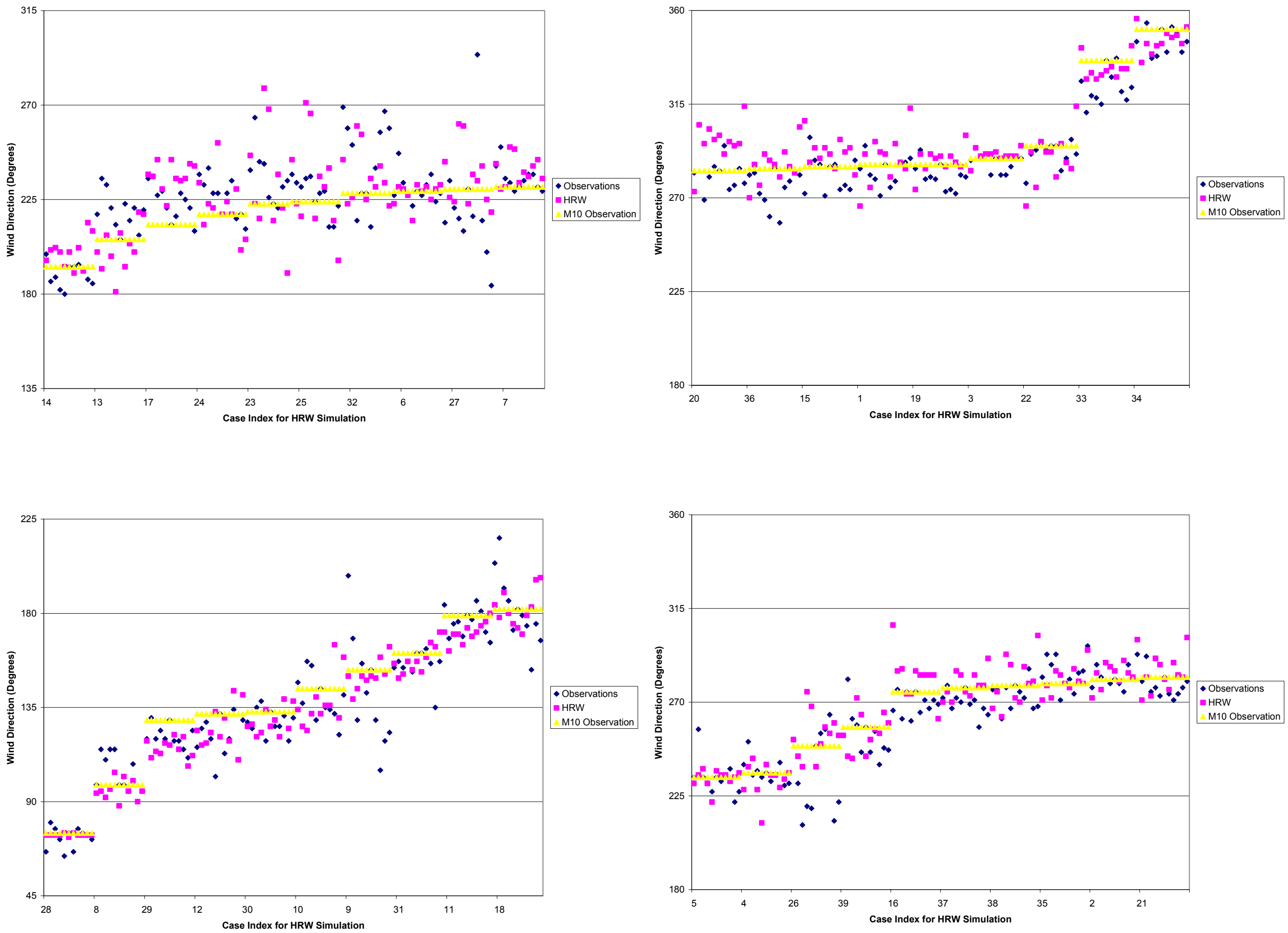

Figure 1. Observed and HRW modeled Wind Direction for the 39 Cases listed in Table 1. Note that the abscissa is ordered by the M10 Wind Direction.

Models and Data

The reader is referred to Ball and Johnson (1978) for a review of the basis of the HRW model. Slight modifications have been made to this code base (Byers 1985) though the underlying numerical scheme remains essentially unchanged. Briefly, the HRW model is a limited area diagnostic wind model that requires minimal meteorological information to produce a wind field estimate. Typical domains are on the order of 5 km by 5 km with a horizontal resolution of 100 m. Input required to initialize the model is a single wind speed and direction, a potential temperature profile, and an effective terrain height which is the combination of both the underlying terrain and the morphological features contained within the model domain.

The other models being considered, the spatially homogeneous models (SHM), are very simple. This type of 'model' is often used to drive transport and diffusion models for a quick assessment of a hazardous plume's fate. One cannot overemphasize the simplicity of the SHM. We note that there are obvious disadvantages to using simple homogeneous estimators should the complexity of the underlying terrain increase, should the domain of interest become relatively large, or should the atmospheric dynamics from larger scale features, such as frontal passages, influence the flow.

For convenience, we can denote the HRW model and the different SHMs we consider as

HRW Model: M1 $=\overrightarrow{V_1}(u_1, v_1)$,

SHM (M10): M2 $=\overrightarrow{V_2}(u_2, v_2)$,

SHM (μ): M3 $=\overrightarrow{V_3}(u_3, v_3)$,

SHM (LRO): M4 $=\overrightarrow{V_4}(u_4, v_4)$.

Here $\overrightarrow{V_1}(u_1, v_1)$ is a wind vector field produced by the HRW model (**M1**) as briefly described above. The second model (**M2**) is a spatially homogeneous model which produces a wind field $\overrightarrow{V_2}$ equal to the observed wind vector at station M10. The field $\overrightarrow{V_2}$ is constant throughout the domain (*i.e.* $(u_2 = \text{const}, v_2 = \text{const})$). Likewise, the third model (**M3**) refers to a homogeneous wind field, $\overrightarrow{V_3}$ equal to the mean observed wind, *i.e.*,

$$u_3 = \frac{1}{n}\sum_{k=1}^{n} u_k,$$
$$v_3 = \frac{1}{n}\sum_{k=1}^{n} v_k.$$

Here n is the total number of observations that were used in the previous regression studies of the MADONA data by Cionco and Byers (1995). Typically n was 11, though there were two cases, Case 9 and Case 23, where n was 10 and 9, respectively. The fourth model considered (**M4**) called the least representative observation model (LRO) is represented by a constant wind vector field $\overrightarrow{V_4}$, which is the homogeneous vector field corresponding to the observation which produces the lowest value of the linear correlation coefficient r. This coefficient will be discussed shortly. The LRO and spatial mean cases are used to establish a baseline of the possible worst case and best case scenarios, respectively.

The data used in this analysis was previously presented in Cionco and Byers (1995). We do further analysis of the observed-simulated data pairs in this report. The domain of interest was Porton Down, UK. The terrain could be described as gentle to moderate rolling hills.

Methodology

To investigate the relatively high linear regression correlation coefficients previously reported, we would like to establish a benchmark which can be compared against on a relative basis. This initial benchmark will use the linear regression correlation coefficient. We compare the regression coefficients for both wind speed and wind direction as computed by Cionco and Byers (1995) to those produced by computing the linear regressions for the SHM realizations described in the previous section. With this initial benchmark we hoped to be able to estimate the value added by running the HRW model for the MADONA case study.

A simple linear regression analysis can be expressed as

$$y = mx + b.$$

Here x is the observed value of either wind speed or wind direction. The variable y is the corresponding value produced by one of the four models M1, M2, M3, or M4. There are four measures that one can consider when using linear regression techniques to evaluate models. These four measures are m – the slope of the linear model, b – the y-intercept of the linear model, r – the correlation coefficient of the model, and *SEE* – the standard error estimate of the model. If the linear model were a perfect representation of the correspondence between the measured and modeled values, we would have the following:

$$\begin{aligned}&m = 1,\\&b = 0,\\&r = 1, \text{and}\\&SEE = 0.\end{aligned}$$

Typically, when using linear regression, we are concerned with *m*, *r*, and *SEE*. The key indicator used by Cionco and Byers (1995) was the correlation coefficient *r*. This is defined as follows:

$$r = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 \cdot \sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2}}.$$

The standard error estimate is defined as:

$$SEE = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}\left(y_i^{\text{Model}} - y_i^{\text{Linear Model}}\right)^2}.$$

We also look at other traditional statistical measures used in model evaluation studies such as bias, maximum and average absolute error, and the root mean square error (RMSE). They provide a more direct comparison of the differences between the measurements and the model outputs and are independent of the regression analysis, used to evaluate the model performance.

We define the Bias of a model to be the following:

$$Bias = \frac{1}{n}\sum_{i=1}^{n}\left(y_i^{\text{Model}} - y_i^{\text{Obs}}\right)$$

Note that the Bias gives a gross estimate of the model's tendencies. For example when modeling wind direction, we can estimate whether the model tends to favor being clockwise or counter-clockwise of the observed value. Knowing this characteristic, a modeler could then either correct the model as appropriate or inform users of the known bias. We define the Maximum Absolute Error *MAE* as:

$$MAE = \max_{i=1}^{n}\left|y_i^{\text{Model}} - y_i^{\text{Obs}}\right|$$

This particular error estimate is a measure of the worst case that a modeler could expect and could possibly establish a bound on whether or not to use the information produced by or derived from the model output. We define the Average Absolute Error *AAE* as follows:

$$AAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i^{\text{Model}} - y_i^{\text{Obs}}\right|$$

This error estimate is similar to the bias estimator described above and can be used in a similar manner. Finally, the Root Mean Square Error *RMSE* is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i^{\text{Model}} - y_i^{\text{Obs}}\right)^2}$$

This is the traditional root mean square error and is generally thought of as a good estimator for model performance.

Results and Discussion

Table 2 lists the values of *r*, *SEE*, *Bias*, *MAE*, *AAE* and *RMSE* for the wind direction and wind speed for the four models. As can be seen in table 2, even the homogeneous model using the Least Representative Observation (LRO) yields a regression coefficient that is quite high, indicating two things. First, this is an indicator that the underlying terrain, while not flat, is in a sense fairly benign, confirming our estimate earlier that the terrain could be described as gentle to moderate rolling hills. Second, this indicates that the correlation coefficient is a relatively weak measure in determining the validity of the model. It should also be noted from the table above that the model, in terms of the correlation coefficient performed, worse when compared to the homogeneous field used to initialize it.

Table 2. List of the statistical values r, SEE, Bias, MAE, AAE, and RMSE for wind direction and wind speed for the four models considered.

	r	*SEE*	*Bias*	*MAE*	*AAE*	*RMSE*
Wind Direction						
HRW	**0.9701**	**16.3**	**2.97**	**60**	**12.2**	**17.1**
SHM (M10)	**0.9811**	**12.5**	**1.98**	**64**	**9.8**	**13.3**
SHM(μ)	**0.987**	**10.3**	**-0.01**	**72.6**	**7.3**	**10.4**
SHM(LRO)	**0.9666**	**16.7**	**1.88**	**110**	**10.9**	**17.6**
Wind Speed						
HRW	**0.8567**	**1.02**	**0.33**	**3.6**	**0.85**	**1.13**
SHM (M10)	**0.9149**	**0.77**	**0.17**	**3.1**	**0.64**	**0.86**
SHM(μ)	**0.9357**	**0.65**	**0**	**2.9**	**0.53**	**0.69**
SHM(LRO)	**0.8454**	**0.83**	**-0.29**	**3.1**	**0.86**	**1.13**

Since the correlation coefficient is a fairly weak estimator in measuring the model performance for the MADONA studies, it is appropriate to look at other measures to ascertain the model performance. These other measures, as described previously, are the Bias, *MAE*, *AAE*, and *RMSE*. The values for these measures for both wind speed and wind direction are also reported in table 2. As is indicated in table 2, in each case the statistical measure becomes slightly worse once the homogeneous field is relaxed through the running of the HRW model, except for the wind direction maximum absolute error. Here, in this one case, we see a slight improvement. Recalling the caveats of the simple SHM model presented earlier, we investigated the relatively large *MAE* for both the HRW model and the SHM. These values were corresponding values for station M12 for Case 27 of the model runs. An inspection of the five minute averaged data revealed that the reason for these high values was most likely a frontal boundary with a wind shift propagating through the domain from west to east. Further analysis showed that the next largest *MAE* for the SHM was 48 degrees (Case 9), while there were three cases (9, 26 and 27), where the *MAE* for HRW was 48 degrees or greater. This further analysis also showed that there were 7 cases (9, 12, 18, 23, 26, 27, and 32) where the SHM and HRW both had an MAE greater than 30 degrees. There were an additional 9 cases where the HRW *MAE* exceeded 30 degrees while the SHM *MAE* did not.

It was noted earlier in our discussion of figure 1 that for a number of the observations, the corresponding model value produced by HRW was on the opposite side of the M10 observation. Specifically, of the 426 regression data pairs, 164 or slightly more than 1 in 3, exhibited this behavior. Recalling that HRW is initialized with a homogeneous wind field equal to the station M10 observation {*i.e.,* SHM(M10)}, we see that HRW actually 'corrected' the wind field in the wrong direction more than a third of the time. Slightly less than half, specifically 196 cases, were relaxed or 'corrected' in the appropriate direction. The remaining cases consisted mostly of those observation-model data pairs associated with a station M10 observation. Here the model relaxed the vector field away from the observation.

There were other measures of model performance that were investigated that are not reported here. These included looking at each case individually to see whether the model captured the range (max, min) of the observations for the wind speed and wind direction, as well as the mean, the median, and the quartiles. There was no discernable pattern to make any significant conclusion, though the mean values were typically within reason.

Conclusions

By these yardsticks, there is little to no value added by running HRW to estimate the wind field in this particular case. The correlation coefficients derived from the linear regression model for the Spatially Homogeneous Model for both the station M10 observation and the observations spatial mean for each case were slightly better than those that were previously reported. The reason for the high correlations previously reported is associated with the already high correlation of the observed winds at all of the sites with the homogeneous field used to initialize HRW for any one run. We can see that, statistically, the model performance is actually slightly worse than the homogeneous estimate used to initialize the model, and this holds true for all statistical measures that we have considered. As a measure of the HRW model performance, the correlation coefficient by itself and in combination with the standard error is quite weak and is not a good indicator of the validity of the model.

References

Ball, J. A. and S. A. Johnson, 1978: Physically-based High-Resolution Surface Wind and Temperature Analysis for EPAMS, Report ASL-CR-78-0043-1.

Byers, J. H., 1985: High Resolution Wind (HRW) Program for Wind, Temperature, and Turbulence Analysis, Report submitted to ASL under Contract DAAD07-82-C-0009. PSL, NMSU.

Cionco, R. M. and J. H. Byers, 1995: Initializing the HRW Model with Meteorological Inputs from Different Sites to Optimize Model Output. Proceedings of the 11th Symposium on Boundary Layers and Turbulence, AMS, Boston, Mass.

Cionco, R. M., W. aufm Kampe, C. Biltoft, J. H. Beyers, C. G. Collins, T. J. Higgs, A. R. T. Hin, P.-E. Johansson, C. D. Jones, H. E. Jorgensen, J. F. Kimber, T. Mikkelsen, K. Nyren, D. J. Ride, R. Robson, J. M. Santabarbara, J. Streicher, S. Thykier-Nielsen, H. van Raden, and H. Weber, 1999: An Overview of MADONA: A Multinational Field Study of High Resolution Meteorology and Diffusion over Complex Terrain, BAMS Vol. 80, No. 1, January 1999.

Distribution list

Admnstr
Defns Techl Info Ctr
ATTN DTIC-OCP (Electronic copy)
8725 John J Kingman Rd Ste 0944
FT Belvoir VA 22060-6218

US Army Rsrch Lab
ATTN AMSRD-ARL-CI-OK-T Techl Pub (2 copies)
ATTN AMSRD-ARL-CI-OK-TL Techl Lib (2 copies)
ATTN AMSRD-ARL-CI-EM D Garvey
ATTN AMSRD-ARL-CI-EM S Chang
ATTN AMSRD-ARL-CI-EM C Williamson (10 copies)
ATTN IMNE-ALC-IMS Mail & Records Mgmt (2 copies)