Metamorphic Testing of Monte Carlo Modeling Program

Junhua Ding, Tong Wu, Dianxiang Xu,
Jun Q. Lu, Xin-Hua Hu

May 23rd, 2011

Non-testable Software System

- Software system without available test oracle
    - An oracle doesn't exist for a program
    - an oracle is potentially available, but the efforts to get the oracle are impractical

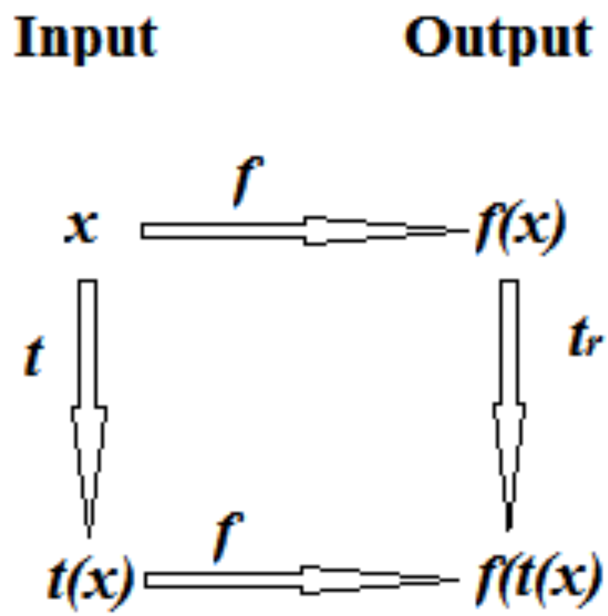
E. Weyuker, "On testing non-testable programs". The Computer Journal, 25(4), pp 465-470.

Test “Non-testable” System

- Dual coding
- Using unapproved results
- Using special cases

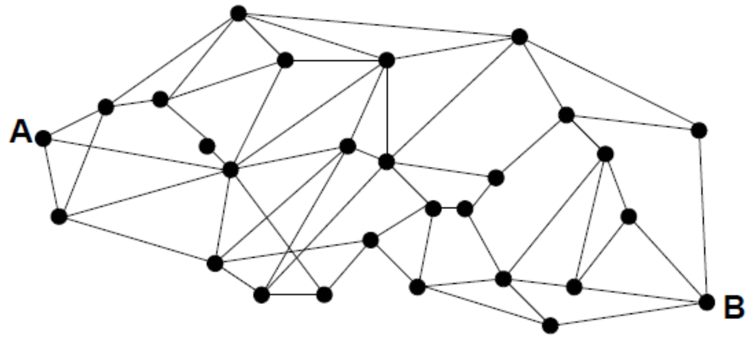
E. Weyuker, "On testing non-testable programs". The Computer Journal, 25(4), pp 465-470.

Metamorphic Testing



R1(x, t(x, f(x))) ➔ *R2(f(x), f(t(x, f(x))))*

Example of Metamorphic Testing



- $S_{AB} = \{P1, P2, \ldots, Pn\}$ and $n \geq 1$
- One of the expected MRs:
  $Pi \in S_{AB}$. Randomly select a vertex X in path Pi. Then, run the program to get the shortest paths from A to X (S_{AX}) and from X to B(S_{BX}). For any path $Q \in S_{AX}$ and any path $Q' \in S_{BX}$, the concatenation of Q and Q' must be an element in S_{AB}.

Z. Q. Zhou, D. H. Huang, T.H. Tse, Z. Yang, H. Huang, and T. Y. Chen "Metamorphic Testing and Its Applications," Proceedings of the 8th International Symposium of Future Software Technology, 2004.

Limitation of Metamorphic Testing

- How to ensure quality of metamorphic testing?
- How to find more useful information from satisfaction of a metamorphic relation?

What is Self-checked Metamorphic Testing

- The self-checked metamorphic testing is a metamorphic testing extended with evaluating the adequacy of testing coverage criteria during the test process.

Advantages

- Act as criteria to evaluate the identification of metamorphic relations. (need more MRs?)

- Act as criteria to evaluate the generation of test cases. (need more test cases?)

- Possibly increase fault-revealing rate.

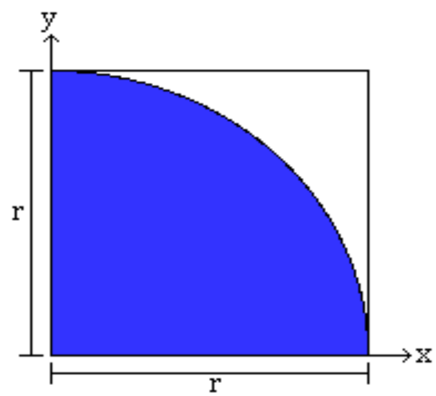
Perform Self-checked Metamorphic Testing

- Define metamorphic relations.
- Generate metamorphic testing inputs.
- Select test coverage criteria.
- Instrument code for checking test coverage criteria.
- Conduct testing.

What is Monte Carlo Simulation

- Monte Carlo method is a computerized mathematical technique that uses random numbers and probability statistics to solve problems.
- Any computational approach involves algorithms that contains random numbers and repeated sampling belongs to Monte Carlo method.

Example of Monte Carlo Simulation



$$\frac{\text{\# darts hitting shaded area}}{\text{\# darts hitting inside square}} = \frac{\frac{1}{4}\pi r^2}{r^2} = \frac{1}{4}\pi$$

or

$$\pi = 4\,\frac{\text{\# darts hitting shaded area}}{\text{\# darts hitting inside square}}$$

x=(random#)

y=(random#)

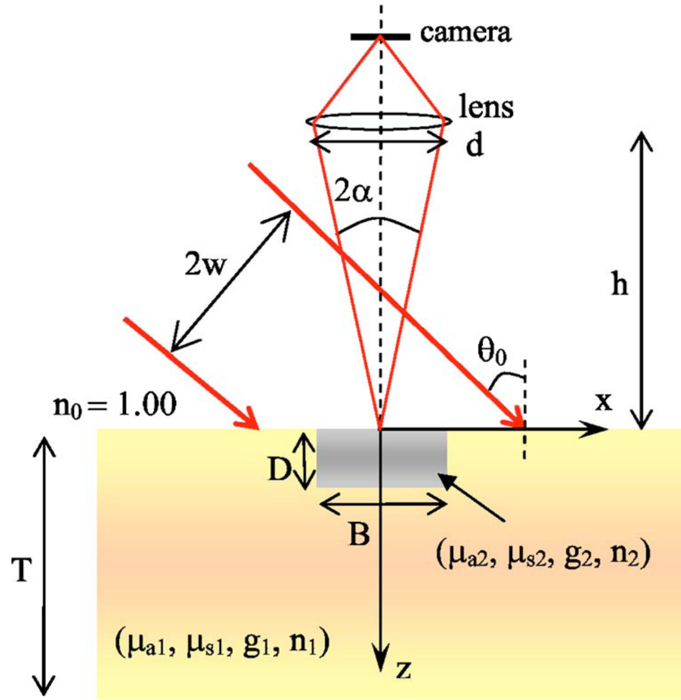
dist=sqrt(x^2 + y^2)

if dist.from.origin
(less.than.or.equal.to) 1.0

let hits=hits+1.0

http://www.chem.unl.edu/zeng/joy/mclab/mcintro.html

The biological model



- C. Chen, J.Q. Lu, K. Li, S. Zhao, R.S. Brock, X.H. Hu, “Numerical study of reflectance imaging using a parallel Monte Carlo method”, Medical Physics, vol. 34, pp. 2939-2948, 2007.

Validation of Monte Carlo Program Modeling Photon Propagation

- Homogenous Media
  - Beer-Lambert law
  - Van de Hulst table
  - RTE and approximation methods

- Heterogeneous Media
  - Special values, unapproved oracles

Program Structure

- There are 5 modules, about 40 subroutines or functions in the program, and the total lines of Fortran 90 code is about 1600.
  - Monte_main.f90 - the main program including the code calling the MPI functions.
  - Monte_go.f90 includes the subroutines and functions to check if the photons hit the different optical boundaries in the turbid medium and record current photon status and position.
  - Monte_sub.f90 is the module for all utility subroutines that do the calculation for the simulation.
  - Monte_io.f90 is the file for input/output subroutines.
  - Monte_define.f90 contains all the definitions for objects and constants.

General Information

- Two structural test coverage criteria, *function coverage* and *branch coverage*, were checked.

- Module *monte_go*, *monte_sub*, and *monte_io* were checked with structural coverage information.

Coverage Code Instrumentation

- Statements for evaluating the test coverage criteria are instrumented into the program under test so that test coverage information can be generated whenever a test input is executed.

```
subroutine HitOutSideCy(Pos_p,Dir,s,r,Reg,Hit,s_go,s_left)
IMPLICIT NONE
real(8), intent(in) :: Pos_p(3), Dir(3), s
real(8), intent(in) :: r, Reg(2)
real(8), intent(out) :: s_go, s_left
integer(1), intent(out) :: Hit
real(8) :: Pos(3), Pos2(3), dist(2),t1,t2,temp
   output%fc(1)=output%fc(1) +1
!change coordinates
Pos=Pos_p
if ((cy%Pos(1).NE.0.0).OR.(cy%Pos(2).NE.0.0)) then
  output%branch(1)=output%branch(1) +1
Pos(1)=Pos_p(1)-cy%Pos(1)
Pos(2)=Pos_p(2)-cy%Pos(2)
    ...
```

General Information

- 5 metamorphic relations (MR) are selected for the testing.
- For each metamorphic relation, at least two test sets with different inputs were examined to lower the possibility that satisfies the metamorphic relations accidentally by special inputs.
- For semi-infinite (only consider $z \geq 0$) heterogeneous media, which the parallel Monte Carlo code simulated, phantoms thickness is infinite ($T \rightarrow \infty$). In simulations, T was set to 100mm.
- For all the test cases in this study, we used the following parameters: w = 12.5 mm, d = 25 mm, FOV = 41.2 mm X 41.2 mm, and 201 X 201 grid cells on the surface.

Metamorphic Relations

- MR1: Contrast C value decreases when refractive index n2 value increases.
- MR2: Contrast C value decreases when anisotropy factor g2 value increases.
- MR3: Contrast C value increases when albedo α2 value increases, where $\alpha = \mu s/(\mu s + \mu a)$.
- MR4: For each pixel along the x axis P(x,0) on the image, the averaged reflectance R(x,0) will decrease if the numerical aperture (NA) decreases, where NA = sinα.
- MR5: For each pixel along the x axis P(x,0) on the image, the averaged reflectance R(x,0) will decrease if the incident light angle θ0 increases.

General Information

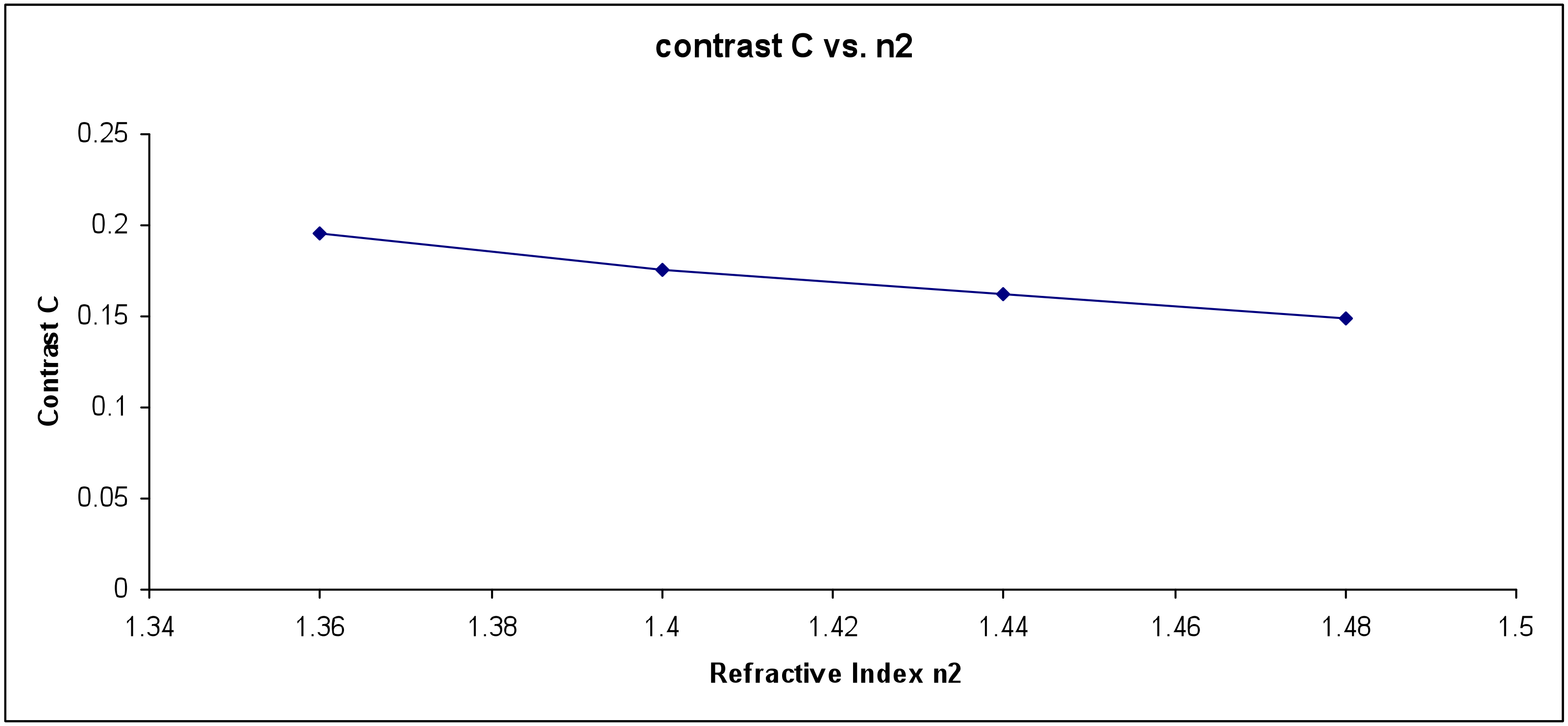
- Five metamorphic relations (MR1-5) have been selected. MR1- MR4 contain two test sets. MR5 has four test sets instead. Each test set made with four test cases.
- Although the numeric results of the simulations could be compared directly, the results were plotted for easier and quicker checking.
- F-go, F-sub, and F-io means the coverage of all functions or routines in module *monte_go*, *monte_sub* and *monte_io* respectively.
- B-go, B-sub, and B-io means the coverage of all branches in module *monte_go*, *monte_sub* and *monte_io* respectively.

Simulation Results of MR1

- MR1: contrast value decreases when *n2* value increases.
- Simulation results for test set HT-MR1-1

Test Case	**n_2**	**Contrast C**	**F-go**	**F-sub**	**F-io**	**B-go**	**B-sub**	**B-io**
HT-MR1-1-T1	1.36	0.19589	14 100%	14 100%	13 100%	126 75.9%	34 60.7%	27 90%
HT-MR1-1-T2	1.40	0.175588	14 100%	14 100%	13 100%	124 74.7%	34 60.7%	26 86.7%
HT-MR1-1-T3	1.44	0.162571	14 100%	14 100%	13 100%	127 76.5%	34 60.7%	26 86.7%
HT-MR1-1-T4	1.48	0.148713	14 100%	14 100%	13 100%	127 76.5%	34 60.7%	26 86.7%

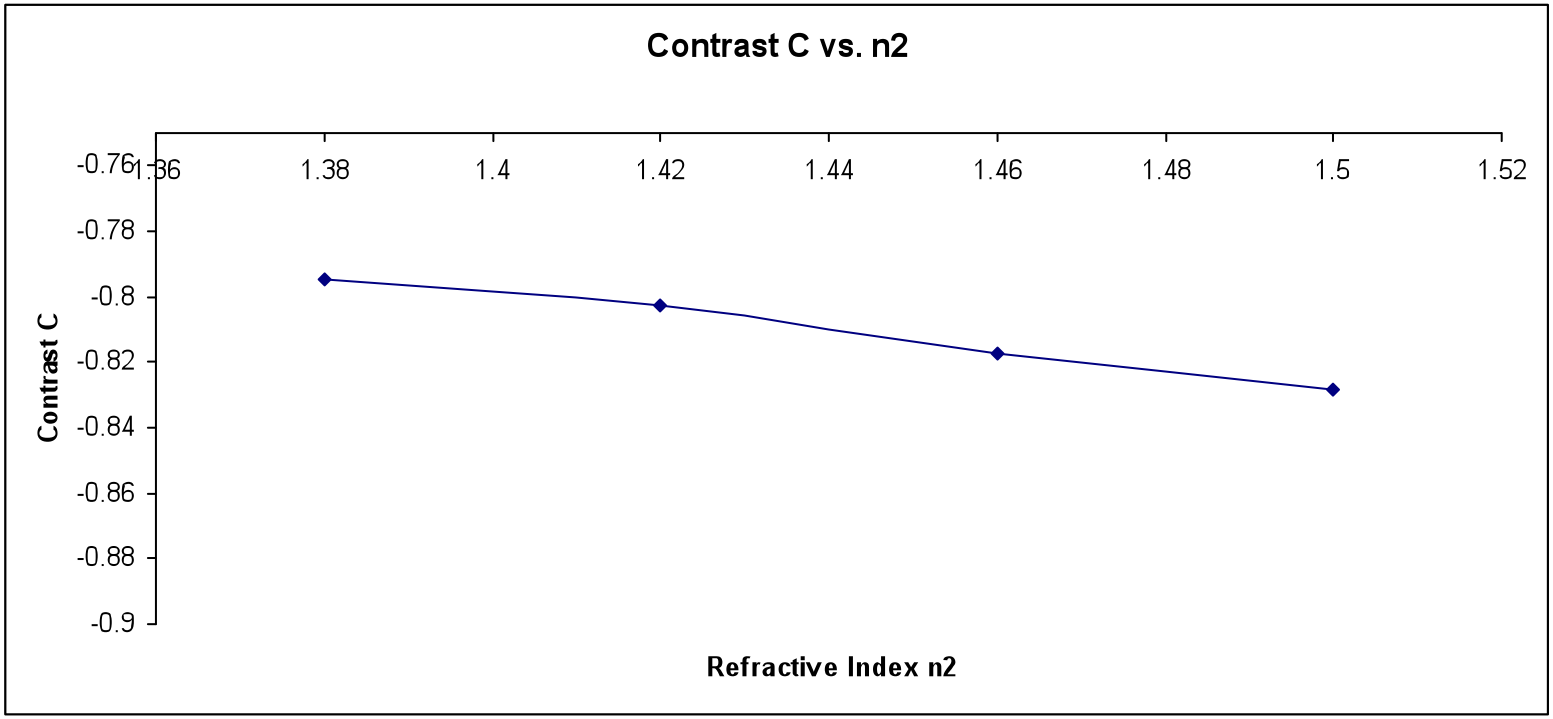
Simulation Results for Test Set HT-MR1-1(cont'd)



Simulation Results for Test Set HT-MR1-2

Test Case	n_2	Contrast C	F-go	F-sub	F-io	B-go	B-sub	B-io
HT-MR1-1-T1	1.38	-0.794871	14 100%	14 100%	13 100%	127 76.5%	34 60.7%	26 86.7%
HT-MR1-1-T2	1.42	-0.802847	14 100%	14 100%	13 100%	129 77.7%	34 60.7%	26 86.7%
HT-MR1-1-3	1.46	-0.817642	14 100%	14 100%	13 100%	129 77.7%	34 60.7%	26 86.7%
HT-MR1-1-T4	1.50	-0.82838	14 100%	14 100%	13 100%	129 77.7%	34 60.7%	26 86.7%

Simulation Results for Test Set HT-MR1-2 (cont'd)



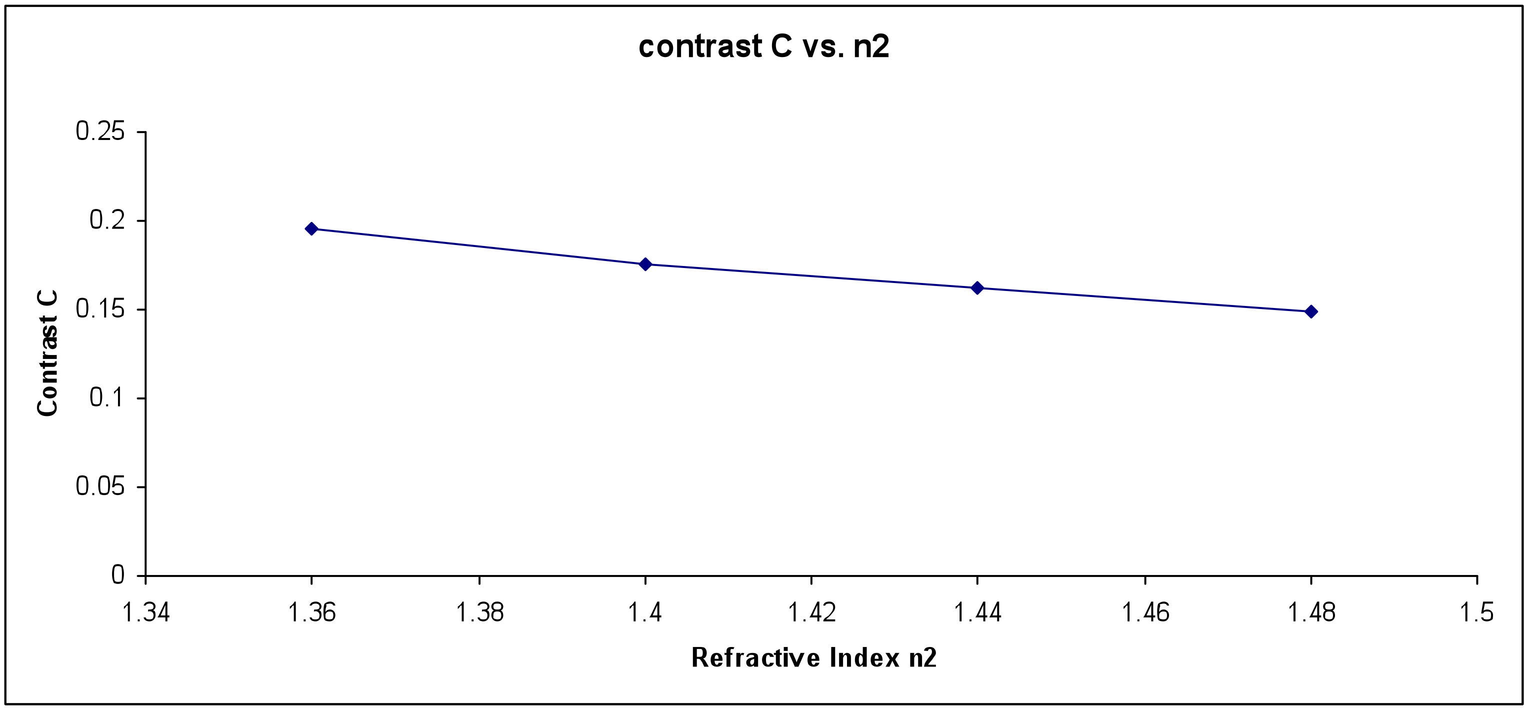
Simulation Results of MR2

- MR2: contrast value decreases when *g2* value increases.

Simulation results for test set HT-MR2-1

Test Case	g_2	Contrast C	F-go	F-sub	F-io	B-go	B-sub	B-io
HT-MR2-1-T1	0.1	0.557547	14 100%	14 100%	13 100%	124 74.7%	34 60.7%	26 86.7%
HT-MR2-1-T2	0.3	0.513292	14 100%	14 100%	13 100%	124 74.7%	34 60.7%	26 86.7%
HT-MR2-1-T3	0.6	0.391112	14 100%	14 100%	13 100%	124 74.7%	34 60.7%	26 86.7%
HT-MR2-1-T4	0.9	0.0907334	14 100%	14 100%	13 100%	125 75.3%	34 60.7%	26 86.7%

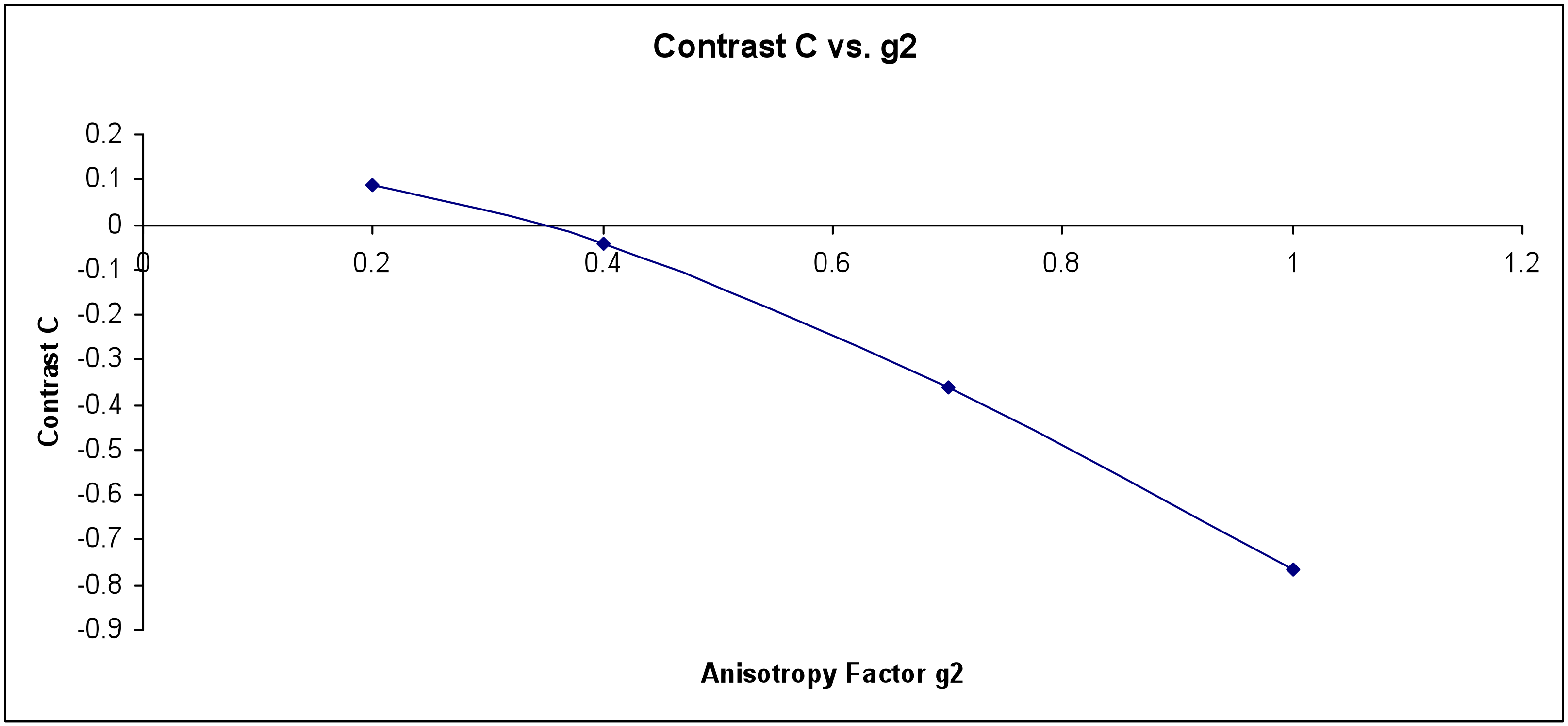
Simulation Results for Test Set HT-MR2-1 (cont'd)



Simulation Results for Test Set HT-MR2-2

Test Case	g_2	Contrast C	F-go	F-sub	F-io	B-go	B-sub	B-io
HT-MR2-2-T1	0.2	0.085616	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%
HT-MR2-2-T2	0.4	-0.042526	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%
HT-MR2-2-T3	0.7	-0.362704	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%
HT-MR2-2-T4	1.0	-0.767007	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%

Simulation Results for Test Set HT-MR2-2 (cont'd)



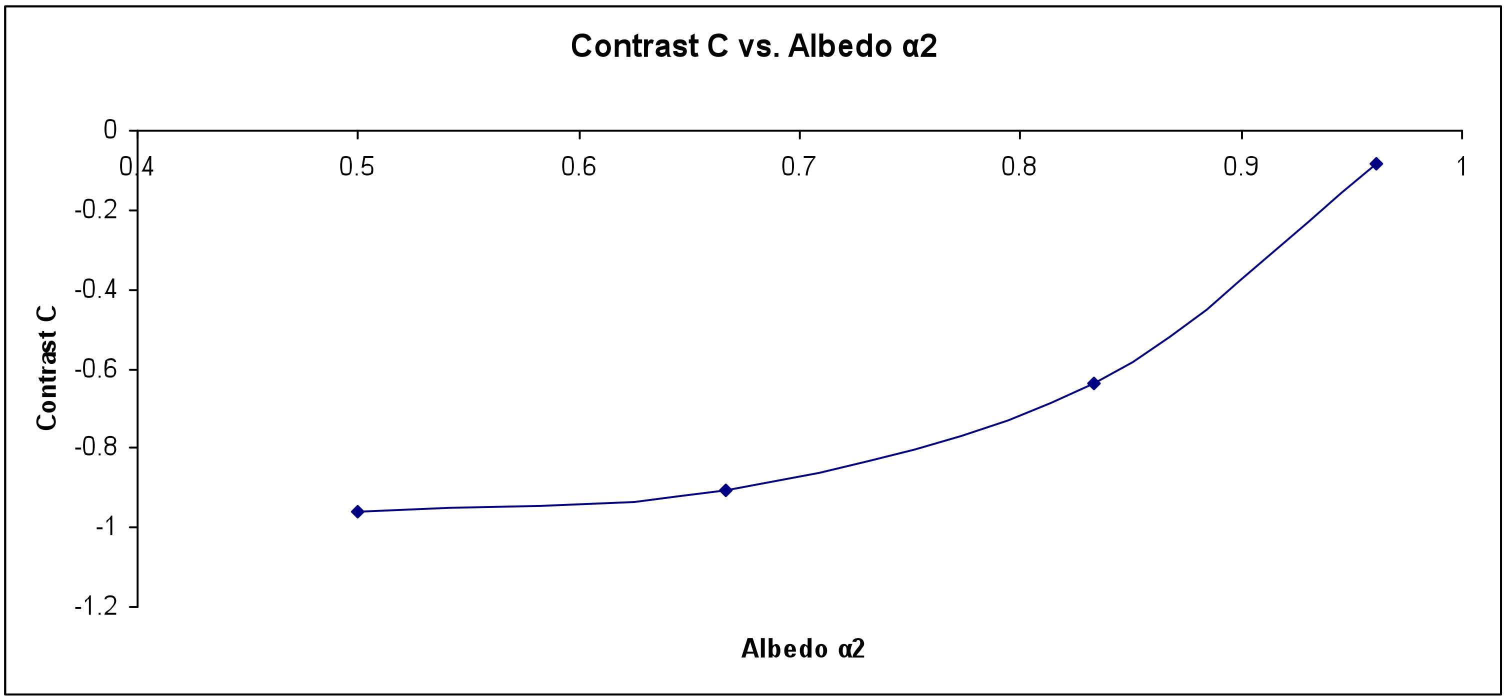
Simulation Results of MR3

- MR 3: contrast value increases when albedo α2 value increases.

Simulation results for test set HT-MR3-1

Test Case	μ_{a2}	α_2	Contrast C	F-go	F-sub	F-io	B-go	B-sub	B-io
HT-MR3-1-T1	0.02	0.9615	-0.0820255	14 100%	14 100%	13 100%	124 74.7%	34 60.7%	26 86.7%
HT-MR3-1-T2	1.00	0.8333	-0.637966	14 100%	14 100%	13 100%	124 74.7%	34 60.7%	26 86.7%
HT-MR3-1-T3	2.50	0.6666	-0.904051	14 100%	14 100%	13 100%	124 74.7%	34 60.7%	26 86.7%
HT-MR3-1-T4	5.00	0.50	-0.961337	14 100%	14 100%	13 100%	125 75.3%	34 60.7%	26 86.7%

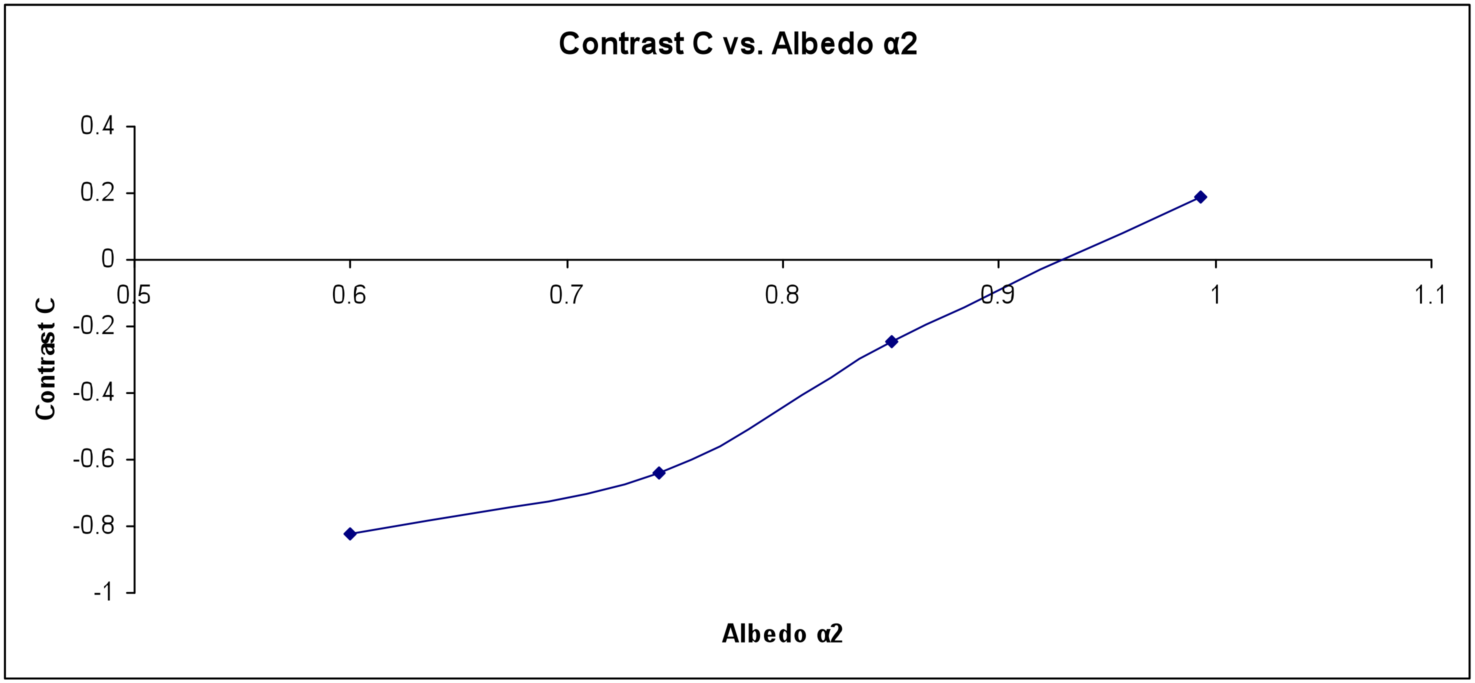
Simulation Results for Test Set HT-MR3-1 (cont'd)



Simulation Results for Test Set HT-MR3-2

Test Case	μ_{a2}	α_2	Contrast C	F-go	F-sub	F-io	B-go	B-sub	B-io
HT-MR3-2-T1	0.02	0.9933	0.185817	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%
HT-MR3-2-T2	0.50	0.8500	-0.247233	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%
HT-MR3-2-T3	1.20	0.7430	-0.639169	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%
HT-MR3-2-T4	2.00	0.60	-0.82205	14 100%	14 100%	13 100%	125 75.3%	37 66.1%	26 86.7%

Simulation Results for Test Set HT-MR3-2 (cont'd)



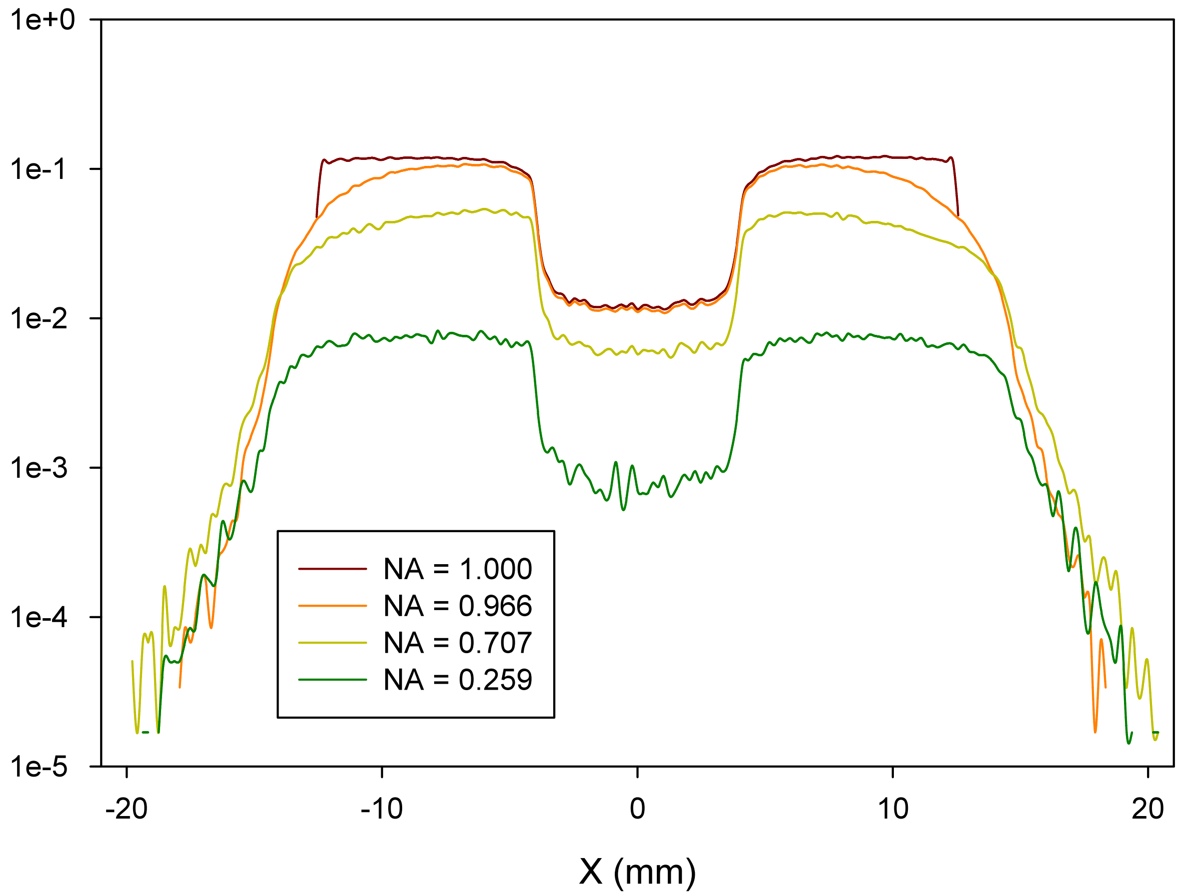
Simulation Results of MR4

- MR 4: for each pixel along the x axis P(x,0) on the image, will decrease if the numerical aperture (NA) decreases.

Coverage information for test set HT-MR4-1

Test Case	**Height (mm)**	**NA**	**F-go**	**F-sub**	**F-io**	**B-go**	**B-sub**	**B-io**
HT-MR4-1-T1	0.00	1.000	14 100%	14 100%	13 100%	122 73.5%	34 60.7%	26 86.7%
HT-MR4-1-T2	3.35	0.966	14 100%	14 100%	13 100%	122 73.5%	34 60.7%	26 86.7%
HT-MR4-1-T3	12.50	0.707	14 100%	14 100%	13 100%	122 73.5%	34 60.7%	26 86.7%
HT-MR4-1-T4	46.65	0.259	14 100%	14 100%	13 100%	122 73.5%	34 60.7%	26 86.7%

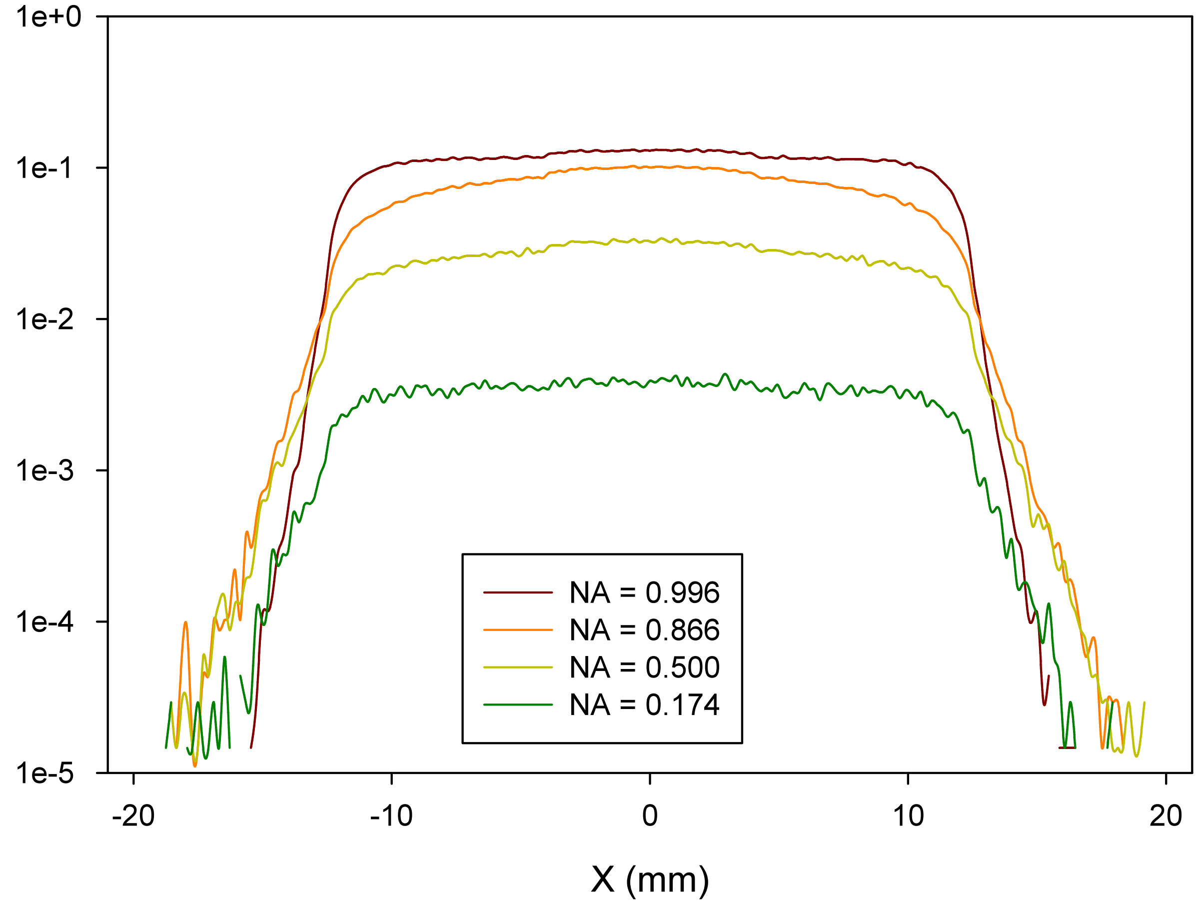
Simulation Results for Test Set HT-MR4-1



Coverage Information for Test Set HT-MR4-2

Test Case	Height (mm)	NA	F-go	F-sub	F-io	B-go	B-sub	B-io
HT-MR4-2-T1	0.00	0.996	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%
HT-MR4-2-T2	3.35	0.866	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%
HT-MR4-2-T3	12.50	0.500	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%
HT-MR4-2-T4	46.65	0.174	14 100%	14 100%	13 100%	124 74.7%	37 66.1%	26 86.7%

Simulation Results for Test Set HT-MR4-2



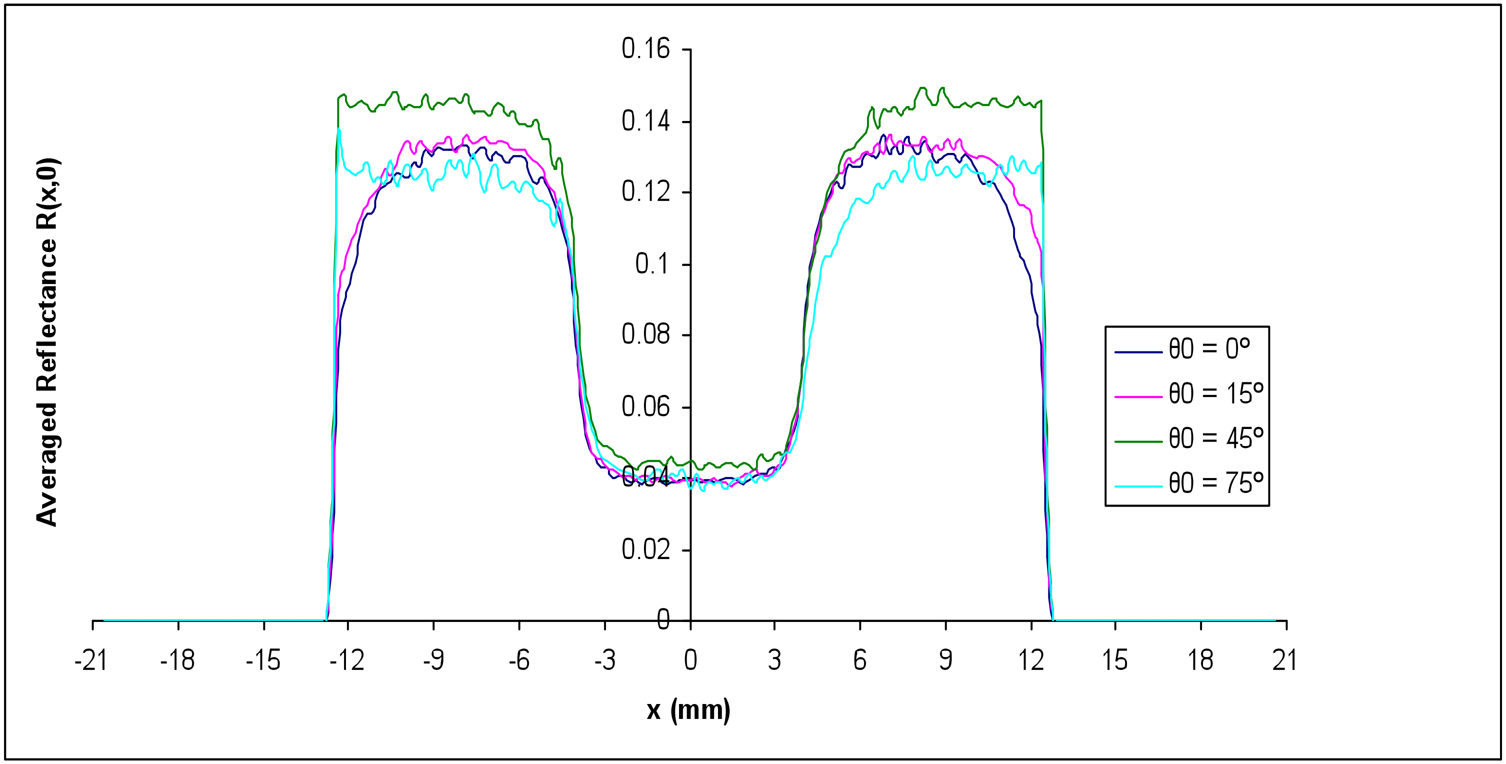
Simulation Results of MR5

- MR 5: For each pixel along the x axis P(x,0) on the image, the averaged reflectance R(x,0) will increase if the incident light angle θ0 increases.

Coverage information for test set HT-MR5-1

Test Case	θ_0 (º)	F-go	F-sub	F-io	B-go	B-sub	B-io
HT-MR5-1-T1	0	14 100%	14 100%	13 100%	124 73.5%	37 66.1%	26 86.7%
HT-MR5-1-T2	15	14 100%	14 100%	13 100%	124 74.7%	34 60.7%	26 86.7%
HT-MR5-1-T3	45	14 100%	14 100%	13 100%	124 74.7%	34 60.7%	26 86.7%
HT-MR5-1-T4	75	14 100%	14 100%	13 100%	124 74.7%	34 60.7%	26 86.7%

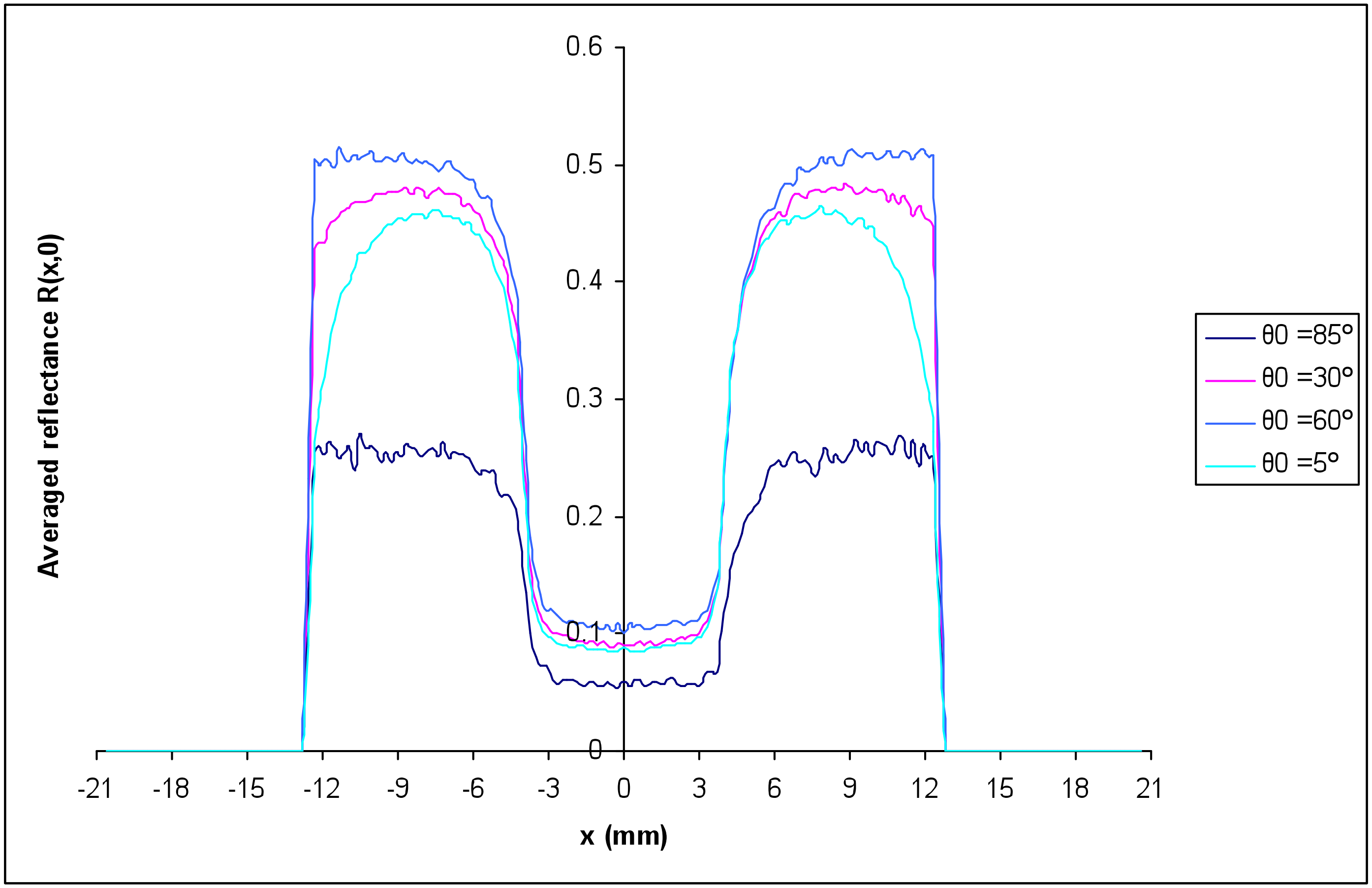
Simulation Results for Test Set HT-MR5-1



Coverage Information for Test Set HT-MR5-2

Test Case	θ_θ (º)	F-go	F-sub	F-io	B-go	B-sub	B-io
HT-MR5-2-T1	5	14 100%	14 100%	13 100%	123 74.1%	36 64.3%	26 86.7%
HT-MR5-2-T2	30	14 100%	14 100%	13 100%	122 73.5%	34 60.7%	26 86.7%
HT-MR5-2-T2	60	14 100%	14 100%	13 100%	122 73.5%	34 60.7%	26 86.7%
HT-MR5-2-T2	85	14 100%	14 100%	13 100%	122 73.5%	34 60.7%	26 86.7%

Simulation Results for Test Set HT-MR5-2



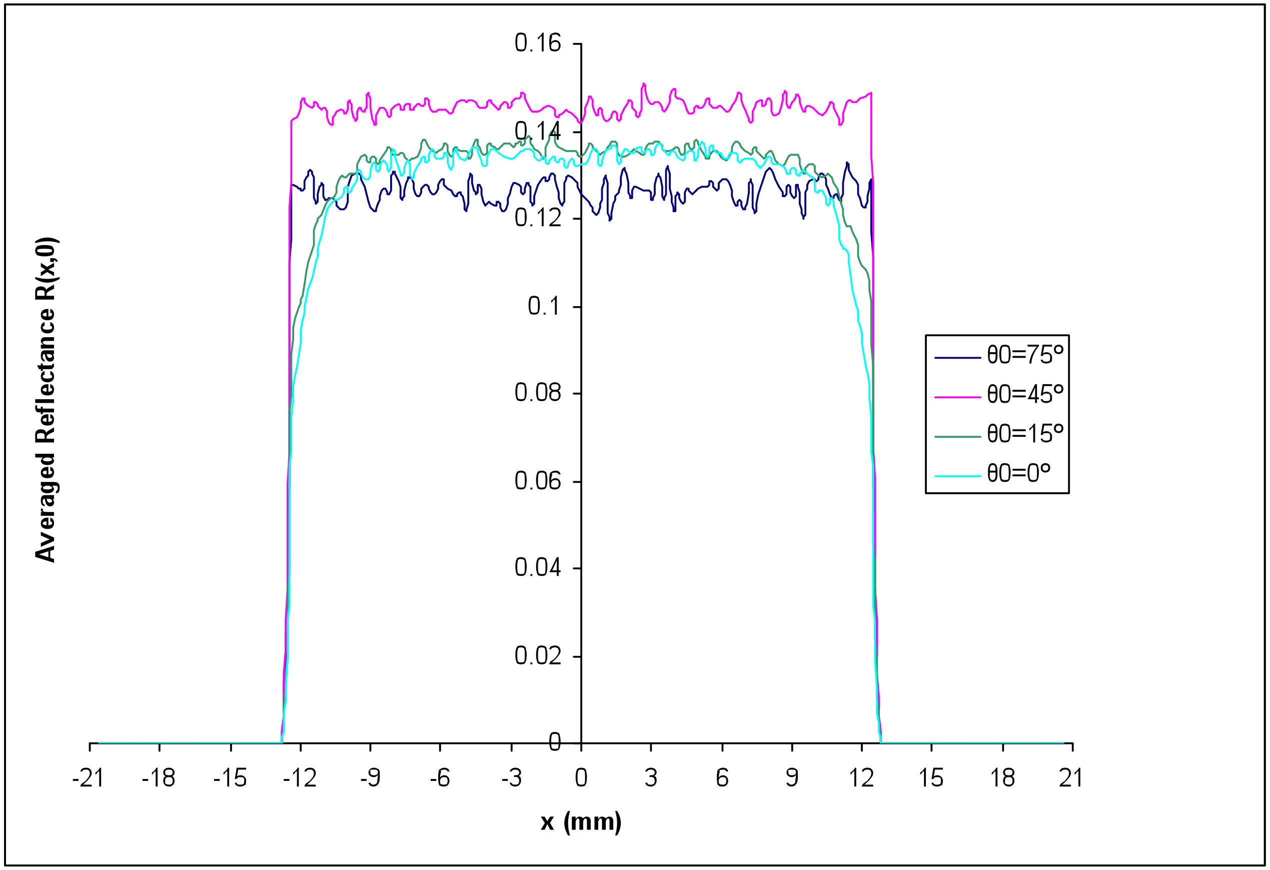
- A quick check shows that the results of both test sets for MR5 could not hold MR5.
- More test sets were created to test MR5.
- If MR5 holds for heterogeneous media, then it should hold for homogenous media as well.
- We tested MR5 for homogeneous media with test sets HM-MR5-1 and HM-MR5-2 to look at whether MR5 is satisfied in the special cases.

Simulation Results of MR5 (cont'd)

Coverage information for test set HM-MR5-1

Test Case	θ_0 (º)	F-go	F-sub	F-io	B-go	B-sub	B-io
HM-MR5-1-T1	0	11 78.6%	14 100%	13 100%	56 33.7%	35 62.5%	26 86.7%
HM-MR5-1-T2	15	11 78.6%	14 100%	13 100%	55 33.1%	32 57.1%	26 86.7%
HM-MR5-1-T3	45	11 78.6%	14 100%	13 100%	55 33.1%	32 57.1%	26 86.7%
HM-MR5-1-T4	75	11 78.6%	14 100%	13 100%	55 33.1%	32 57.1%	26 86.7%

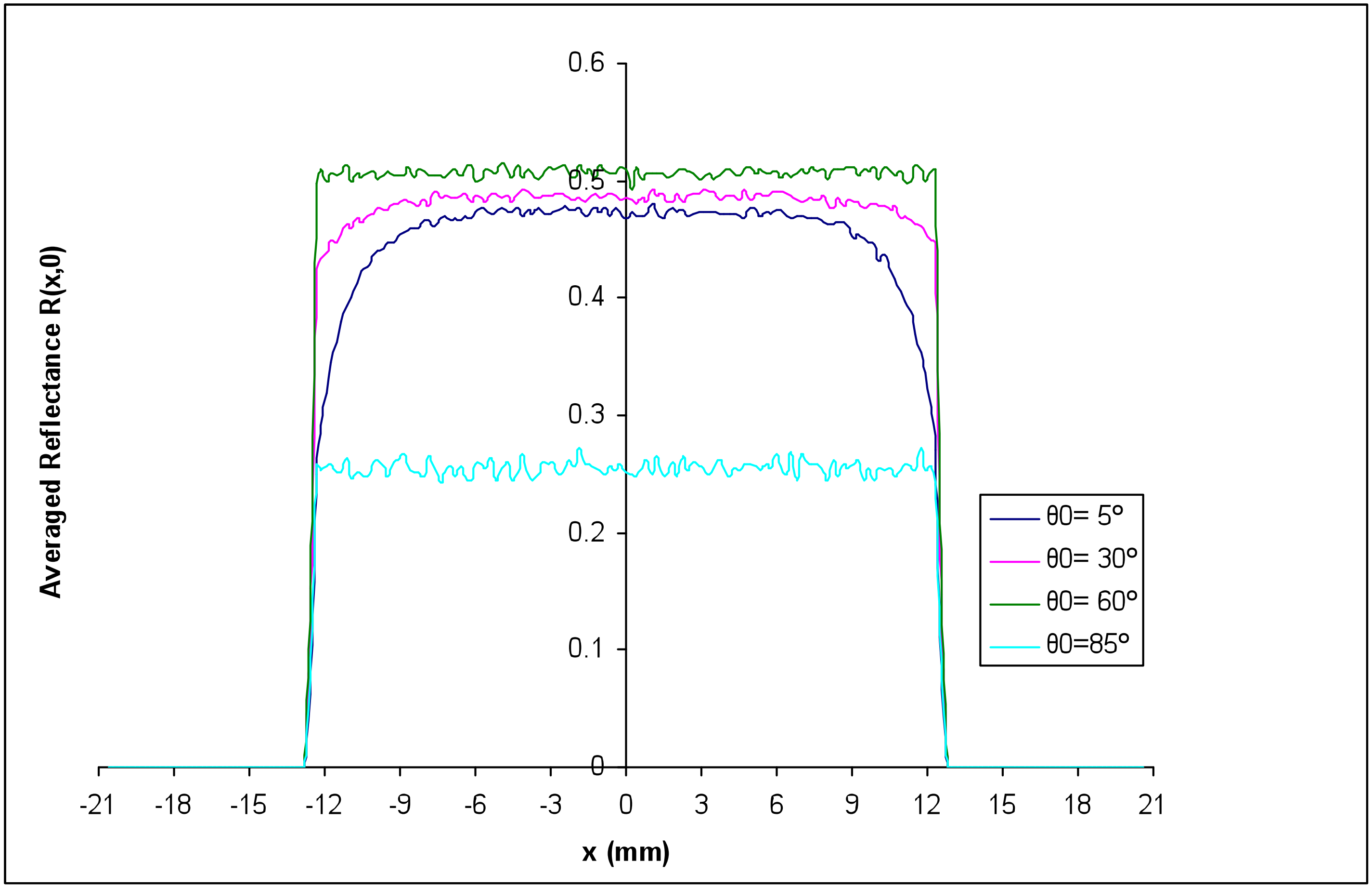
Simulation Results for Test Set HM-MR5-1



Coverage Information for Test Set HM-MR5-2

Test Case	θ_0 (º)	F-go	F-sub	F-io	B-go	B-sub	B-io
HM-MR5-2-T1	5	11 78.6%	14 100%	13 100%	55 33.1%	32 57.1%	26 86.7%
HM-MR5-2-T2	30	11 78.6%	14 100%	13 100%	55 33.1%	32 57.1%	26 86.7%
HM-MR5-2-T3	60	11 78.6%	14 100%	13 100%	55 33.1%	32 57.1%	26 86.7%
HM-MR5-2-T4	85	11 78.6%	14 100%	13 100%	55 33.1%	32 57.1%	26 86.7%

Simulation Results for Test Set HM-MR5-2



Discussion

Self-Checked Metamorphic Testing

MR1 to MR4

- For metamorphic relation MR1 to MR4, simple visual inspections can show that all the inputs and outputs fit the corresponding metamorphic relation very well for the whole eight test sets and 32 test cases.
- No faults was revealed.

Problem with MR5

- For metamorphic relation MR5, all of our test sets showed there was no simple linear change between averaged reflectance *R(x,0)* and *x*, which is different to MR5.
- However, MR5 could be satisfied for some special test cases. This observation further confirms that it is important to evaluate the quality of the test cases and their corresponding metamorphic relations.

Problem with MR5 (cont'd)

- The program for the homogenous media is much simpler than the program for the heterogeneous media.
- Results for both programs have the same patterns, we can narrow the defect code to the program for homogenous media.
- Several experienced programmers independently inspected the program for homogenous media, we could not find the faults.
- From physics theory, we cannot conclude whether MR5 should hold or not.
- We could not find the original version of the code that was experimented for the results in reference. But the latest version of the program does not support MR5.
- Build a real experimental environment to test MR5 could help. However, it is very complex to build. We are working on this now.

Coverage Information

- All functions are easily covered.
- No any test set covers 100% branches.
- The source code was underwent a static inspection with the help of coverage information.
- Most of the code, which has not been covered by the test cases, is used to handle extreme system configurations and the errors happening during the simulation.

Evaluate Test Cases

- If more complex test coverage criteria are considered, then the coverage data can be even more useful.

- The similar coverage data may indicate simply increasing the number of test cases for each metamorphic relation probably have little help to find more faults.

Summary

- The metamorphic testing is extended with checking adequacy of test coverage criteria: *function coverage* and *branch coverage*.
- The adequacy of the test coverage criteria is chosen as the metamorphic testing requirements and it is also served as a guideline for creating metamorphic relations, generating test cases, and finding errors in the program.
- The effectiveness of our approach has been investigated through testing a Monte Carlo modeling program.