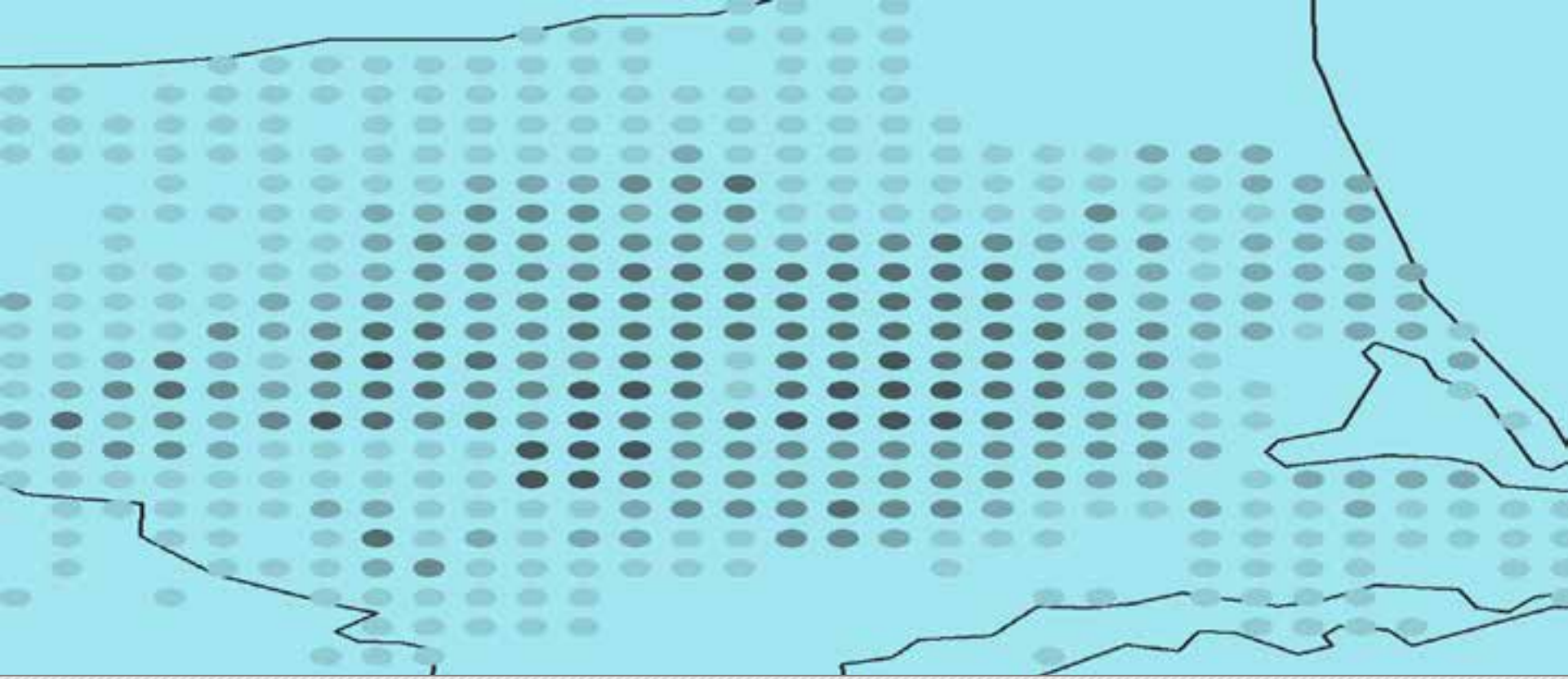


Disaggregation in Risk Models

An approach to improve risk assessment

Jutta Schmieder
Geospatial Solutions, Munich Re

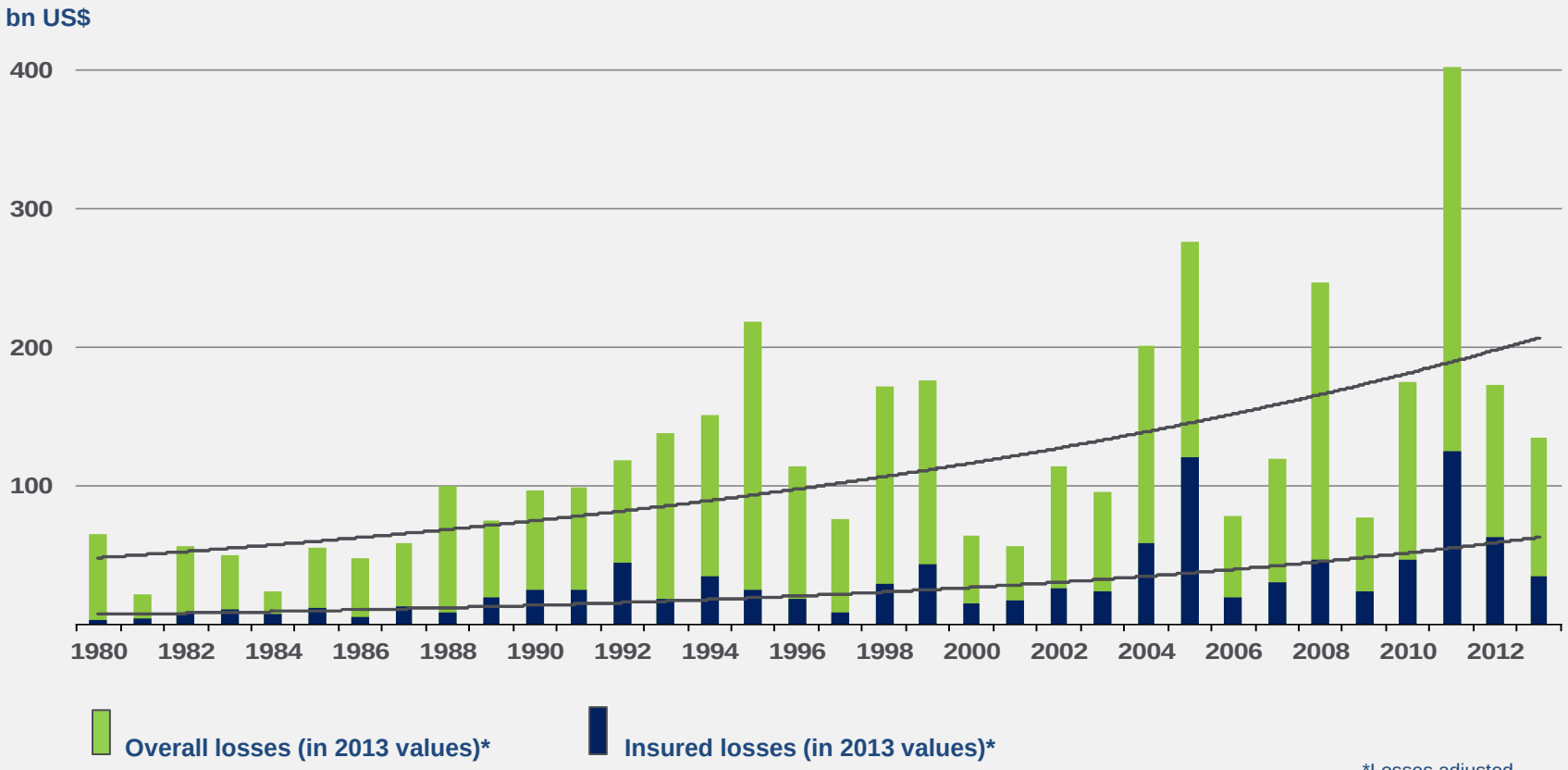
ESRI User Conference - July 15, 2014

Munich RE

Introduction

Munich RE

Natural Disasters: Loss events worldwide 1980 – 2013
Overall and insured losses



*Losses adjusted to inflation based on country CPI.

Introduction

Munich RE

Natural Disasters:

Where do different types of events happen?

How likely is their occurrence?

What damage and impact per event can be expected?

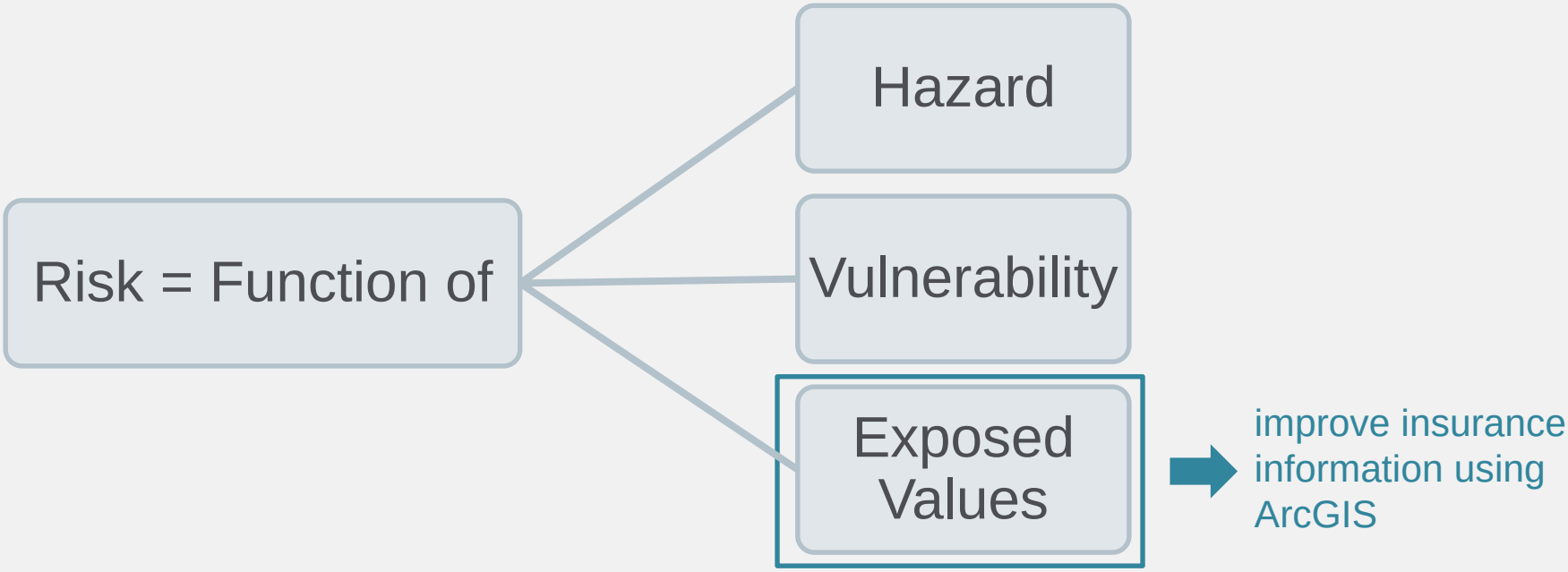
What is the expected insured loss?

-> Munich Re develops its own probabilistic NatCat models

-> Model software: MATLAB

How to assess the risk?

Risk equation:

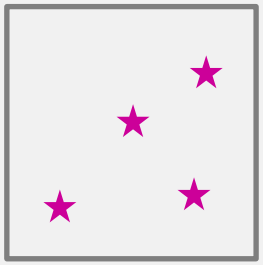


The model output is oly as good as the quality of the data input.

What is the challenge?

Insurer gets clients exposure data as

Hazard and vulnerability data is available at a detailed resolution



detailed data,
e.g. single locations with address information

aggregated data,
e.g. several risks summarized per postal or administrative zone

Challenge!

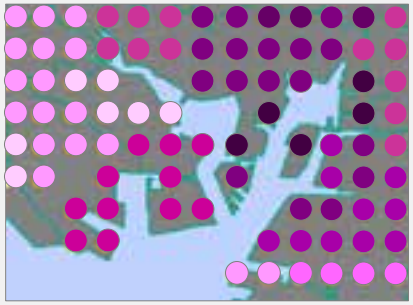
Total Sum Insured per postal zone

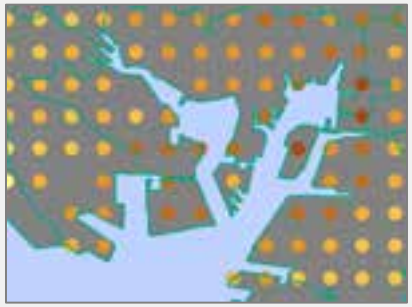
Approach

Total Sum Insured
per postal code

Intelligent Disaggregation

Total Sum Insured
per modelling point





Weighted
modelling
points

Modelling Points

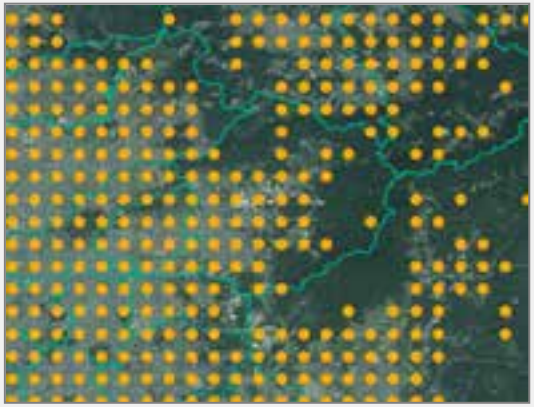
Underlying value distributions

How to get the data together?

Use modelling grid points, which
- represent all relevant aspects of hazard, vulnerability and exposure
- ensure good performance on the part of the model.

Modelling points are designed in a „semi-regular raster“.
Spacing depends mainly on
- peril in general (e.g. flooding needs higher resolution than winter storm)
- local variations of the hazard (e.g. flooding areas ↔ secure areas)
- variation of vulnerability (e.g. high-rise buildings vs. low-rises)
- spatial distribution of the exposure

Areas with high data impact / data variability:
spacing between points is quite small
Homogeneous or non-hazardous areas:
spacing can be quite wide



Side note to modelling points

Munich RE

Generation of modelling points
- design in ArcGIS
- based on several rasters with different spatial resolutions

Steps:
- Generation of several rasters
  (1) a high resolution raster (0.0005 degrees), based on specific exposure and hazard information
  (2) aggregations of raster(1) to coarser resolutions
- Set and apply conditions for cell selection. Conditions are different for each raster resolution (e.g. for resolution “0.0045 degrees”: hazard >5 and industrial buildings >=1 and...).
- Convert selected cells to points and combine to a common point file.

Result: Point file, which covers the modelling area completely,
but in different density.

POINT_X	POINT_Y	XY_ID	resolution
152,06925	-27,00575	152069250_-27005750	0,001
152,04625	-27,00875	152046250_-27008750	0,001
151,99325	-27,00675	151993250_-27006751	0,005
152,04325	-27,22175	152043250_-27221750	0,005
152,18325	-27,02175	152183250_-27021750	0,025
152,03325	-27,04675	152033249_-27046750	0,025

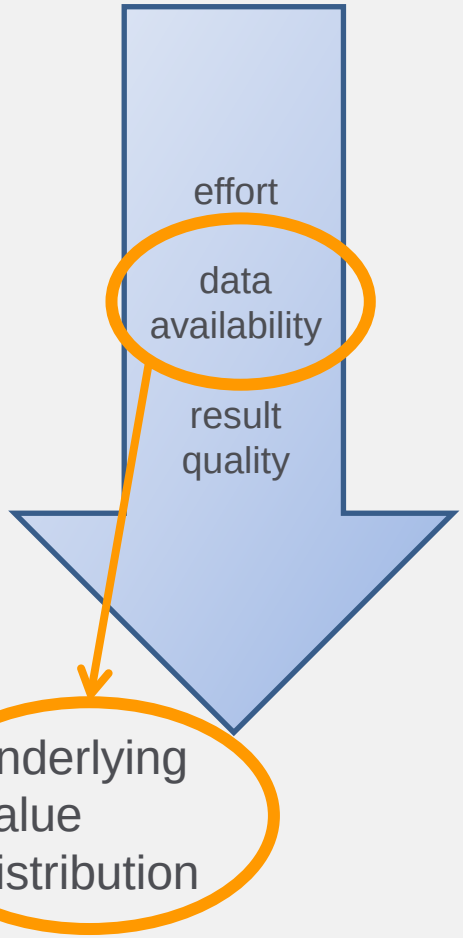
Possibilities of Exposure Disaggregation

Distribute the value per zone (i.e. client's liability data) to the corresponding modelling points

homogeneously

according to population (frequently used)

according to each line of business
(e.g. industrial risks to industrial sites,
commercial risks to commercial areas etc.)



Underlying value distribution

Munich RE

Requirement for intelligent disaggregation: high level background data as basis
→ BUT: No optimum global database easily available, esp. not for industrial and commercial exposure...
→ Underlying value distributions for each line of business (residential, commercial, industrial etc.) have to be generated!

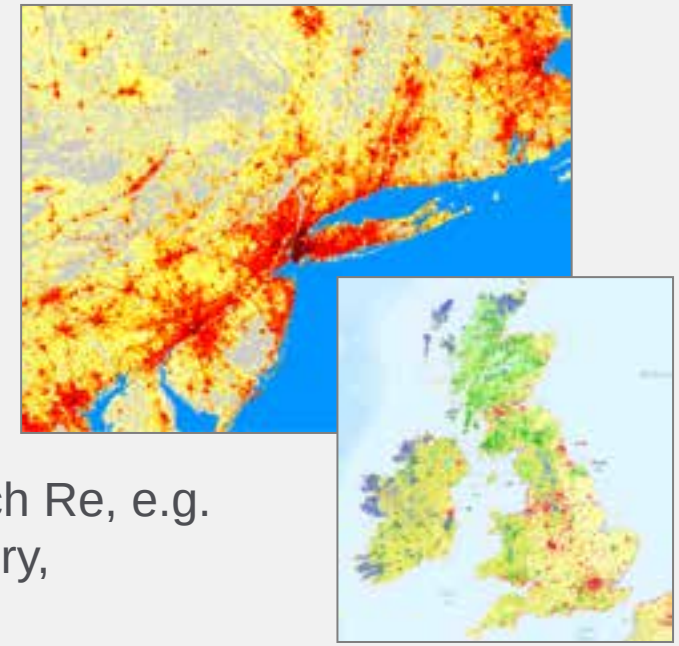
→ research work: each country has to be analyzed for appropriate data sources and suitable kind of data
→ different raw data for each country

→ Development of a common approach to generate underlying exposure distributions (ArcGIS model)
→ however, evaluation methodology has to be individually adapted according to data availability...

Generation of value distributions
Data collection

- Publicly accessible sources, e.g.
  - land-use information, e.g. CORINE data for Europe
  - address attributes (e.g. TOMTOM)
  - databases on ports, airports, power plants etc.
  - Population databases (e.g. 24-hour-population)
  - business databases (incl. addresses and occupancy codes)
  - Statistical data as GDP per region, employees per region etc.

→ frequently not sufficient



- Additional detailed information collected by Munich Re, e.g.
  - industry park polygons based on satellite imagery, the homepages of industrial parks etc.
  - global location database for a variety of industries such as automotive, chemical, electronics, etc.
  - Global Database on critical infrastructure

Generation of value distributions
Data preparation

Munich RE

Challenges:

- Some data include indicator values (e. g. numbers of employed in business databases, numbers of addresses at one point), some data indicate only “yes/no” (e.g. airports)
- Some databases only include points, but are in reality areas (e.g. ports) → buffer points to obtain areas
- Different data resolutions (point data, polygons, various raster)

→ harmonize all data (incl. values as e.g. population) to high resolution raster with 0,0005 deg resolution (= ca. 50m)
→ classify all raster to the same range (e.g. 0-1000) to make raster comparable
→ result: diverse raster with different values, but the same resolution, range and extend
→ raster can be weighted and combined

Goal: High resolution value distribution **per line of business** (RES, COM, IND..)

Generation of value distributions
Data combination (1)

Munich RE

1. Define weights and combinations per line of business
   -> has always to be done individually; depends not only on content but also on data quality and reliability

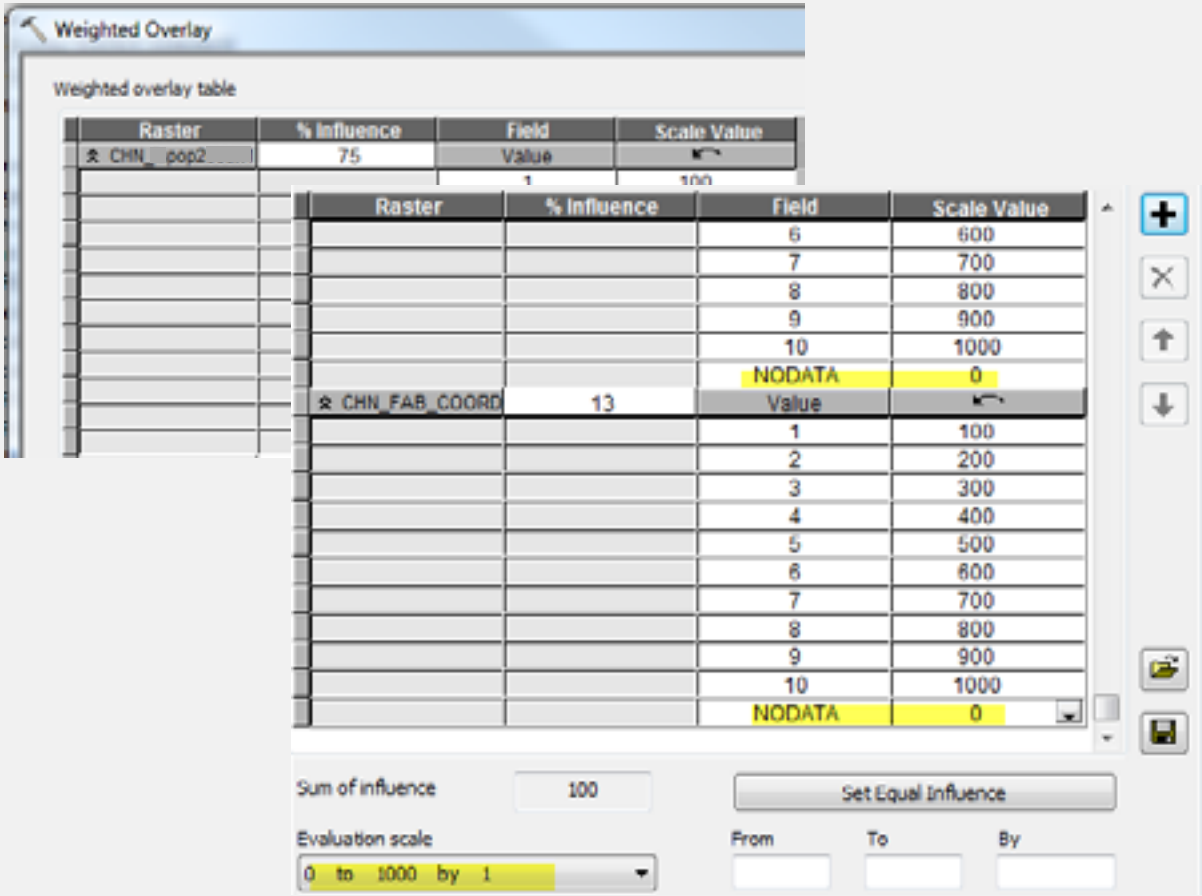
Example (Dummy Weighting):

			Weights		
Input layer (50m resolution)	**Value** per raster point	**Classes**	**RES%**	**COM%**	**IND%**
Population (24-hour pop)	population per raster cell	1-1000	100%	25%	10%
Industrial parks	no exposure / exposure exists	nodata/1000		10%	35%
Landuse "factories"	no exposure / exposure exists	nodata/1000		15%	25%
Landuse "highrises"	no exposure / exposure exists	nodata/1000		30%	
Ports	no exposure / exposure exists	nodata/1000			10%
commercial employees	number of employees per raster cell	1-1000		20%	
industrial employees	number of employees per raster cell	1-1000			20%

Generation of value distributions
Data combination (2)

Munich RE

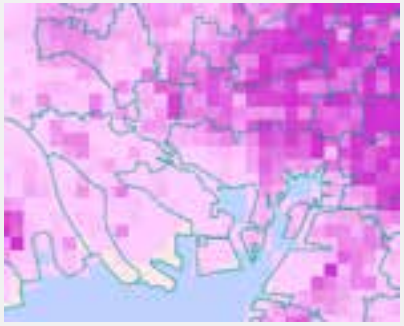
2. Combine raster data by “Weighted Overlay”



Value assignment to modelling points (1)

Result of data combination: Value distributions per Line of Business

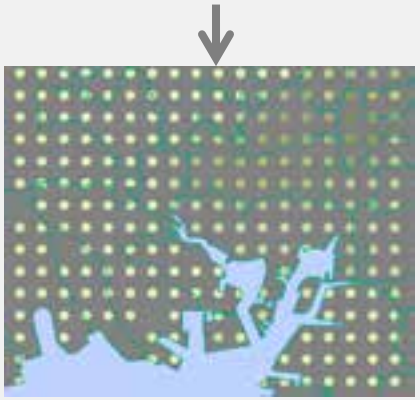
RES

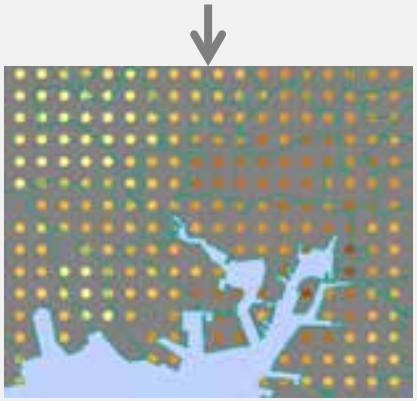


COM

IND

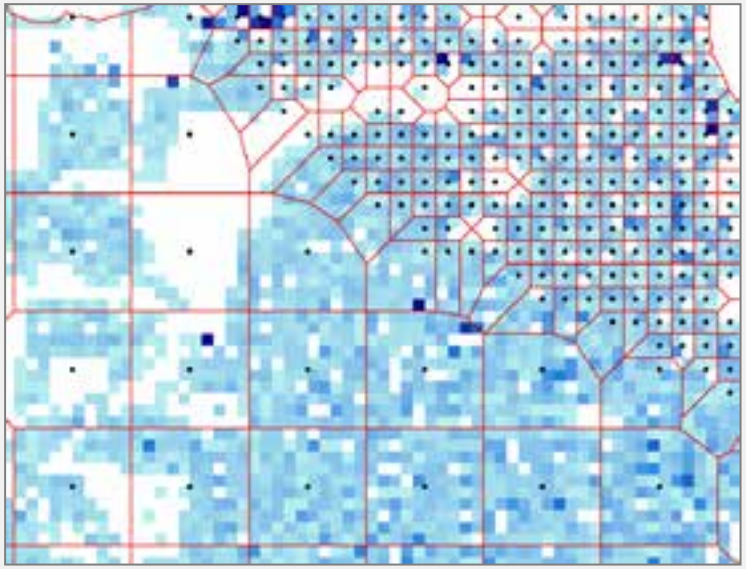
Next Step: assign values to the corresponding modelling points





Value assignment to modelling points (2)

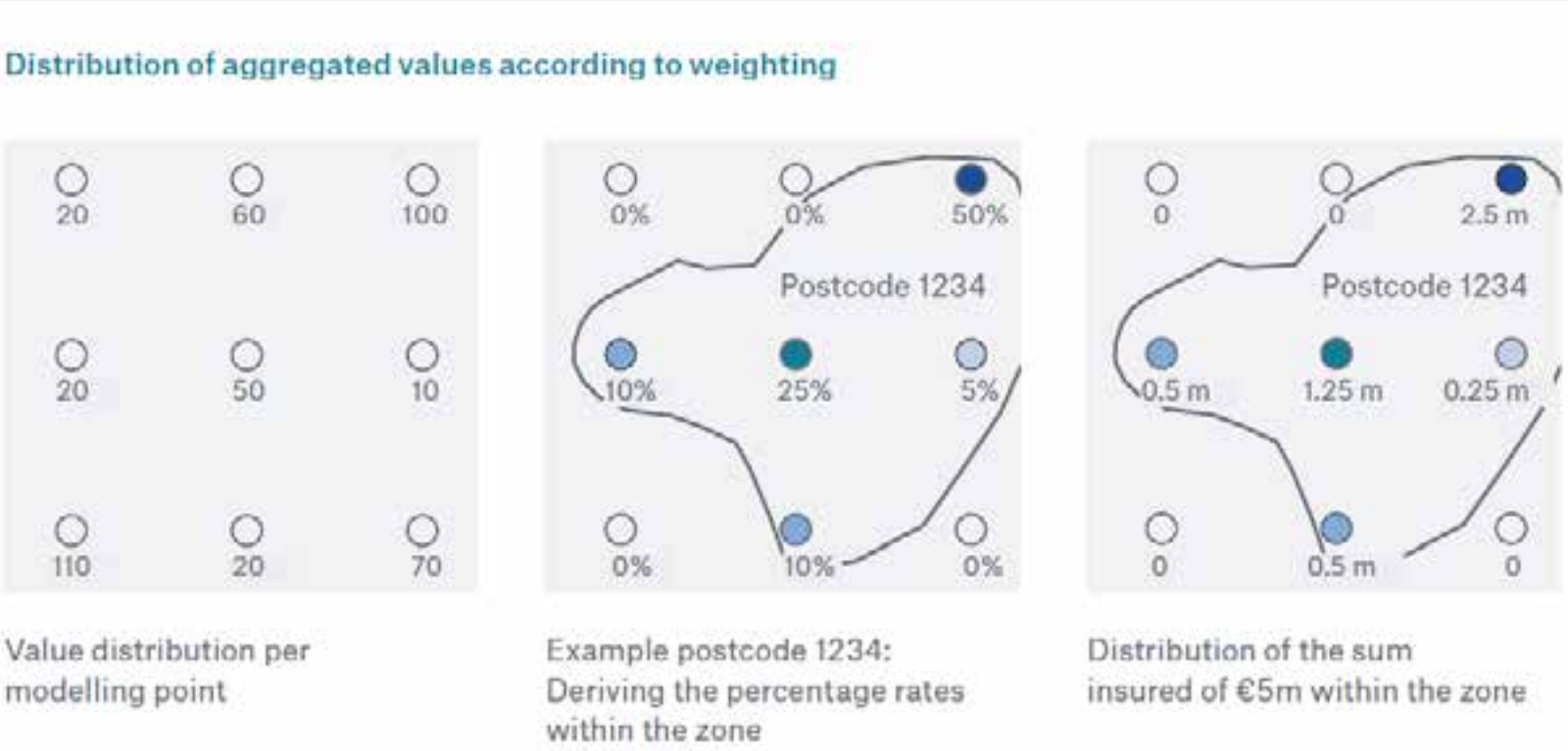
How to attach values to the modelling points?
→ define „catchment areas“ per modelling point
→ use „Calculate Thiessen-Polygons“
for modelling points
→ clip with boarders of regarded area (e.g. China)



→ Calculate “Sum Exposed per modelling point” by „Zonal Statistics as Table“

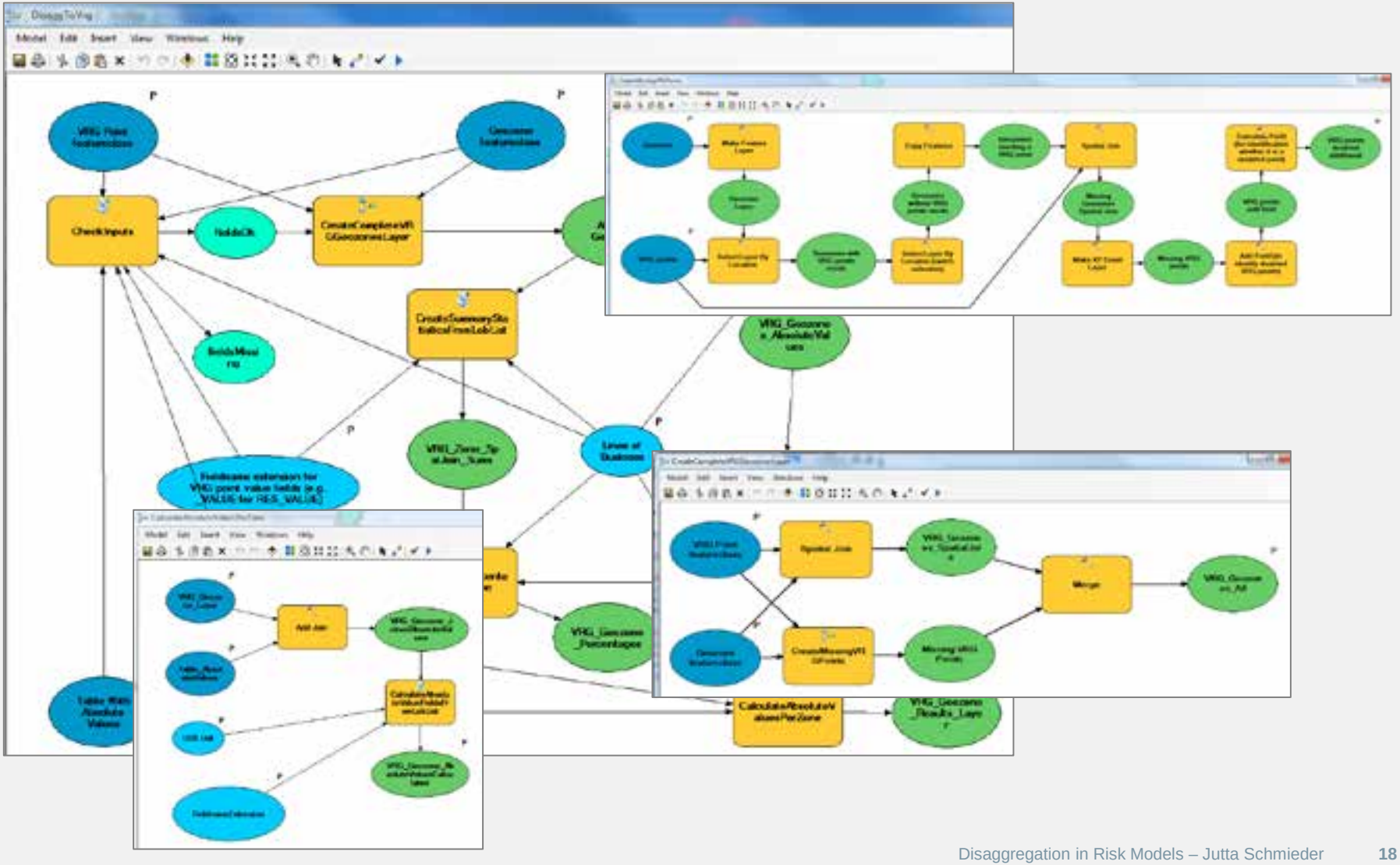
Distribution of liability data

Munich RE



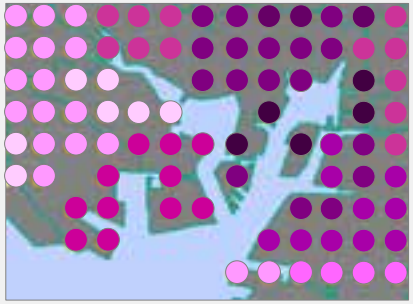
All steps are automated in ArcGIS models

ArcGIS Model including various Submodels



Results

Intelligent Disaggregation



file for implementation in the NatCat models (running in MATLAB):

X	Y	ID_XY	ADM2_ID	RES_ADM2_PC	COM_ADM2_PC	IND_ADM2_PC
141,517583	43,215917	141517583_43215916	01303	0,17	0,033	0,039
141,504083	43,215917	141504083_43215916	01303	0,03	0,045	0,056
141,544583	43,215917	141544583_43215916	01303	0,024	0,036	0,045
141,531083	43,242917	141531083_43242916	01303	0,071	0,04	0,045
141,450083	43,202417	141450083_43202416	01303	0,052	0,094	0,116
141,517583	43,202417	141517583_43202416	01303	0,028	0,06	0,074
141,558083	43,215917	141558083_43215916	01303	0,113	0,068	0,08
141,517583	43,242917	141517583_43242916	01303	0,133	0,058	0,051
141,517583	43,229417	141517583_43229416	01303	0,097	0,027	0,031
141,504083	43,229417	141504083_43229416	01303	0,066	0,044	0,051
141,652583	43,229417	141652583_43229416	01304	0,338	0,376	0,375
141,612083	43,175417	141612083_43175416	01304	0,297	0,421	0,4
141,625583	43,202417	141625583_43202416	01304	0,366	0,203	0,224
140,086583	41,541917	140086583_41541916	01331	0,002	0,029	0
140,113583	41,433917	140113583_41433916	01331	0,16	0,089	0,092
140,100083	41,433917	140100083_41433916	01331	0,227	0,09	0,097
140,127083	41,433917	140127083_41433916	01331	0,05	0,148	0,138

→ Liability values of one zone (“01304”) are distributed to 3 modelling points (sum = 100%)

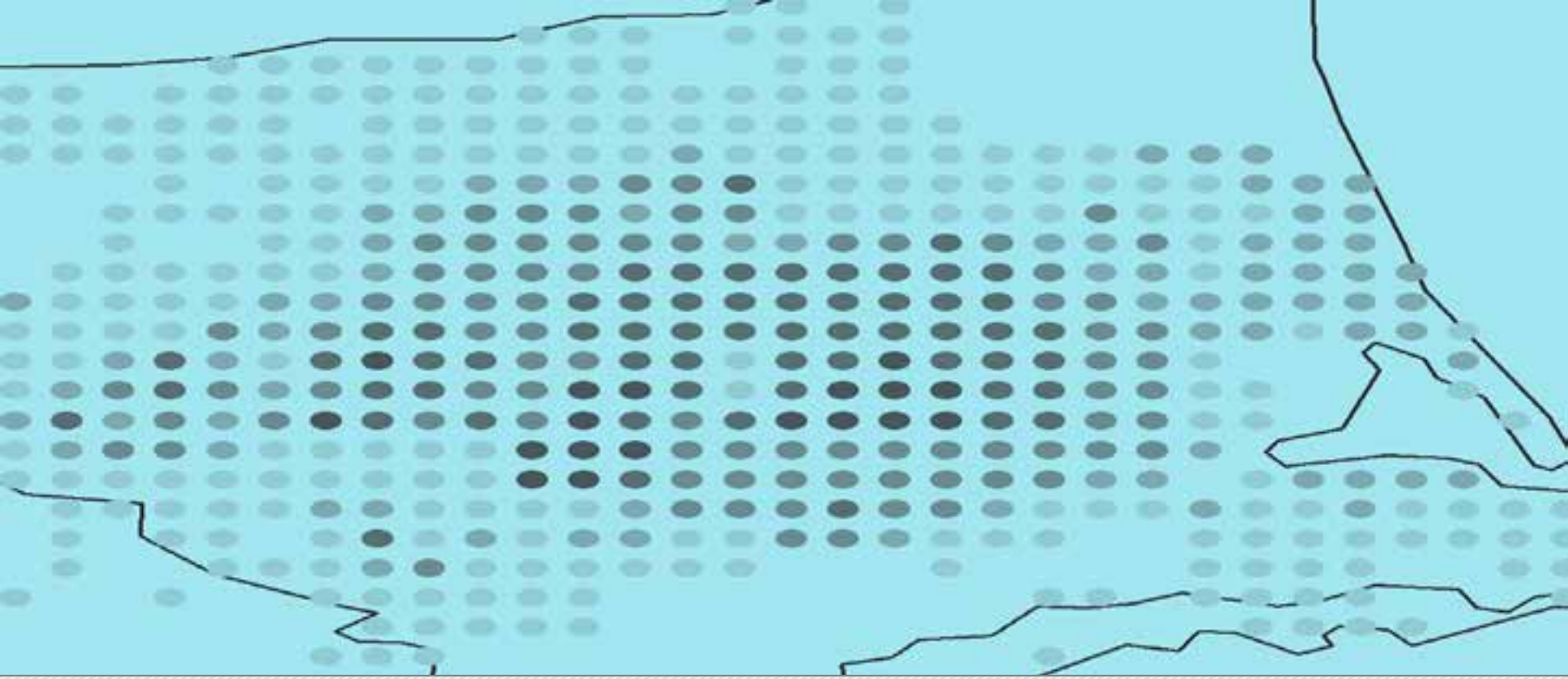
Final Remarks

Disaggregation can only approximate the actual spatial location of the insured objects
-> a real distribution is not reflected necessarily
-> the distribution will always correspond to the market average

-> impacts in particular large zones: individual liability accumulations, e.g. of regional insurers writing business on a limited basis, cannot be identified.

This problem can only be circumvented by collecting liability data for each location in a detailed way from the outset.

Customer-specific liability distributions and individual modelling results that best reflect reality can then be obtained.



Thank you very much for your attention!

Jutta Schmieder
Geospatial Solutions, Munich Re

Munich RE