REPRESENTATIVE METEOROLOGICAL DATA FOR AERMOD: A CASE STUDY OF WRF-EXTRACTED DATA VERSUS NEARBY AIRPORT DATA

Brian Holland, Tiffany Stefanescu, Qiguo Jing, and Weiping Dai
BREEZE Software/Trinity Consultants, Dallas, TX, USA

1. INTRODUCTION

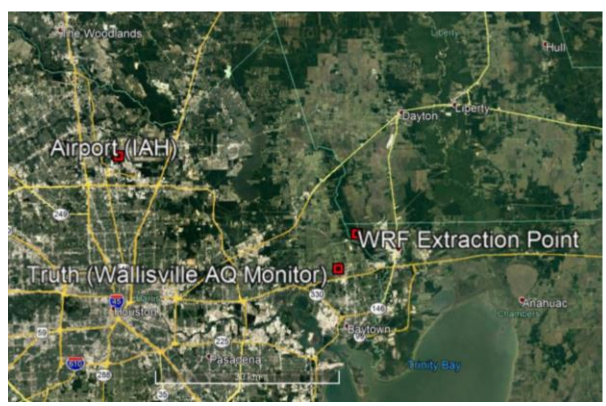
Historically, meteorological data for near-field air dispersion modeling (such as with AERMOD) has come primarily from the closest airport station to the facility being modeled, or from purpose-built "onsite" stations located at or near the facility. In areas where nearby observational data is not available or where meteorological conditions change rapidly with distance, these typical data sources become less representative of the actual facility location, introducing substantial error.

Recent changes to U.S. EPA's Appendix W air dispersion modeling document have opened the possibility of increased use of mesoscale meteorological model data (WRF or MM5) as an alternative source of meteorological data for near-field air dispersion modeling (U.S. EPA 2017). Site-specific mesoscale model data is promising in that it has the potential to eliminate most of the distance-based representativeness error described above. However, this comes at the cost of introducing forecast error from the mesoscale model, which will typically be larger than the observation error of a perfectly-placed surface meteorological station. Weighing the representativeness error of a distant airport meteorological station against that of an imperfect mesoscale meteorological model is a necessary but potentially difficult task in deciding which meteorological data source is most representative of a given location.

This study examines the relative magnitude of the errors in these two meteorological data sources in two case studies: one using a facility located in relatively flat terrain, and another using a facility located in complex terrain. In both cases, an on-site meteorological station is used as "truth". Meteorological data taken from a moderately distant airport station and from the closest grid cell of a WRF model run are compared to the on-site station's observations to quantify the relative error of each. AERMOD model runs are then carried out using each data source (site specific "truth", distant airport, and mesoscale model) to quantify the extent to which error in each meteorological source translates into dispersion model result error.

2. METHODS

For the simple terrain case study, the Wallisville Road air quality monitor location near Houston, Texas (AQS: 48-201-0617) was used as the source location. Onsite data from the monitor was used as an approximation of "true" meteorological conditions at the site. NWS airport meteorological data was taken from George Bush Intercontinental Airport (KIAH), 40 km to the northwest of the site. As the closest and most representative NWS station to Wallisville Road, this is the site that would most likely have been used in a typical real-world regulatory modeling application. The WRF dataset was extracted from the nearest gridpoint of a 12 km resolution national WRF simulation obtained from U.S. EPA (29.871N, 94.960W). The locations of all three sites are shown in Figure 1. Data from January-December 2007 was used for the simple terrain case study, as this was the most recent year available from all three data sources.



**Corresponding author:* Brian Holland, Trinity Consultants, 12700 Park Central Drive Suite 2100, Dallas, TX 75251. bholland@trinityconsultants.com

Fig. 1. Location of the three meteorological data sources (Truth, Airport, and WRF) for the simple terrain case.

For the complex terrain case study, sources were placed at the location of the Wamsutter, Wyoming air quality monitor (AQS: 56-037-0200), which was used to approximate the "true" meteorological conditions at the site. NWS airport data was taken from the Rock Springs, Wyoming Airport (KRKS), 86 km to the west of the site. Of the available NWS station data in this area, Rock Springs is the most representative of the Wamsutter site, but significant distance and differences in elevation and surrounding topography make it less than ideally representative of Wamsutter. This situation, in which the best-available NWS data is far from ideal in representing a project site, is a common occurrence faced by industry and regulators, particularly in areas of the Mountain West where topography and widely-spaced airports mean many possible source locations do not have representative meteorological data readily-available. The WRF dataset for the complex terrain case was extracted from the nearest gridpoint of a 12 km resolution national WRF simulation obtained from U.S. EPA (41.728N, 107.994W). The locations for all three sites are shown in Figure 2. Data from January-December 2008 was used for the complex terrain case study, as this was the most recent year available from all three data sources.

Fig. 2. Location of the three meteorological data sources (Truth, Airport, and WRF) for the complex terrain case.

All six datasets were processed using the latest version of AERMET according to standard U.S. EPA regulatory guidance and recommendations. The WRF data was extracted into simulated surface and onsite point data files using U.S. EPA's MMIF tool, and was then processed through AERMET to ensure as much consistency as possible with the "truth" and airport datasets. For the airport stations, 1-minute wind data was incorporated using AERMINUTE, and all datasets used the same 0.5 m/s wind threshold, with winds below that threshold being treated as calm hours (and thus being ignored by AERMOD). AERSURFACE was used to analyze the land use for the "truth" and airport datasets, while the WRF land use as extracted by MMIF was used for the WRF datasets.

The ADJ_U* option in AERMET, which is intended to offset AERMOD's tendency to over-predict concentrations from near-ground sources under stable, low wind conditions, was applied to the airport and WRF meteorological datasets, in accordance with U.S. EPA guidance U.S. EPA 2016a,b). It was not applied to the "truth" datasets due to the fact that the onsite stations used as "truth" include hourly σ_θ (standard deviation of horizontal wind direction) data. U.S. EPA guidance on use of ADJ_U* recommends that it not be used if direct measurements of turbulence are available. ADJ_U* operates by increasing the surface friction velocity (u*) used by AERMET for the stable atmosphere, low wind speed hours in which AERMOD otherwise would tend to over-predict ground-level concentrations (U.S. EPA 2016a).

AERMOD simulations were performed using each meteorological dataset. Because the impacts of different types of sources can be determined by different meteorological regimes and variables, two different sources were modeled – a ground-level volume source, and a 35 meter stack with 350 K exit temperature, 25 m/s exit velocity, and 1 m diameter. A receptor grid typical of standard regulatory modeling applications was used, with a small receptor-free buffer area around the source location (representing the area inside a facility fenceline), and three tiers of receptors:

-100 m spacing for the first 1 km past the fenceline
-500 m spacing from 1-5 km past the fenceline
-1000 m spacing from 5-10 km past the fenceline

Terrain data was incorporated in the modeling via the AERMAP utility. Building downwash was not incorporated. AERMOD simulations were carried out for a one-year period using the six datasets. Regulatory default settings were used in AERMOD, and maximum 1-hour, 24-hour, and annual concentrations were modeled.

3. RESULTS

3.1 Comparison of Meteorological Data

In the simple terrain case, generally similar wind patterns were present in all three meteorological datasets. Wind roses for the three datasets are shown in Figure 3. The most notable difference in patterns of wind direction was the increased frequency of the prevailing SSE/SE wind pattern in the WRF dataset. SSE/SE winds were present 32% of the time in the WRF dataset, compared to 25% of the time in the Airport dataset and 26% of the time in the Truth dataset. Low wind speeds were less frequent in the WRF data (12.6% < 1.54 m/s) and particularly in the Airport data (7.0% <1.54 m/s) compared to the Truth dataset (23.7% <1.54 m/s). In addition to underrepresenting low winds, the Airport dataset also overrepresented high wind speeds relative to the Truth dataset.

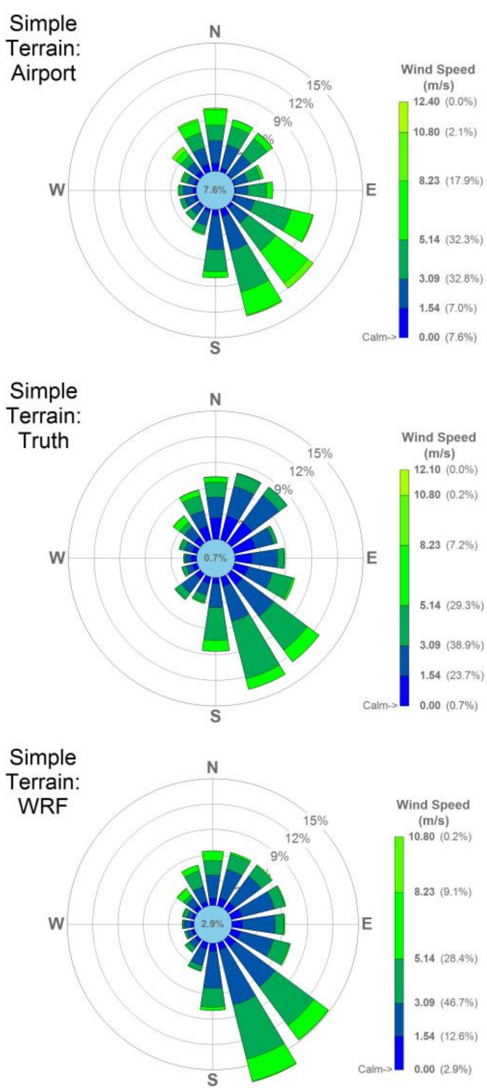
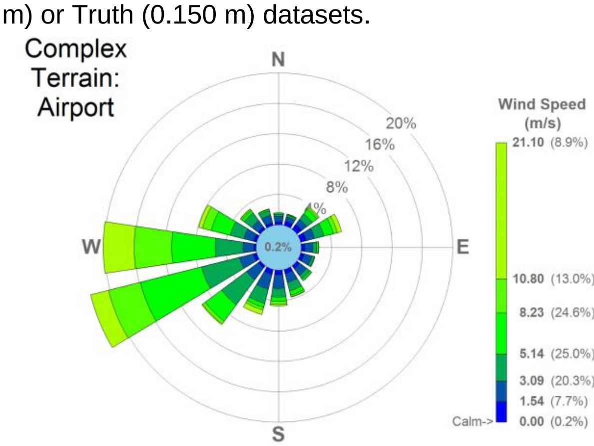


Fig. 3. Comparison of wind roses for airport, truth, and WRF datasets in the simple terrain case.

This is consistent with the relatively low surface roughness typically found at airports. The surface roughness corresponding to the prevailing wind in this case was in fact lower for the Airport data (0.045 m) than for the WRF (0.168 m) or Truth (0.159 m) datasets. Similarly, average wind speed was highest in the Airport data (3.5 m/s compared to 2.9 m/s in the WRF data and 2.8 m/s in the Truth data). Calm winds, which AERMOD does not model, were most common in the Airport data (7.6%, versus 2.9% in the WRF data and 0.7% in the Truth data).

In the complex terrain case, major differences in wind speed and direction patterns were evident between the three datasets. Wind roses for the three datasets are shown in Figure 4. The Truth dataset shows a prevailing WNW-WSW wind that represents 36% of hours, but a wide range of other wind directions are also common. The WRF dataset captures some of this variability (WNW-WSW represents 35% of hours, but a frequent SE wind in the Truth data is not seen in the WRF data). The airport data gives much heavier weight to the prevailing WNW-WSW wind (51% frequency) and underrepresents other wind directions.

High winds are somewhat more frequent in the Airport dataset than in the Truth dataset. The WRF dataset underrepresents high winds, possibly as a result of insufficient grid resolution to resolve terrain- or thunderstorm-induced winds. As in the simple terrain case, average Airport wind speeds (5.6 m/s) were higher than average WRF (4.0 m/s) or Truth (5.3 m/s) wind speeds, and average surface roughness was lower for the Airport dataset (0.063 m) than for the WRF (0.249 m) or Truth (0.150 m) datasets.



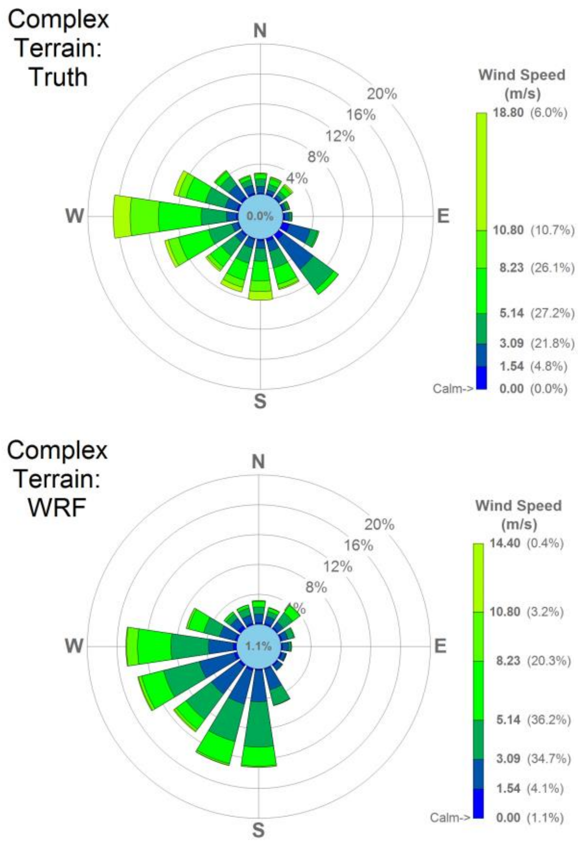


Fig. 4. Comparison of wind roses for airport, truth, and WRF datasets in the complex terrain case.

3.2 Comparison of AERMOD Results

A summary of peak 1- hour, 24-hour, and annual average concentrations is provided in Table 1. The concentration data is normalized so the “Truth” results have a value of 1.0. Thus, higher values represent over-prediction relative to the “Truth”, and lower values represent under-prediction. Results are presented for each case (simple and complex terrain), and for the tall stack source and ground level source.

In the simple terrain case, the WRF dataset consistently over-predicted peak concentrations for the tall stack source and consistently under-predicted peak concentrations for the ground level source. The Airport dataset over-predicted concentrations for the tall stack in the 24-hour and annual periods, but under-predicted the maximum 1-hour concentrations. As with the WRF data, the Airport dataset consistently under-predicted concentrations for the ground level source. With the exception of the ground level source annual averaging period, the WRF dataset consistently produced more conservative results than the Airport dataset.

In the complex terrain case, the tall stack results were more mixed, but the pattern of ground level results consistently being under-predicted by both the Airport and WRF datasets was present.

Maximum Annual Concentration				
Source Group	Simple Terrain		Complex Terrain	
	Airport	WRF	Airport	WRF
Tall Stack	1.34	1.67	1.28	0.80
Ground Level	0.52	0.45	0.50	0.39
Maximum 1-Hour Concentration				
Source Group	Simple Terrain		Complex Terrain	
	Airport	WRF	Airport	WRF
Tall Stack	0.85	1.29	0.85	1.21
Ground Level	0.14	0.16	0.21	0.29
Maximum 24-Hour Concentration				
Source Group	Simple Terrain		Complex Terrain	
	Airport	WRF	Airport	WRF
Tall Stack	1.37	1.70	0.86	1.10
Ground Level	0.37	0.42	0.24	0.22

Table 1. Summary of maximum ground level concentrations in each case, normalized so the “Truth” concentration is 1.00.

Normalized Bias (1-Hour Concentrations)				
Source Group	Simple Terrain		Complex Terrain	
	Airport	WRF	Airport	WRF
Tall Stack	-20%	30%	2%	-12%
Ground Level	-81%	-63%	-45%	-35%
Normalized RMSE (1-Hour Concentrations)				
Source Group	Simple Terrain		Complex Terrain	
	Airport	WRF	Airport	WRF
Tall Stack	34%	49%	47%	38%
Ground Level	124%	110%	126%	119%

Table 2. Bias and RMSE, normalized based on the average “Truth” concentration.

Normalized bias and RMSE were also calculated for 1-hour concentrations, treating the modeled onsite calculations as “Truth”. These results are shown in Table 2. Similar to the findings for maximum concentrations, both the Airport and WRF datasets showed a consistent under-prediction bias for the ground level source, and lower bias for the tall stack source. Normalized RMSE for the WRF dataset was lower than for the Airport dataset with the exception of the simple terrain, tall stack case.

Q-Q plots of 1-hour concentration data are presented in Figure 5 (simple terrain) and Figure 6 (complex terrain). The general trends of error and bias described above can be seen.

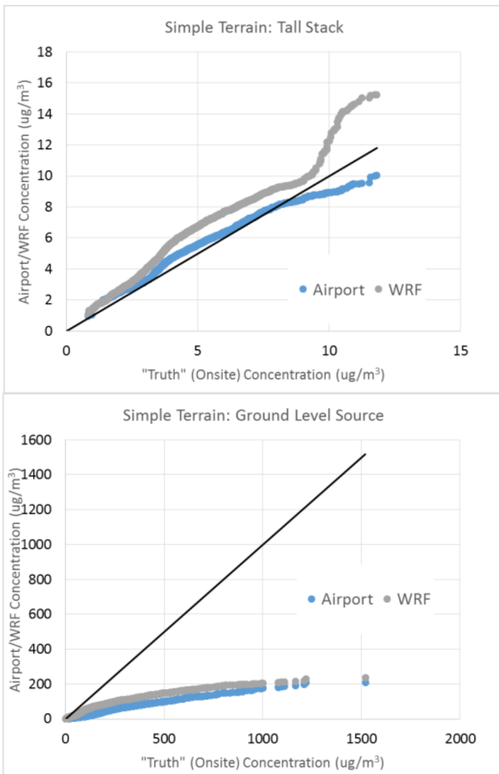
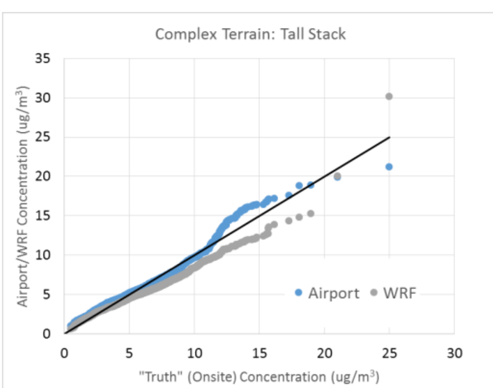


Fig. 5. Q-Q plots for 1-hour concentrations resulting from each source type in the simple terrain case.



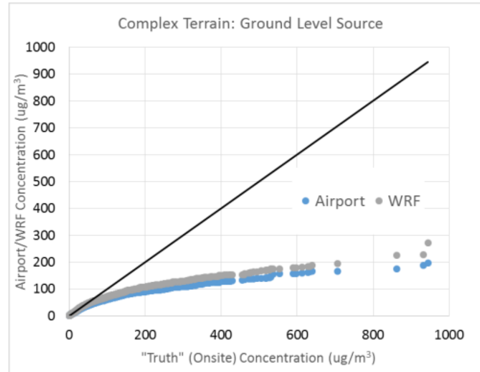


Fig. 6. Q-Q plots for 1-hour concentrations resulting from a tall stack and ground level source in the complex terrain case.

4. DISCUSSION

4.1 Effectiveness of WRF-Derived Meteorological Data versus Traditional Airport Data

This study examined two cases: a simple terrain case in which the airport meteorological data typically used in regulatory AERMOD modeling would generally be considered appropriate and representative of the modeled source location, and a complex terrain case in which the best available airport meteorological data is quite distant (80 km) and not clearly representative due to different terrain than the actual source location.

For each case, three sets of AERMOD modeling was performed: using onsite meteorological data from each source location, using data from the airport that would typically be used in a regulatory modeling application, and using data derived from a 12 km resolution WRF model simulation. The AERMOD results using Airport and WRF data were each compared to the results using onsite data, considering the results using onsite data to be "Truth" because onsite data is the preferred meteorological data source both on a scientific basis and in the eyes of U.S. EPA regulatory guidance.

AERMOD model accuracy when using WRF-derived data was approximately equal to accuracy when using Airport meteorological data. This was true both for the simple terrain case in which the Airport would generally be considered a representative data source, and for the complex terrain case in which the Airport might not be considered an acceptable data source. The details of the relative performance of the WRF-derived and Airport datasets varied among source type, averaging period, and assessment metric, but they were broadly equal in quality. Given that RMSE for the WRF-derived data was consistently lower than the Airport data in the complex terrain case, it appears that, as would be expected, the benefits of using WRF data over Airport data increase as the degree to which the Airport location is representative of the source location decreases. These findings that AERMOD performance using WRF data is at least as good as AERMOD performance using Airport data are consistent with the findings of U.S. EPA's evaluation of WRF and MMIF-derived meteorological data performance (U.S. EPA 2016c).

4.2 Applicability of ADJ_U to Onsite Meteorological Datasets That Include Partial Turbulence Data*

Possibly the most interesting results in this study actually did not relate to the intended study objective of comparing AERMOD performance using Airport and WRF-derived meteorological data, but to the poor performance of both Airport and WRF data when modeling a ground-level source. Both data sources resulted in large under-predictions of maximum ground level concentrations when compared to AERMOD results using onsite meteorological data.

The cause of this large discrepancy appears to be due to the decision, made in accordance with U.S. EPA regulatory guidance, to use the ADJ_U* AERMET option when processing the Airport and WRF-derived datasets, but not when processing the onsite "Truth" datasets. More discussion of ADJ_U* can be found in Section 1.

The large change in AERMOD performance for ground level sources seen in the cases where ADJ_U* is applied is not unusual, and is in keeping with the findings of U.S. EPA's thorough evaluation of the benefits of ADJ_U*. Thus, unless the multiple case studies used in U.S. EPA's evaluation of ADJ_U* are somehow fundamentally different than the two cases examined here, it is likely the case that for the ground level source in this case, the low concentrations produced by the WRF and Airport datasets are in fact more accurate than the high concentrations produced by the "Truth" onsite datasets. Recall that in this case, the decision not to apply ADJ_U* to the onsite datasets was a regulatory "gray area": ADJ_U* is supposed to be applied when turbulence data is not measured at the onsite station, but is not supposed to be applied when turbulence data is available. In this case, a small amount of turbulence data (σ_θ) was available. Thus, this case would seem to suggest that when σ_θ is the only available turbulence data at an onsite station, ADJ_U* should in fact be applied, as the AERMOD results will otherwise be likely to produce the over-predictions of concentrations for ground level sources that are found when modeling without any observed turbulence data. Figure 7 shows a Q-Q plot with ADJ_U* applied to the onsite "Truth" data for the simple terrain ground level source case that, when compared to the same plot in Figure 5 that did not include ADJ_U*, shows that performance of both WRF and Airport data is comparable to the onsite "Truth" data if ADJ_U* is applied to the onsite data.

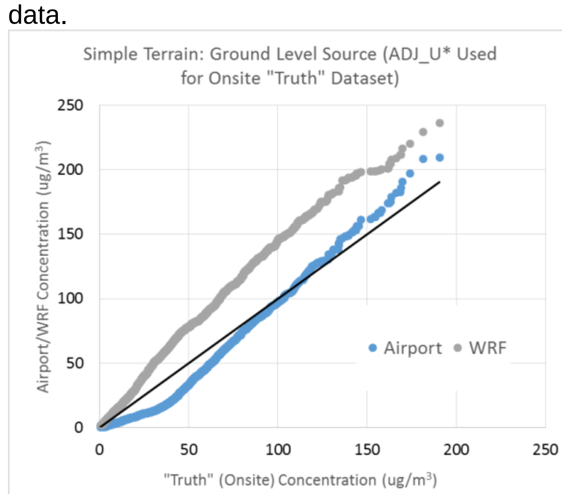


Fig. 7. Q-Q plots for 1-hour concentrations resulting from a ground level source in the simple terrain case, with ADJ_U* applied to the onsite ("Truth") meteorological dataset.

5. REFERENCES

U.S. EPA, 2016: User's Guide for the AERMOD Meteorological Preprocessor (AERMET). EPA-454/B-16-010. Online: https://www3.epa.gov/ttn/scram/7thconf/aermod/aermet_userguide.pdf

U.S. EPA, 2016: Guidance on Use of the Mesoscale Model Interface Program (MMIF) for AERMOD Applications. EPA-454/B-16-003. Online: https://www3.epa.gov/ttn/scram/appendix_w/2016/MMIF_Guidance.pdf

U.S. EPA, 2016: Evaluation of Prognostic Meteorological Data in AERMOD Applications. EPA-4354/R-16-004. Online: https://www3.epa.gov/ttn/scram/appendix_w/2016/MMIF_Evaluation_TSD.pdf

U.S. EPA, 2017: Revisions to the Guideline on Air Quality Models: Enhancements to the AERMOD Dispersion Modeling System and Incorporation of Approaches To Address Ozone and Fine Particulate Matter. 40 CFR§51, Appendix W. Online: https://www3.epa.gov/ttn/scram/appendix_w/2016/AppendixW_2017.pdf