I know what you’re thinking...
We’re in Las Vegas, we could’ve had Celine Dion!
My WSC street cred

- Attended every WSC since 1983 (70%).
- Helped drive *Proceedings* to WSC ’83, wrecking the van in the hotel garage.
- Arrested at WSC ’84 for bringing “liquid refreshment” into the hotel for the TIMS College on Simulation Mixer.
- Worried Secret Service at WSC ‘89 by running at Sen. Richard Lugar on the Mall.
- Pasted 2470 running heads & page numbers on the WSC ‘91 *Proceedings* by hand (my arms & wrists still hurt).
- Broke WSC ’97 budget by forgetting $\text{Correlation}[\text{number of tracks, attendance}] > 0$.
- So Ernie, what could possibly go wrong?
In the beginning...

...there was the 1967 “Conference on the Applications of Simulation Using GPSS” which became the WSC we know today.

- This year we have a great history track that celebrates the mythology & legends of WSC.
- But let’s take a moment to salute GPSS:
  - GPSS = General Purpose Simulation System
  - Came free with IBM mainframes.
  - Had an integer clock, a uniform distribution, you drew the network diagram by hand, and you ran it from cards.
  - But without GPSS we might be at a machine learning conference today.
GPSS!

  - Written by the famous Tom Schriber!

Case study 2D: One-line, one-server queueing system with feedback.
Fun facts! (about the first conference)

- Registration was $30 in advance, $35 at the door.
- Cost for a hotel night was $18 (in New York!).
- 401 attendees.
- Conference had to borrow $700 to get off the ground.
- There were no Proceedings.
- All the organizers were smart, all the attendees were good looking, and all the talks were above average.
- I believe we are still here because of great organization, great talks, great cooperation and great relevance.
Great organization

Board of Dictators

- IBM SHARE
- ASA
- ASIM
- ACM/SIGSIM
- IEEE/SMCS
- IISE
- I-Sim
- NIST
- SCS

General Patton

Program Nerd

All-Business Chair

Publicity Hound

Proceedings Lackeys

Track Nerd-ettes
Great talks

- “Only Wet Babies like Change”
  2003 keynote
- “The ‘Tell-Us-the-Answer-You-Want’ Problem”
  1983
- “Better Simulation with COBOL”
  Fake news
- “A Simulation of the Operations of an Illegal Numbers Game”
  1974
- “What to Simulate First: The Chicken or the Egg?”
  Fake news
- “Linda Arouses a Sleeping Barber”
  2002
More great talks

- “Simulation, Drugs and Rock & Roll”
  
  Fake news

- “When is a Satellite not a Toaster?”
  1991

- “But, Mr. President - Is It ETHICAL?”
  1986

- “Lose all the Variance You Want without Being (data) Hungry”
  Fake news

- “The Daughter of Celia, the French Flag, and the Firing Squad”
  1973

- “Training Socialist Managers by Enterprise Simulation”
  1974
Great cooperation

\[
E_\psi [g(X)] = \int g(x) \frac{\psi(x)}{\varphi(x)} \varphi(x) \, dx
\]

```python
>>> def car(env):
...     while True:
...         print('Start parking at %d' % env.now)
...         parking_duration = 5
...         yield env.timeout(parking_duration)
...         print('Start driving at %d' % env.now)
...         trip_duration = 2
...         yield env.timeout(trip_duration)
```

SAS® Simulation Studio
With great RELEVANCE comes great responsibility

- WSC 2013 Titan’s Talk: “The Simulation Curmudgeon”
  1. Why do we build simulation models as if they will never change?
  2. Why do we treat simulation like poor man’s queueing theory?
  3. Why do we fit input distributions like it’s 1922?
  4. Why can’t I talk to my simulation?

  1. Simulation analytics
  2. Parallel simulation
  3. Simulation to support decisions

The World has changed more than WSC has

- **Then...**
  - Batch runs on a single processor with limited memory and storage.
  - Simulation imitates queueing theory: long-run average performance of stationary systems for design.
  - Data gathered as needed.
  - Modeling & simulation viewed as a technical competency.

- **Now...**
  - Rent 1000s of processors; store and search vast quantities of results.
  - Support needed for real or near-real-time decisions, possibly made algorithmically, for operations.
  - Data gathered all the time.
  - “Analytics” viewed as a core competency.
To remain relevant, we need better answers to harder problems in a more timely way.
Some challenges...

1) All models are wrong, but some models are wronger
2) Simulation: the glitter or the glue?
3) All the world’s a database
4) Mutually assured relevance
CHALLENGE #1

All models are wrong, but some models are *wronger*
George Box said...

- “Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful?”
  - From his book *Empirical Model-Building and Response Surfaces*
- We often quote this to justify what we do.
  - But did we take the wrong message from Box?
- “… how wrong do they have to be to not be useful?”
  - Box’s statement is not a blanket license to use models.
  - “Better answers” requires a better idea about “how wrong.”
Simulation = Inputs + Logic

- The **inputs** are the lowest level of unexplained variation.
  - Modeled by distributions, often estimated from data.

- Everything else is **logic**: the “art” part that is true-enough or not, right level or wrong level.
  - Tocher (1963) *The Art of Simulation.*
Thinking about “how wrong”

- **Old Idea:** Common random numbers and estimating differences.

- **New Idea:** Model risk & inference to the resolution of the model.

All models are wrong ...
All models are wrong ...

Old school simulation modeling
Inputs & Logic revisited

- The Sloan Valve video is really just a big dataset.
- Instead of (younger) me with a notepad, the “logic” could be “learned” from dissecting the video.
  - Fixed objects, dynamic objects, what follows what and with what regularity.
  - Analysis of hours of video would reveal inconsistencies, worker differences, rare events, etc. that I would never observe.
  - When “logic” is like a statistical model, then I can infer “how wrong!”
- Did I just throw us under the machine learning bus?
Is simulation just “machine learning?”

- **Question**: What is the essential difference?
- **Answer**: The meaning of $x$.
  - We observe at most **one** $x$.
  - We need to **model** what happens at new $x$’s.

\[
\text{minimize } \sum (\beta^T f(x_i) - y_i)^2
\]
Is simulation just “machine learning?”

**Question:** What is the essential difference?

**Answer:** The meaning of $x$.
- We observe at most one $x$.
- We need to model what happens at new $x$’s.

$$\text{minimize } \sum \{ \beta^T f(x_i) - y_i \}^2$$

All models are wrong …
“Learning” our simulations

- Statistical models can be learned from data.
  - Push as much of the simulation model as sensible to being an “input.”

- But we are interested in more than the observable I-O relationship.
  - Embedded in the data is a control $x$ that we want to change.

- The “art” part is what a change in $x$ will do to the I-O relationship.

- Our value add:
  - Modeling focused on the impact of changes.
  - Fully recognizing uncertainty to generate robust, defensible solutions.
Tweets from our future: 2029

President Justin Timberlake
@TheNSyncPrez
Dudes at Federal Emergency Modeling & Simulation Agency (FEMSA) iden
Vegas power grid before blackouts happen. Better than @TomCru
“Minority Report.”

President Justin Timberlake
@TheNSyncPrez
Check out @TheRealTShriber podcast
“Inside learning simulations: How it works
and why it matters.” #SchriberRules
CHALLENGE #2

Simulation: the glitter or the glue?
Think BIG

- Many of society’s most important challenges — healthcare, global terrorism, income inequality, world food supply, power distribution, unfiltered tweets — are systems-of-systems problems.
  - What is our role in solving these “harder” problems?

- Strategies that seem doomed to fail:
  - One big, fully detailed model of everything.
  - A simple, stylized model that provides “insight.”

- “Models of models” seem to be required.
  - Should simulation be the glitter or the glue?
An argument for glue...

- We naturally account for **time**.
- Level of detail can be **variable** and **our choice**.
- **Uncertainty** matters; if it can be modeled then we can include it.
- We do **realizations**, and realizations **parallelize**.
- We have a history of **combined** discrete-event, continuous-state, agent-based, and hybrid modeling.
- **Simulation is arguably the least restrictive modeling & analysis paradigm.**
Cool proof-of-concept: IBM Splash

- “The Splash project provides a prototype platform for combining existing heterogeneous simulation models and datasets to create composite simulation models of complex systems, thereby facilitating cross-disciplinary modeling, simulation, and optimization.”

- The Splash philosophy is to take the world as it is (i.e., lots of individual models) and to solve the assembly problem.

Why are we “not yet ready for prime time”

- We are not good at policy optimization or system control.
- We typically think in terms of a single-resolution model and one time scale, that tries to get everything right.
- Computation does matter (“computationally tractable”).
- Our dominant model-building paradigm does not easily scale up to large, data-driven, parametric, multi-resolution models.
  - Watch closely as the keynote speaker now commits professional suicide.
Drag & drop ‘til you drag & drop

- Drag and drop with embedded animation has been good to us, but going forward it is a bottleneck.
  - Glitter!

- More data-driven input and structure
  - Relationships that build the models, not just models that build the relationships

- Models easily created, updated & explored
  - Easy to change; easy to vary
  - Reference models rather than analysis models

- Models connected to the real world
  - Call out to the world & “symbiotic simulation”

- Aggregate & disaggregate time, space and data
  - Input and output to multiple models
  - “Believe your limit theorems” — PW Glynn

- Retain as much I/O data as possible

- But Python is not the answer either.
Tweets from our future: 2042

 breaks news! @GlobalSim shows...
CHALLENGE #3

All the world’s a database
What’s wrong with this? A thought experiment

- Think about something *important* that you would simulate.
  - Big is bad; small is good.
- You run some number of replications of one scenario.
- You look at the average and a histogram of performance.
- What is the first question that comes into your mind when you see *this*?
  - “Why did that happen?”
    - Do we actually ask this question?
    - Would we know how to find the answer?
Why the world loves “the machine”

- *Conditional statements* are useful for making money:
  - **Unconditional**: Customers spend $87.12 on average vs.
    - **Conditional**: A university professor from the Midwest spends $1.35 on average

- “Big Data” is less about *quantity* and more about *coverage* of a space of *covariates*.

- Unconditional has been ok because *design* is more forgiving than *operations*.
  - For macro changes the “main effects” matter, but control is about “dynamics.”

- If we want to be taken *seriously* in the future, then we have to take our *simulation-generated data seriously* too.
Taking simulation data seriously

- Last year Titan Susan Sanchez talked about **data farming**:
  - Experiments to explore **huge** factor spaces of complex systems.
  - Recognizes our differences from “field data” analytics:
    - We can **design** the experiment to **grow** the data.
    - High-performance computing ≠ Infinite computing.
    - And we can go back for more.
Making simulation runs you don’t “need”

- **Proposition:** One purpose of a simulation is to fill a database with detailed, relevant, dynamic observations of systems that do not yet exist to be mined for insight & solutions as needed.

- **Translation:** Simulation everywhere, all the time.

- Same decision made repeatedly in evolving conditions.
  - Forecast-driven designs

- Very expensive simulations covering a big factor space.
  - Low-discrepancy/low-uncertainty designs

- Want to decide if real-world experiments are worthwhile.
  - Designs that mimic real experiments.

- Detect trouble before it happens.
  - Designs that seeks danger close to where we are now
Tweets from our future: 2052

Schriber Institute
@SimMeSimYou

Trending: World archive of simulated data exceeds archive of real data for the first time. #MineTheGap

5:11 PM · 4 Feb 2052
133 RETWEETS 1,170 LIKES

Schriber Institute
@SimMeSimYou

Trending: Tune your iBrain implant to “Brontobyte simulation: How it works and why it matters” by @TheRealTSchriber

1:15 PM · 24 May 2052
181 RETWEETS 2,470 LIKES
CHALLENGE #4

Mutually assured relevance
Benchmarks

- **Optimization** and **Statistics** are two communities that are very similar to us.
- Yet the connections seem better for them.
- Why?
## Statistics

- **The competitive advantage comes from providing deeper and more-robust insight than others can provide.**

- **Implications**
  - Need to implement near the cutting edge.
  - Worth employing methodologists; worth it to methodologists to work with data & software.
  - Intensive computation is OK.

| Not concerned with the **creation of data**. |
| Vendors compete with free options that have extensive and continually evolving public libraries to provide better answers. |

<table>
<thead>
<tr>
<th>SAS</th>
<th>Minitab</th>
<th>Python</th>
<th>R</th>
</tr>
</thead>
</table>
Optimization

- A universally agreed upon definition of "the problem" allows the model to be separated from the solver.

Implications
- Competition is on better solvers for harder problems.
- Connection to latest OR & CS research is a competitive advantage.
- Intensive computation is expected.

Optimize $f(x)$ such that $x \in C$
What about the good guys (simulation)?

- The focus has been on generating the **right data** (including animation) not analysis of a common model.

- **Researchers** pretend that...

- **Users** understand modeling takes time, but can be impatient if the simulation itself does.

- It is a competitive advantage for **vendors** to differentiate based on better modeling.
It takes a SimCity

- “Better answers to harder problems in a more timely way” won’t happen separately in each group.
  - **Researchers**: Machine learning research involves real (i.e., hard) data sets, but we don’t often work with real simulations.
  - **Vendors**: R, Python, etc. succeed because “better answer” has become more important than “easy to use.”
  - **Users**: Ad hoc modeling & simulation methods really can produce misleading results that underachieve your objectives.

- The good news is that all the key players are at WSC, every year.
Tweets from our future: 2067

WinterSim.org  
@WinterSim100

After 99 years, WSC eliminates tracks! “Distinctions without a difference,” says WSC Gen Chair Emma Page.
2:35 PM · 1 Dec 2067

359 RETWEETS  3,122 LIKES

WinterSim.org  
@WinterSim100

Tom Schriber’s avatar to party like its 1999 at 100th Winter Simulation Conference! #InsideTomSchriberHowHeWorks&WhyItMatters
1:19 AM · 4 Dec 2067

7590 RETWEETS  31k LIKES
“To a large degree, it [simulation] developed as a counter element in the emerging computer culture — a tool for support in an area that aimed toward the practical rather than the theoretical, and toward the complex rather than the simple.”


Computing, data & acceptance of scientific management work in our favor.

- Provide robust, credible solutions that comprehensively hedge risks.
- Be the foundation for modeling the most difficult societal problems.
- Produce data analytics before there are data.
- Deliver theory in support of practice, and practice in support of theory.

- We do all that and our chances are pretty good.
THANK YOU
SLAVE LABOR
Yujing Lin • Linda Pei • Mark Semelhago • Eunhye Song

SAGE WISDOM
Paul Fishwick (modeling guru) • Robert Sargent (history spoilers)
Dave Goldsman (joke tester) • Russell Barton (ego deflator)

TEST MARKETS
2017 INFORMS Simulation Society Workshop:
“Toward an Ecosystem of Models and Data”

2017 BIRS Workshop: “Future Research Directions in
Digital Simulation Methodology for the Next 10 Years”

Supply Chain Optimization Tech Talk @ Amazon

INSPIRATION
Peter Glynn • Peter Haas • Jeff Hong • Leon McGinnis
Shane Henderson • Henry Lam • Susan Sanchez
And the many who have given their time and talents to WSC