Army's Night Vision Integrated Performance Model (NVIPM)

NVIPM is not validated.

The current Future Command modelers have never explained why the Targeting Task Performance (TTP) metric accepted by the Army to replace the Johnson Criteria has been dropped. Without explanation, without any supporting experiment data, without citing any theory, and without being open and transparent about their actions, the current Army modelers have dropped the validated TTP and substituted the different metric in NVIPM.

We know from participants evaluating the NVIPM metric that validation failed. True, the Future Command refuses to publish those results, so "proof" does not exist. However, the fact that no data supporting NVIPM has ever been published is a certain indicator of NVIPM failure. Now, journal articles are appearing (8-11) claiming that the validation of the original TTP, the validated TTP, actually supports the failed NVIPM. How is that possible?

Although there are a number of problems with NVIPM, the most serious is the re-definition of Contrast Threshold Function (CTF). The original TTP used the definition accepted by the vision science community; CTF is the measured sine wave threshold using enough sine wave periods to get a "single frequency" threshold value. Vision scientists have found that at least seven (7) sine wave periods and perhaps as many as thirteen (13) are needed to get an accurate value.

The Army Future Command now use a modified CTF that depends on the angular size of the target on the display. Whatever number of sine wave periods fits on the viewed target, that is the number that determines threshold value. Why? The reason has never been disclosed by the Future Command modelers.

After the change, all attempts at validating the modified TTP failed. The only recourse for the current Army modelers is to falsely claim that the original validation of the TTP using the accepted CTF definition now applies to their mutilated CTF. That false claim is now widely accepted by a misled technical community.

Einstein said: ***"Explanations should be as simple as possible, but no simpler."***

We will try to explain what the Future Command has done.

Discussion outline:

- TTP metric fundamentals and validation.
- What the Future Command modelers changed.
- Validation data, or the lack thereof, for NVIPM.

The validated logic of the original TTP metric

The original TTP is a frequency domain model that compares eye threshold at each frequency to the spatial frequency content of the target at each spatial frequency. See Reference 7. Eye threshold is measured by vision scientists and called the Contrast Threshold Function (CTF). A single frequency is a sine wave with infinite periods, but vision scientists have discovered that between seven and eleven sine wave periods are enough to get an accurate threshold measurement.

Also, most researchers measuring CTF ensure that the eye is adapted to the display luminance, so even given that eleven sine wave periods are presented, the display luminance is made the same as the sine wave pattern over a large area.

The CTF of several people with good eyesight are measured one at a time at left below and the average CTF used to represent the average soldier at right. That is, we represent the CTF of the soldier at right with the measured sine wave thresholds. We have a typical observer with good eyesight.

The soldier's eyes adapt to the display even though he is fixating the tank. Luminance adaptation does not occur only on a fixated object. The eye adapts to the display luminance, and the display luminance establishes the threshold contrast values at each spatial frequency.

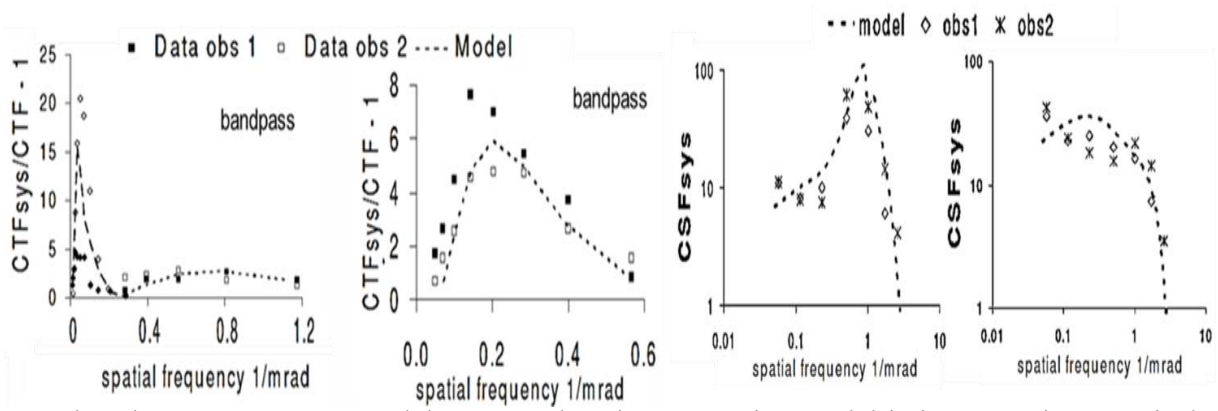
The tank is represented by the Fourier Transform of the tank or a set of vehicles is represented by a Normal Distribution C_{TGT}. TTP compares tank frequency content to frequency threshold of the eye. C_{TGT} is target contrast, CTF_{sys} is CTF adjusted for imager noise and blur, ε is spatial frequency, low and cut are frequencies where C_{TGT} is less than CTF_{sys}.

$$\theta = \int_{\xi_{low}}^{\xi_{cut}} \sqrt{\frac{C_{TGT}}{CTF_{sys}(\varepsilon)}}\, d\varepsilon \qquad PID = \frac{2}{\sqrt{\pi}} \int_{0}^{\Theta/\Theta 84} e^{-t^2}\, dt$$

θ84 = task difficulty parameter

Note, after initially using what is referred to as a Target Transform Probability Function (TTPF) that curve fits PID predictions to metric values, we finally realized that the error function (erf) relates PID to metric value if the IQM is actually valid.

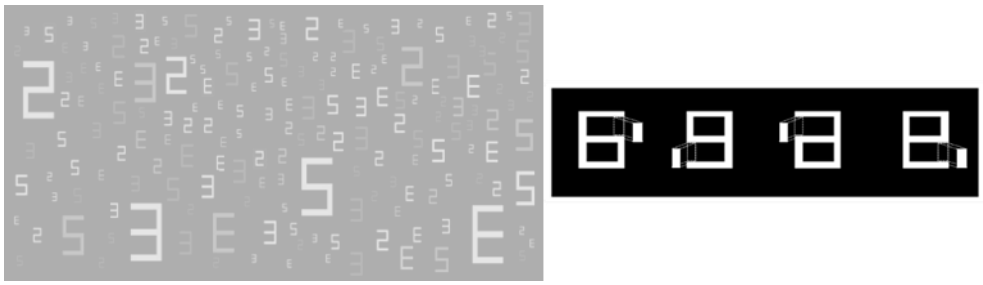
The importance of CTF_{sys} is obvious, as is the importance of getting the mathematical description of C_{TGT} right. Since the TTP metric uses the standard definition of CTF, the literature contains experiment data taken by independent researchers of the effect of noise on CTF of the naked eye. We compared our CTF_{sys} model to the data of independent researchers, and some of the results are shown here. See [7] for details and additional examples.



Note that the current Army modelers state that the TTP noise model is incorrect because it does not predict NVIPM CTF behavior.. Of course not; NVIPM uses a non-standard definition of CTF. We have validated our model to standard CTF data. The statements by current Army modelers that our noise model is incorrect are baseless and outlandish.

We get naked eye CTF using Barten's numerical CTF algorithm [13]. Barten's algorithm is used because an independent researcher[14] says it is accurate and easy to use, as long as the display size and luminance are entered correctly. Our selection of Barten's CTF numerical generator was because it provides CTF of young people with good eyesight and is easy to use. Our selection of CTF algorithm had nothing whatsoever to do with Barten's IQM work. Any suggestion otherwise is an attempt to distract and obscure the fact that NVIPM has no foundation in experiment or theory. None.

Consider the two character sets in the next figure. The current Army modelers do not seem to understand that these two types of target sets must be treated differently mathematically. The set on the left is represented in the metric calculation by a Normal Distribution with an average contrast, the set on the right is represented by the Fourier Transform of a rectangle. My point here is that these modeling professionals (sic) have muddled many things, not all of which will be covered in detail in this discussion. We call the character set at left diverse and the character set at right specific.

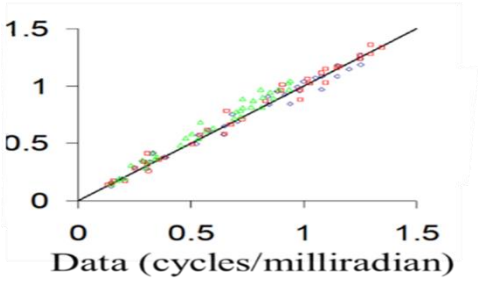


See **RHVElectro-Optics.com** for more details on the differences between NVIPM and TTP.

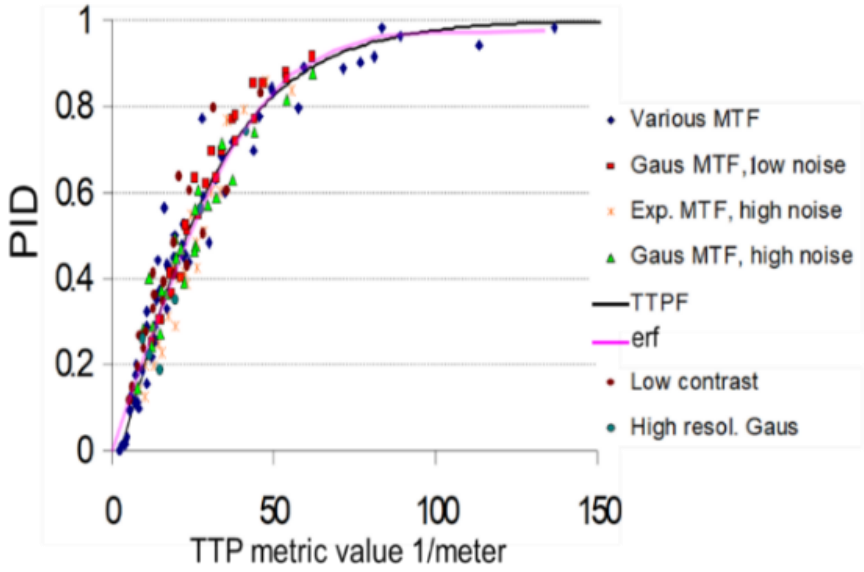
TTP Validation Summary

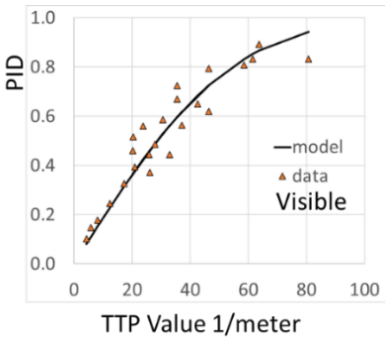
Experiment details are provided in the references.

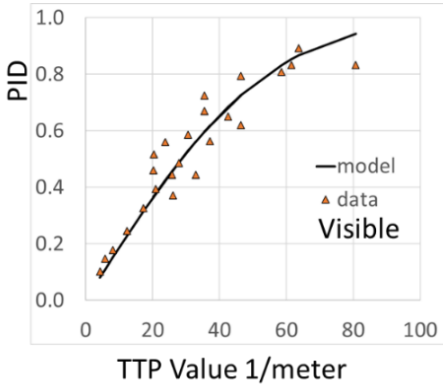
In [1-3], three experienced observers viewed Air Force 3-bar charts to establish limiting resolution of image intensifiers versus chart illumination. The experiment used chart contrasts of 1.0 and 0.4. Data were collected with and without laser protective eyewear situated between the goggle and eye that reduced apparent luminance by a factor of ten. Light to the eye varied from as little as 3.4E-4 fL to 35 fL. Chart illuminance varied from 2.88E-6 to 3.39E-3-foot candles, and that variation in illumination means that the image varied from noise to resolution limited.

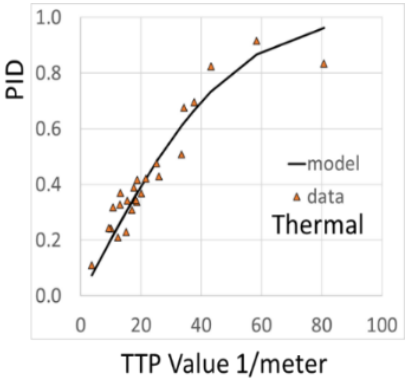


The figure below compares the original TTP PID predictions to data collected with various shape blurs, contrasts, and noise levels. The subsequent figures show predictions for colored noise. See [1-4].

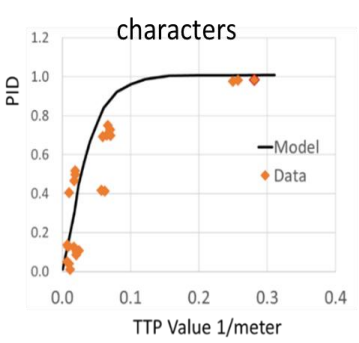


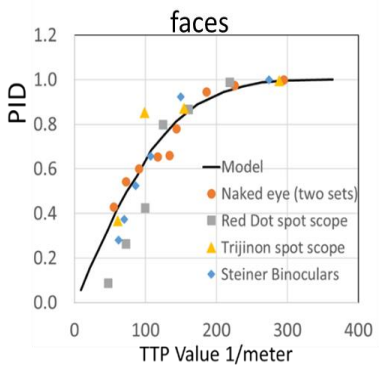


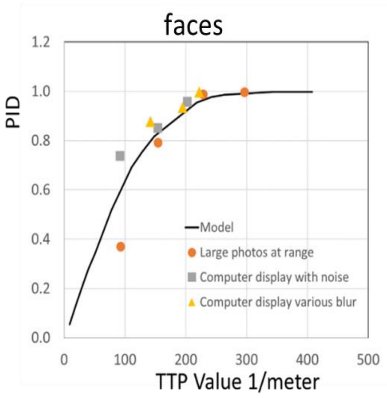


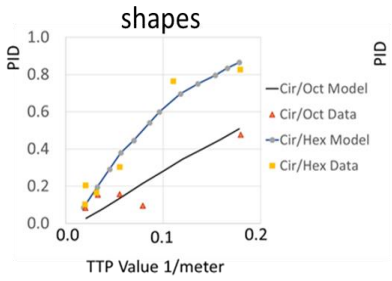


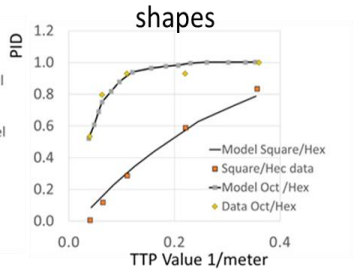
The graphs below show predictions for characters, faces, and shapes. We are particularly impressed with predicting performance through spotting scopes and binoculars. The current modelers have thrown away the performance in the graphs above and below. Why?





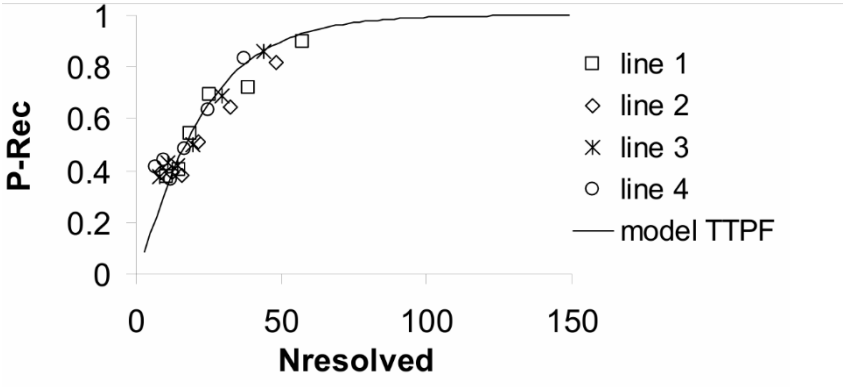






Cir is Circle, Hex is Hexagon
Oct is Octagon

The next graph shows recognition of armored tracked versus armored wheeled versus wheeled trucks [3].



The next two graphs show field test results using tracked tactial vehicles. Details are in [3]. Field results are more accurate for reasons described in [25].

NVThermIP modeled thermal imagers and was part of a set of models that included an image intensifier model. Over the last forty years, Army aviators have flown with imager intensifiers (ANVIS) and both first and second generation thermal imagers [24]. The table shows TTP

predictions for pilotage utility and the results of pilot surveys. TTP predicts aviator experience with pilotage sensors. This data for the original TTP and not NVIPM.

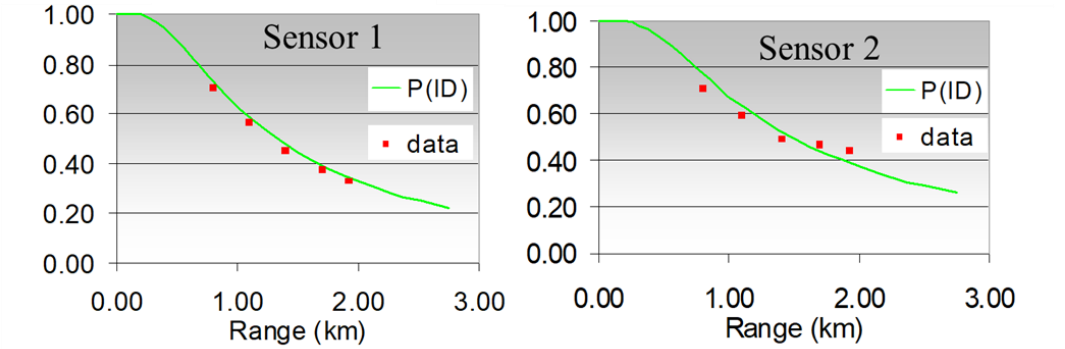
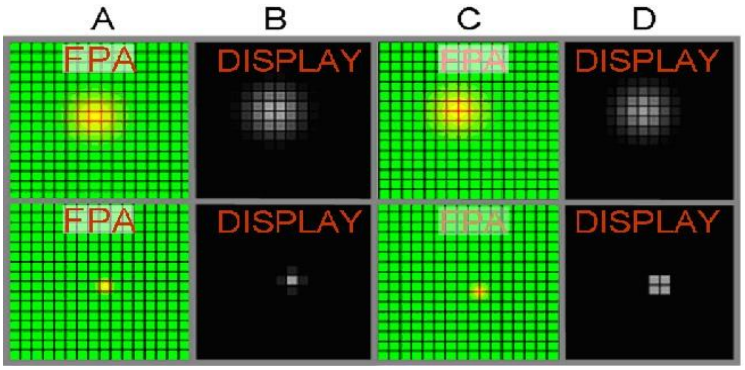


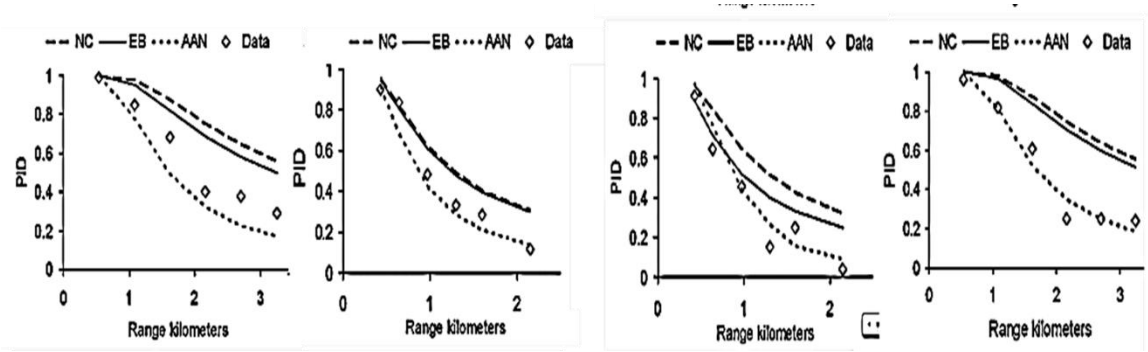
Table 1. Aviator assessment of imager performance.

Sensor	Condition	TTP	Pilot assessment
ANVIS	overcast starlight	1.05	Poor
ANVIS	starlight	3.3	Fair
ANVIS	Quarter Moon	5.8	Good
1st Gen thermal	0.1 K scene	1.0	Poor
1st Gen thermal	1.0 K scene	2.6	fair
1st Gen thermal	5.0 K scene	3.1	fair
2nd Gen thermal	0.1 K	2.4	fair
2nd Gen thermal	1.0 K	5.9	Good
2nd Gen thermal	5.0 K	6.7	Good

Alaising has been treated wrong since I and others made a mistake in the first sampling experiments. I have not been able to get my coauthors to agree on the mistake, but the problem is obvious and the solution published [15,16]. It is not possible to believe that aliasing does not corrupt a picture, given any significant amount of alsising, that is.

In the picture below at top, a big blur is moved from the center of one pixel to the intersection of four pixels, and the display result changes little, At bottom, a small optical blur is moved and the change in display obvious. The idea that this behavior cannot affect PID is simply wrong. When and how much can be calculated using the algorithms in [15,16]. The incorrect model used in NVIPM is laeled EB and the correct model AAN. NC means no corection. The grpahs are exampes of TTP AAN prediction accuracy, and more examples are presented in [15,16].





Regardless of the many successful validation efforts, the Future Command modelers and their management (sic) have dropped the TTP metric. They have never provided any explanation other than a few words in NVIPM documentation where they present a misleading representation of facial PID data.

NVIPM Validation Summary

The data presented above under the title "TTP" does not apply to NVIPM even though they ursurp the TTP name. None of the references [1-7,15,16] refer to NVIPM even though some peer reviewed journal articles would suggest otherwise [8-11]. The complete lack of anything to discuss here is the result of the Future Command unwillingness to publish what they have learned, that NVIPM is not an accurate predictor of anythng. All we can say is NVIPM is unvalidated, and that negative statement is "proved" by the lack of evidence in support of NVIPM.

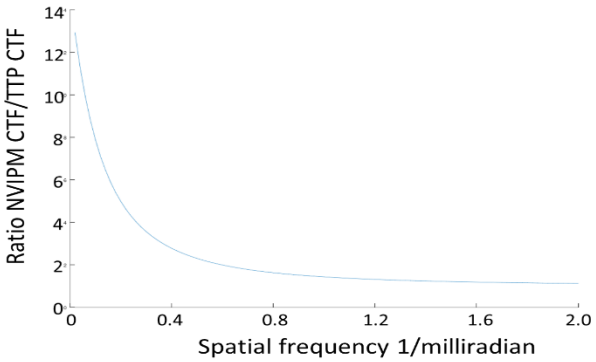
The Optical Engineerng paper [9] requires some discussion, simply because the reviewers and editors let the authors claim things in the conclusion unsupported by anything in the paper.

First, the CTF experiment description clearly demonstrates the non-standard CTF used by NVIPM. The authors state:

> "The CTF is measured at each spatial frequency, ξ, using a displayed sinusoidal stimulus pattern with an amplitude (Lp), a mean luminance (L0), **and a stimulus size defined by the apparent target angle (w) equal to the square root of the target area in degrees."**

A one meter target at 1 kilometer viewed through a 10X spotting scope would subtend 0.57 degrees. The figure shows the ratio of NVIPM CTF to TTP CTF (TTP CTF is used by the vision community). Note that at low spatial frequencies there are few sine wave periods in 0.57 degrees and threshold are an order of magnitude high. The threshold of NVIPM CTF decreases as spatial freqquency increases because there are more cycles to look at.

The point of the ratio of CTF plot is that metric calculation is very different for NVIPM than TTP. The larger the target is on the display, the closer the NVIPM calculation is to the TTP calculation.
.



Why NVIPM uses the non-standard CTF definition is not discussed in [9] or anywhere else.

The CTF experiment described in [9] is riddled with problems, but the authors do not claim that the experment has any relationship to NVIPM, and they dismiss the CTF results themselves. Given that the authors of [9] have not tied the work to NVIPM and have not used the results in any way, it seems like describing our many problems with the CTF experimet simply wastes time.

The authors published the results of three of eightyone TTP experiments. They have the PID data on all eightyone but have published the results for three. Why not credit at least those three experiments as validation even though the errors are larger than the TTP predictions? Because all three presented the targets with a four degree target angle. Those three experiments presented the targets large enough on the display that the NVIPM CTF approaches the standard CTF. That is, the current Army modelers have published only those NVIPM results where the model approaches the TTP model. That selection of experiments makes our point, not theirs.

There is an unsupported statement in [9] that the original TTP has trouble with adjusting to different target angles. Totally untrue. There are all kinds of target angles considering not just [4] experiments but [4-7], [15,16], and the field test results over range. Again, the current Army modelers make simple, declarative statements unsupported by any discussion or examples. There are many other such statements in [9] and various conference articles over the years. Simple, unsupported statements, all of which ignore TTP validation literture.

The authors of [9] describe a two-hand-hand-held PID experiment where they use the PID data to calibrate a new empirical constant. The authors admit in the body of the paper that there are unresolved issues with the noise model, but the point here is that no validation data was presented. In the [9] conclusion, once again, simple declarative statements about how validated their model is.

The NVIPM CTF_{sys} model is complete nonsense and, of course, no data is or can be presented in support of that model. The Future Command has published no validation of the NVIPM predictions, although Optical Engineering editorial staff and reviewers have let them make unsuppported statements in the [9] conclusions. Read the paper, it is not hard, there is nothing there, and the confusing part is that the authors actually admit to serious, unsolved problems with the NVIPM model, but the reviewers still let the future Command modelers claim success in the conclusion section.

Conclusions

There is a long and successful history of TTP validation including predicting the CTF in noise data of independent researchers, predicting PID of target sets of tactical vehicles, faces, characters, and shapes. The fact that the erf function relates metric value to PID is an indicator of model viability as is predicting the assessment of aviators about the performance of pilotage aids.

NVIPM does not use the TTP metric. We understand that the Army calls the metric in NVIPM "TTP." Doing otherwise would separate NVIPM for any semblance of legitimacy.

The Army has never explained the transition of their model away from TTP. They have published no supporting data, no theory, and no explanation. At this point, it appears there is no stopping the dishonest destruction of valid research by a disinterested and lazy Army bureaucracy.

Certainly Future Command management (sic) will do nothing, and clearly the current Army modelers are getting their way with everyone in a position to do anything. Why is that? It appears that no one really cares about these models or understands their value, but what caliber of techncial community knowingly lets good work languish and sloppy work thrive?

1. R. VOLLMERHAUSEN, “INCORPORATING DISPLAY LIMITATIONS INTO NIGHT VISION PERFORMANCE MODELS,” IN IRIS PASSIVE SENSORS, PROCEEDINGS OF THE INFRARED INFORMATION SYMPOSIUM, ENVIRONMENTAL RESEARCH INSTITUTE OF MICHIGAN, VOL. 2, PP. 11–31. 13, (1995).

2. ELECTRO-OPTICAL IMAGING: SYSTEM PERFORMANCE AND MODELING, LUCIEN M. BIBERMAN, ED., CH 12, “MODELING THE PERFORMANCE OF IMAGING SENSORS,” RICHARD VOLLMERHAUSEN, SPIE PRESS, BELLINGHAM (2000).

3. R. H. VOLLMERHAUSEN, E. JACOBS, J. HIXSON, AND M. FRIEDMAN, “THE TARGETING TASK PERFORMANCE (TTP) METRIC: A NEW MODEL FOR PREDICTING TARGET ACQUISITION PERFORMANCE,” TECH. REP. (DTICAMSEL-NV-TR-230, U.S. ARMY RESEARCH, DEVELOPMENT AND ENGINEERING COMMAND, COMMUNICATIONS AND ELECTRONICS RESEARCH, DEVELOPMENT AND ENGINEERING CENTER, NIGHT VISION AND ELECTRONIC SENSORS DIRECTORATE, 10221 BURBECK RD., FT. BELVOIR, VA 22060-5806 (2005).

4. R. H. VOLLMERHAUSEN, E. JACOBS, AND R. G. DRIGGERS, “NEW METRIC FOR PREDICTING TARGET ACQUISITION PERFORMANCE,” OPT. ENG. (BELLINGHAM) 43, 2806–2818 (2004).

5. R. VOLLMERHAUSEN AND A. L. ROBINSON, “MODELING TARGET ACQUISITION TASKS ASSOCIATED WITH SECURITY AND SURVEILLANCE,” APPL. OPT. 46, 4209–4221 (2007).

6. RH VOLLMERHAUSEN, S MOYER, K KRAPELS, RG DRIGGERS, JG HIXSON, “PREDICTING THE PROBABILITY OF FACIAL IDENTIFICATION USING A SPECIFIC OBJECT MODEL,” APPLIED OPTICS 47 (6), 751-759 (2007).

7. RICHARD H. VOLLMERHAUSEN, “REPRESENTING THE OBSERVER IN ELECTRO-OPTICAL TARGET ACQUISITION MODELS,” OPTICS EXPRESS 17 (20), 17253-17268 (2009)

8. BRADLEY L. PREECE, JEFFREY T. OLSON, JOSEPH P. REYNOLDS, JONATHAN D. FANNING, “IMPROVED NOISE MODEL FOR THE US ARMY SENSOR PERFORMANCE METRIC,” PROC. OF SPIE VOL. 8014 801406 (2011)

9. BRADLEY L. PREECE, JEFFREY T. OLSON, JOSEPH P. REYNOLDS, JONATHAN D. FANNING, AND DAVID P. HAEFNER, “HUMAN VISION NOISE MODEL VALIDATION FOR THE U.S. ARMY SENSOR PERFORMANCE METRIC,” OPTICAL ENGINEERING 53(6), 061712 (2014)

10. ROBERT SHORT, DUKE LITTLEJOHN, JOHN BAILEY, AND RONALD DRIGGERS, ”INFRARED SENSOR PERFORMANCE WITH BOOST AND RESTORATION FILTERING,” APPLIED OPTICS, VOL. 60, ISSUE 3, PP. 571-579 (2021)

11 RONALD DRIGGERS, GLENN GORANSON, STEVE BUTRIMAS, GERALD HOLST AND ORGES FURXHI, ’SIMPLE TARGET ACQUISITION MODEL BASED ON FΛ/D,” OPTICAL ENGINEERING, VOL. 60(2), 023104-1 (2021

12. RICHARD H. VOLLMERHAUSEN, “NIGHT VISION INTEGRATED PERFORMANCE MODEL: IMPACT OF A RECENT CHANGE ON THE MODEL’S PREDICTIVE ACCURACY,” OPTICS EXPRESS, VOL. 24, NO. 21, 23654-23666 (2016)

13. P. G. J. BARTEN, “FORMULA FOR THE CONTRAST SENSITIVITY OF THE HUMAN EYE,” PROC. SPIE 5294, 231–238 (2004) (PAPER AVAILABLE ON THE WEB AT HTTP://WWW.SPIE.ORG).

14. R. J. BEATON, AND W. W. FARLEY, “COMPARATIVE STUDY OF THE MTFA, ICS, AND SQRI IMAGE QUALITY METRICS FOR VISUAL DISPLAY SYSTEMS,” ARMSTRONG LAB., AIR FORCE SYSTEMS COMMAND, WRIGHT-PATTERSON AFB, OH, REPORT AL, TR-1992–0001, DTIC ADA252116, (1991)

15. RICHARD H. VOLLMERHAUSEN, RONALD G. DRIGGERS, AND DAVID L. WILSON, “PREDICTING RANGE PERFORMANCE OF SAMPLED IMAGERS BY TREATING ALIASED SIGNAL AS TARGET-DEPENDENT NOISE,” J. OPT. SOC. AM. A/VOL. 25, NO. 8 (2000)

16. RICHARD H. VOLLMERHAUSEN, RONALD G. DRIGGERS, AND DAVID L. WILSON, “PREDICTING RANGE PERFORMANCE OF SAMPLED IMAGERS BY TREATING ALIASED SIGNAL AS TARGET-DEPENDENT NOISE: ERRATUM,” J. OPT. SOC. AM. A/VOL. 25, NO. 8 (2000)

17. RG DRIGGERS, R VOLLMERHAUSEN, B O’KANE, “EQUIVALENT BLUR AS A FUNCTION OF SPURIOUS RESPONSE OF A SAMPLED IMAGING SYSTEM: APPLICATION TO CHARACTER RECOGNITION,” J. OPT. SOC. AM. A, VOL. 16 (5) ,1026-1033 (1999)

18. R. A. MOSES AND W. M. HART, “THE TEMPORAL RESPONSIVENESS OF VISION,’ IN ADLER’S PHYSIOLOGY OF THE EYE: CLINICAL,” APPLICATION, MOSBY (1987)

19. N.S. NAGARAJA, “EFFECT OF LUMINANCE NOISE ON CONTRAST THRESHOLDS, J. OPT. SOC. AM. 54(7), 950-955 (1964)

20. D. G. PELLI, ‘‘WHY USE NOISE?’’ J. OPT. SOC. AM. A 16, 647–653 (1999)

21. M. RAGHAVAN, ‘‘SOURCES OF VISUAL NOISE,’’ PHD DISSERTATION, SYRACUSE, NY, (1987)

22. H. DAVSON, PHYSIOLOGY OF THE EYE, 5TH ED., MACMILLAN, LONDON (1990)

23. PETER G. . BARTEN, CONTRAST SENSITIVITY OF THE HUMAN EYE AND ITS EFFECTS ON IMAGE QUALITY, [PHD THESIS 1 (RESEARCH TU/E / GRADUATION TU/E), INSTITUTE FOR PERCEPTION RESEARCH, EINDHOVEN]. TECHNISCHE UNIVERSITEIT EINDHOVEN. HTTPS://DOI.ORG/10.6100/IR523072 (2005)

24. RH VOLLMERHAUSEN, T BUI, “USING A TARGETING METRIC TO PREDICT THE UTILITY OF AN EO IMAGER AS A PILOTAGE AID,” INFRARED IMAGI

25. RICHARD K. MOORE, EDDIE L. JACOBS, CARL E. HALFORD, “CELL BALANCING FOR VEHICLE IDENTIFICATION PERCEPTION EXPERIMENTS AND CORRECTING FOR CELL IMBALANCE IN TEST RESULTS,” PROC. OF SPIE VOL. 6543 65430D-6 , 2007.23. PETER G. . BARTEN, CONTRAST SENSITIVITY OF THE