From Measurement to Scientific Data Analysis
IMPRS Lecture Series “Space Instrumentation”

Alex Feller

2010-10-25
Outline

1. Basic aspects of measurement and error analysis
   - What is a measurement?
   - Accuracy, precision
   - Errors, estimates, uncertainties
   - Summary

2. Data reduction
   - Example: Sunrise
   - Basic reduction steps
   - Image restoration

3. Where to go from here?
What is a measurement?

"Classical" definition
The process of estimating or determining the magnitude of a physical quantity relative to a unit of measurement.

Representational theory
The correlation of numbers with entities that are not numbers.

Information theory
A set of observations that reduce uncertainty where the result is expressed as a quantity.
Accuracy vs. precision

Accuracy
- Systematic errors
- Can be improved by calibration

Precision
- Random errors (noise)
- Can be improved by repeated measurements
Accuracy vs. precision

**Accuracy**
- Systematic errors
- Can be improved by calibration

**Precision**
- Random errors (noise)
- Can be improved by repeated measurements
Gaussian and Poisson error distributions

- Photon noise!
- From “dice game” (binomial dist.) to radiative transition (Poisson dist.)
- \( P_p(x; \mu) \rightarrow P_G(x; \mu, \sigma = \sqrt{\mu}) \) for \( x \) large
Estimates and uncertainties

Mean value as maximum likelyhood estimate (MLE)

- Model for measurements $y_i, i = 1 \ldots N$:

\[ y_i = \mu + n_i \]

with i.i.d. Gaussian noise $n_i$ and free parameter $\mu$.

- Estimate:

\[ \hat{\mu} = \bar{y} = (1/N) \sum_{i=1}^{N} y_i \]
Estimates and uncertainties

Mean value as maximum likelyhood estimate (MLE)

- Probability for realizing the observed set of measurements $y_1, \ldots, y_N$:
  \[
P(y_1, \ldots, y_N) = \prod_i P_G(y_i - \mu; 0, \sigma) \propto \exp \left\{ -\frac{1}{2\sigma^2} \sum_i (y_i - \mu)^2 \right\}
  \]

- Maximum probability i.e. maximum likelihood for model to yield the measurements $y_i$:
  \[
  \chi^2 = \sum_i (y_i - \hat{\mu})^2 \text{min.} \Rightarrow \hat{\mu} = \bar{y}
  \]

“Least squares”
Estimates and uncertainties

General maximum likelihood estimates

- Model for measurements $y_i, i = 1 \ldots N$:
  
  $$y_i = y(x_i; \mu_j) + n_i(\sigma_i)$$

- with Gaussian noise $n_i(\sigma_i)$ of standard deviations $\sigma_i$
- and with free parameters $\mu_j$. 
Estimates and uncertainties

- Probability for realizing the observed set of measurements $y_1, \ldots, y_N$:

$$P(y_1, \ldots, y_N) = \prod_i P_G(y_i - y(x_i); 0, \sigma_i)$$

$$\propto \exp \left\{ - \frac{1}{2} \sum_i \left( \frac{y_i - y(x_i)}{\sigma_i} \right)^2 \right\}$$

- Maximum likelihood estimates for parameters $\mu_j$:

$$\chi^2 = \sum_i \left( \frac{y_i - y(x_i; \mu_j)}{\sigma_i} \right)^2 \min.$$
Estimates and uncertainties

Uncertainty of the mean value?

- Estimate:
  \[ \hat{\mu} = \bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i \]
- Error propagation: \( \sigma_{\mu} = \frac{\sigma}{\sqrt{N}} \)
- 68.3% confidence interval: \( \bar{y} \pm \sigma_{\mu} \)

Uncertainty of a general model parameter?

- \( \chi^2 = \sum_i \left( \frac{y_i - y(x_i; \mu_j)}{\sigma_i} \right)^2 \) min. \( \rightarrow \hat{\mu}_j \)
- Evaluate variations of \( \chi^2 \) as a function of the parameter
- Monte Carlo simulation
- …
Calibration

Goal: reduce systematic errors!

But calibrations are measurements of their own which come with their proper measurements errors → error propagation!

2 types of calibration measurements

1. Determine the correlation between the actual measurand and the representative measurand.
   Examples:
   - wavelength calibration of spectra: wavelength ↔ pixels
   - polarimetric calibration: polarization ↔ intensity

2. Determine the influence quantities which
   - alter the physical quantity directly (e.g. movement of the spacecraft → Doppler shifts)
   - or which bias the output signal of the instrument (e.g. dark current, flatfield).
Summary - Questions to be asked about a measurement

- Exact measurement task and quantity to be measured?
- Measurement principle?
- Absolute or relative measurement?
- Direct or indirect measurement?
- Required accuracy and precision? Sensitivity?
- What is the measuring environment? What are the influencing quantities? Which calibrations are needed?
Data reduction example: Sunrise

Key Parameters

- Telescope diam.: 1m
- Spatial res.: 0.1" (coin in 45 km!)
Data reduction example: Sunrise

Key Parameters

- Telescope diam.: 1m
- Spatial res.: 0.1"
  (coin in 45 km!)
- 6 days flight at 37 km altitude
Data reduction example: Sunrise

Key Parameters

- Telescope diam.: 1m
- Spatial res.: 0.1" (coin in 45 km!)
- 6 days flight at 37 km altitude
Data reduction example: Sunrise

Key Parameters

- Telescope diam.: 1m
- Spatial res.: 0.1" (coin in 45 km!)
- 6 days flight at 37 km altitude
- Sunrise Filter Imager (SuFI):
  - 214 nm, 300 nm, 312 nm (OH), 388 nm (CN), 397 nm (Ca II)
  - FOV: 14" x 40"

Metrology
- What is a measurement?
- Accuracy, precision
- Errors, estimates, uncertainties
- Summary

Data reduction
- Example: Sunrise
- Basic reduction steps
- Image restoration

Where to go from here?
Data reduction example: Sunrise

Key Parameters

- Telescope diam.: 1m
- Spatial res.: 0.1" (coin in 45 km!)
- 6 days flight at 37 km altitude
- Sunrise Filter Imager (SuFI):
  - 214 nm, 300 nm, 312 nm (OH), 388 nm (CN), 397 nm (Ca II)
  - FOV: 14" x 40"
- IMaging Magnetograph Experiment (IMaX):
  - Full Stokes spectropolarimetry in Fe I 525 nm
  - FOV: 50" x 50"
<table>
<thead>
<tr>
<th>Before</th>
<th>After</th>
</tr>
</thead>
</table>

### Sunrise/SuFl data - from raw to final

#### Metrology
- What is a measurement?
- Accuracy, precision
- Errors, estimates, uncertainties
- Summary

#### Data reduction
- Example: Sunrise
  - Basic reduction steps
  - Image restoration

#### Where to go from here?
Sunrise/SuFI data - from raw to final

<table>
<thead>
<tr>
<th>Before</th>
<th>After</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Before Image" /></td>
<td><img src="image2.png" alt="After Image" /></td>
</tr>
</tbody>
</table>

**Metrology**
- What is a measurement?
- Accuracy, precision
- Errors, estimates, uncertainties
- Summary

**Data reduction**
- Example: Sunrise
  - Basic reduction steps
  - Image restoration

**Where to go from here?**
Sunrise/SuFl data - from raw to final

**Before**

**After**

Sunrise/SuFl data - from raw to final

Before

After

Sunrise/SuFl data - from raw to final

Before

After

Sunrise/SuFl data - from raw to final

Before

After

Sunrise/SuFl data - from raw to final

Before

After

Sunrise/SuFl data - from raw to final

Before

After
Sunrise/SuFl data - from raw to final

Cont. 300 nm

Level 0

Raw data!
Sunrise/SuFI data - from raw to final

Cont. 300 nm

Level 1

- Dark image correction
- Flatfield correction
- Correction for residual defects
  - Median filtering (cosmic rays rem, . . .)
  - Low-pass filtering (scratches, . . .)
  - Fourier filtering (fringes, electronic interferences, . . .)
Sunrise/SuFI data - from raw to final

Cont. 300 nm

Level 2,3

Phase Diversity restoration
Sunrise/SuFI data - from raw to final

<table>
<thead>
<tr>
<th>Cont. 300 nm</th>
<th>Time series</th>
</tr>
</thead>
</table>
|              | • Frame selection  
              | • Cross-correlation |

**Measurement to Scientific Analysis**

A. Feller

**Metrology**
- What is a measurement?
- Accuracy, precision
- Errors, estimates, uncertainties
- Summary

**Data reduction**
- Example: Sunrise
  - Basic reduction steps
  - Image restoration

Where to go from here?
Dark and flatfield correction

Level-0 image \( (I_0) \)

Dark and flatfield corr. image \( (I) \)

\[
I_0(x, y) = F(x, y) \cdot I(x, y) + D(x, y; T, \Delta t)
\]

- \( I_0 \): obs. level-0 image
- \( I \): input image
- \( F \): flatfield
- \( D \): dark image
- \( \Delta t \): integration time
- \( T \): detector temperature
**Dark and flatfield correction**

### Dark image \((I = 0)\)

\[
I_0(x, y) = F(x, y) \cdot I(x, y) + D(x, y; T, \Delta t)
\]

- \(I_0\) \(\text{obs. level-0 image}\)
- \(I\) \(\text{input image}\)
- \(F\) \(\text{flatfield}\)
- \(D\) \(\text{dark image}\)
- \(\Delta t\) \(\text{integration time}\)
- \(T\) \(\text{detector temperature}\)
Dark and flatfield correction

Dark image ($I = 0$), lower-right quadrant

$$I_0(x, y) = F(x, y) \cdot I(x, y) + D(x, y; T, \Delta t)$$

$I_0$ \hspace{1cm} obs. level-0 image
$I$ \hspace{1cm} input image
$F$ \hspace{1cm} flatfield
$D$ \hspace{1cm} dark image
$\Delta t$ \hspace{1cm} integration time
$T$ \hspace{1cm} detector temperature
Dark and flatfield correction

Flatfield ($/\text{const}$)

$$l_0(x, y) = F(x, y) \cdot l(x, y) + D(x, y; T, \Delta t)$$

- $l_0$ obs. level-0 image
- $l$ input image
- $F$ flatfield
- $D$ dark image
- $\Delta t$ integration time
- $T$ detector temperature
Dark and flatfield correction

Level-0 image ($I_0$)

Dark and flatfield corr. image ($I$)

\[ I \propto \frac{I_0 - D}{F - D} \]

- $I_0$: obs. level-0 image
- $I$: input image
- $F$: flatfield
- $D$: dark image
- $\Delta t$: integration time
- $T$: detector temperature
Fringe filtering

IMaX Image

Power spectrum

Where to go from here?
Fringe filtering

Filtered IMaX Image

Filtered power spectrum

Metrology
- What is a measurement?
- Accuracy, precision
- Errors, estimates, uncertainties
- Summary

Data reduction
- Example: Sunrise
- Basic reduction steps
- Image restoration

Where to go from here?
Image restoration

Heritage from ground-based solar observations

Ca II H line center, recorded at Swedish Solar Telescope; time span 10 min., cadence 4s (J. Hirzberger)
Image restoration

Heritage from ground-based solar observations

Ca II H line center, recorded at Swedish Solar Telescope; time span 10 min., cadence 4s (J. Hirzberger)
Phase Diversity - Principle

Paxman et al. 1992
Phase Diversity - Principle

What we have

- of the unknown true Sun
- with unknown aberrations

- of the same unknown true Sun
- with the same unknown aberrations
- plus a known defocus

**a conventional image**

**a diversity image**
Phase Diversity - Principle

This allows us to make maximum-likelihood estimates of . . .

the aberrations

in terms of

- pupil function
- or Zernike coefficients

the true solar image
Phase Diversity - a glimpse of the beauty of the algorithm

Remember the earlier discussion of maximum-likelihood estimates?

- Model: $d_k = f \ast t_k + n_k, \quad k = 1, 2$
- Maximum-likelihood error metric (objective function):

$$L = \sum_v |D_1 - FT_1|^2 + |D_2 - FT_2|^2$$

$d_{1,2}$ conventional and diversity image

$f$ true solar image

$t_{1,2}$ point spread functions of conv. and div. image

$n_k$ Gaussian noise

$D_k, F, T_k$ FFTs of $d_k$, $f$ and $t_k$
Phase Diversity - a glimpse of the beauty of the algorithm
But wait a minute . . . how many free parameters do we have? $10^5$?
Phase Diversity - a glimpse of the beauty of the algorithm
But wait a minute . . . how many free parameters do we have? $10^5$?

“Trick I”: Solve explicitly for the aberrations only!

\[ \phi = \sum_i c_i \phi_i \quad (1) \]

\[ T_k = T_k(\phi) = T_k(c_i) \quad (2) \]

This leaves us with typically 21 to 45 free parameters instead of $10^5$!
Phase Diversity - a glimpse of the beauty of the algorithm
But wait a minute . . . how many free parameters do we have? $10^5$?

“Trick I”: Solve explicitly for the aberrations only!

Paxman et al. 1992

```
\[ \phi = \sum_i c_i \phi_i \]  
\[ \rightarrow T_k = T_k(\phi) = T_k(c_i) \]
```

“Trick II”: Use a Zernike expansion for the aberrations!

Expand the pupil phase $\phi$ into a series of Zernike functions $\phi_i$:

Paxman et al. 1992
Phase Diversity - a glimpse of the beauty of the algorithm
But wait a minute . . . how many free parameters do we have? $10^5$?

“Trick I”: Solve explicitly for the aberrations only!

“Trick II”: Use a Zernike expansion for the aberrations!

Expand the pupil phase $\phi$ into a series of Zernike functions $\phi_i$:

$$\phi = \sum_i c_i \phi_i$$

$$\rightarrow T_k = T_k(\phi) = T_k(c_i)$$

This leaves us with typically 21 to 45 free parameters instead of $10^5$!
Phase Diversity - What is it really good for?

- Structures that are not already “visible” in the raw images cannot be recovered!
- Aberrations re-distribute the intensity in the image
- By doing image restoration we want to recover the true intensity distribution (contrast)

Phase Diversity - What is it really good for?

• Structures that are not already “visible” in the raw images cannot be recovered!
• Aberrations re-distribute the intensity in the image
• By doing image restoration we want to recover the true intensity distribution (contrast)
Phase Diversity - What is it really good for?

- Structures that are not already “visible” in the raw images cannot be recovered!
- Aberrations re-distribute the intensity in the image
- By doing image restoration we want to recover the true intensity distribution (contrast)

How reliable is the restoration?
Phase Diversity - What is it really good for?

- Structures that are not already “visible” in the raw images cannot be recovered!
- Aberrations re-distribute the intensity in the image
- By doing image restoration we want to recover the true intensity distribution (contrast)

How reliable is the restoration?

**Influence of noise**

![Graph showing the influence of noise on the normalized RMS contrast with S/N = 1000 and S/N = 100.](image)

- Hirzberger et al. 2010

Metrology
- What is a measurement?
- Accuracy, precision
- Errors, estimates, uncertainties

Summary

Data reduction
- Example: Sunrise
- Basic reduction steps

Image restoration

Where to go from here?
Phase Diversity - What is it really good for?

- Structures that are not already “visible” in the raw images cannot be recovered!
- Aberrations re-distribute the intensity in the image
- By doing image restoration we want to recover the true intensity distribution (contrast)

How reliable is the restoration?

Hirzberger et al. 2010

Restoration artifacts

- Structures that are not already “visible” in the raw images cannot be recovered!
- Aberrations re-distribute the intensity in the image
- By doing image restoration we want to recover the true intensity distribution (contrast)

How reliable is the restoration?

Hirzberger et al. 2010

Image restoration

Where to go from here?

Metrology

- What is a measurement?
- Accuracy, precision
- Errors, estimates, uncertainties
- Summary

Data reduction

Example: Sunrise
- Basic reduction steps

Where to go from here?
Where to go from here?

Recommended reading

- “Numerical Recipes”, W.H. Press et al., Cambridge University Press

Looking for code?

- SolarSoft Library (http://www.lmsal.com/solarsoft/)
- Community has developed many IDL code snippets that can be easily re-used for different purposes - ask around at MPS!