WORKING WITH REAL DATA: GETTING ANALYTIC ELEMENT GROUNDWATER MODEL RESULTS TO FIT FIELD DATA

CONGDON, Roger D., USDA Forest Service, 333 Broadway Blvd SE, Albuquerque, NM 87102, rcongdon@fs.fed.us

CONGDON, Roger D., USDA Forest Service, 333 Broadway Blvd SE, Albuquerque, NM 87102, rcongdon@fs.fed.us

Models of groundwater flow often work best when very little field data exist. In such cases, some knowledge of the depth to the water table, annual precipitation totals, and basic geological makeup is sufficient to produce a reasonable-looking and potentially useful model. However, in this case where a good deal of information is available regarding depth to bottom of a dune field aquifer, attempting to incorporate the data set into the model has variously resulted in convergence, failure to achieve target water level criteria, or complete failure to converge. The first model did not take the data set into consideration, but used general information that the aquifer was thinner in the north and thicker in the south. This model would run and produce apparently useful results. The first attempt at satisfying the data set; in this case 51 wells showing the bottom elevation of a Pacific coast sand dune aquifer, was to use the isopach interpretation of Robinson (OFR 73-241). Using inhomogeneities (areas of equal characteristics) delineated by Robinson's isopach diagram did not enable an adequate fit to the water table lakes, and caused convergence problems when adding pumping wells. The second attempt was to use a Thiessen polygon approach, creating an aquifer thickness zone for each data point. The results for the non-pumping scenario were better, but run times were considerably greater. Also, there were frequent runs with non-convergence, especially when water supply wells were added. Non-convergence may be the result of the lake line-sinks crossing the polygon boundaries or proximity of pumping wells to inhomogeneity boundaries. The third approach was to merge adjacent polygons of similar depths; in this case within 5% of each other. The results and run times were better, but matching lake levels was not satisfactory. The fourth approach was to reduce the number of inhomogeneities to four, and to average the depth data over the inhomogeneity. The thicknesses were varied within 5% of the average until the lake levels were closely matched. This last methodology proved satisfactory and stable. The data were honored and the solver worked relatively quickly; thus preserving the simplicity and speed of the Analytic Element method; and various pumping scenarios were stable.

Introduction

A common problem with groundwater models is dealing with "real" data. Generally it is fairly sparse and interpolation or other interpretations can lead to inappropriate results, or interpretations that make no sense whatever. When using drill data, this can be particularly difficult since the data represent a very small percentage of the project area. Contouring the data set; in this case the aquifer bottom elevation data, is often effective. However, the interpretation inherent in the contouring is subject to errors related to the completeness of the data set, variability of the data set on a local scale. The latter could be the result of unknown variation in the geology of the substrate, such as fracturing, faulting, previous differential weathering, or other factors undetectable by hole depth to substrate information. I investigated three methods of dealing with the aquifer depth information; one using contour interpretations by Robison (1973)*.

Another data interpretation issue was created by detailed bathymetry supplied by Rich Miller, a research assistant at Portland State University in Portland, Oregon. A decision had to be made whether to incorporate a lot of detail from the data into the lake stage files, or to use a simplified version.

* Robison, J. H., 1973, Hydrology of the dunes area north of Coos Bay, Oregon: U.S. Geological Survey Open-File Report, 62 p.

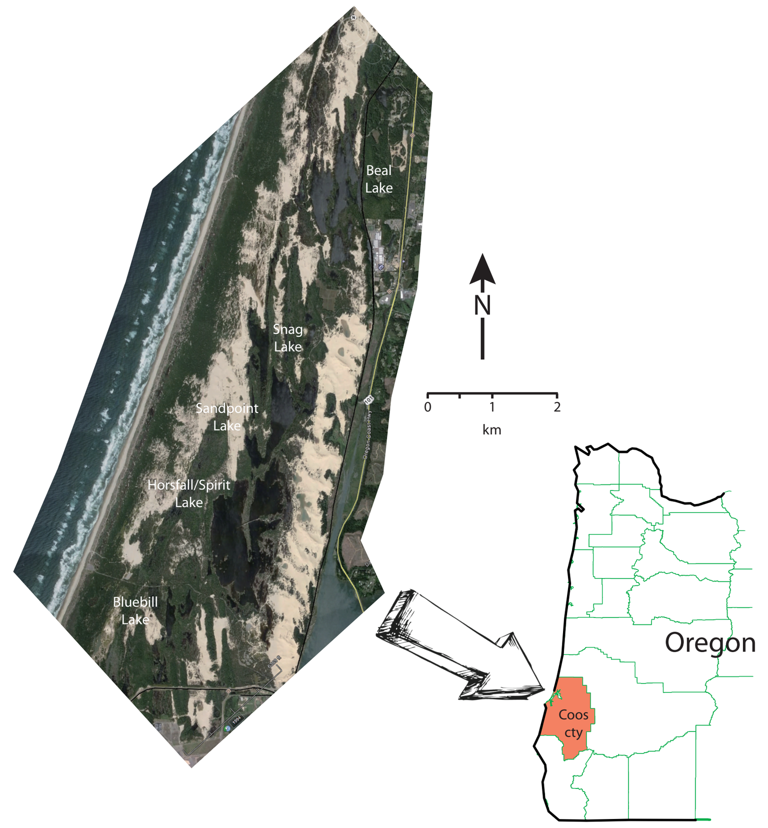


Figure 1. The central part of the coastal dune aquifer. The project is located just to the north of Coos Bay, Oregon, in Coos County. The aquifer is modeled as a uniform, unconfined unit of dune sand. Six water table lakes are also simulated. Their surface level is a reflection of the local position of the water table. Most of the project area is located within the Siuslaw National Forest.

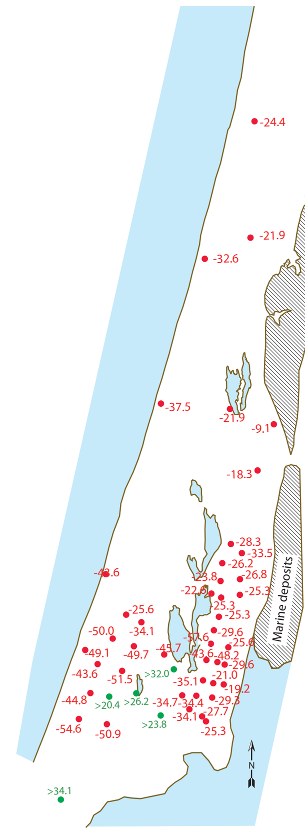


Figure 2. These are the depth to base of the dunal aquifer figures, in meters below mean sea level. The numbers in green, having the "greater than" (>) symbol in front of them, did not reach the aquifer base. Note that most data points are clustered about the central part of the study area.

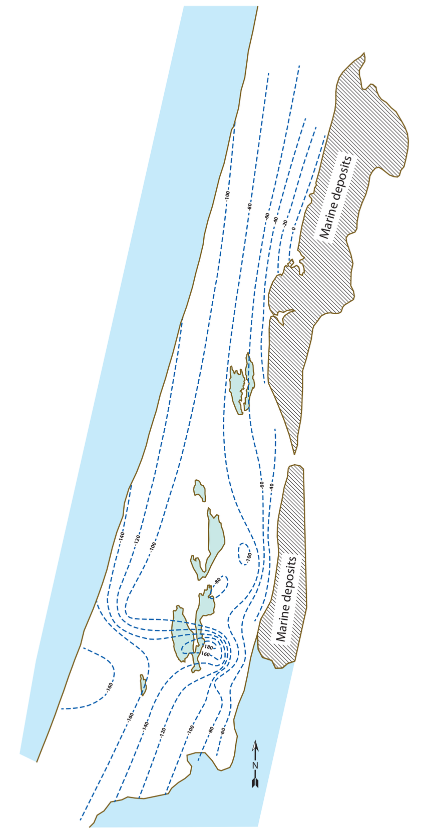


Figure 3. The contour interpretation by Robison (1973), which Jones (1990)* used to construct his steady-state MODFLOW simulation.

* Jones, M.A., 1992, Ground-water availability from a dune-sand aquifer near Coos Bay and North Bend, Oregon. USGS Open-File Report 90-563.

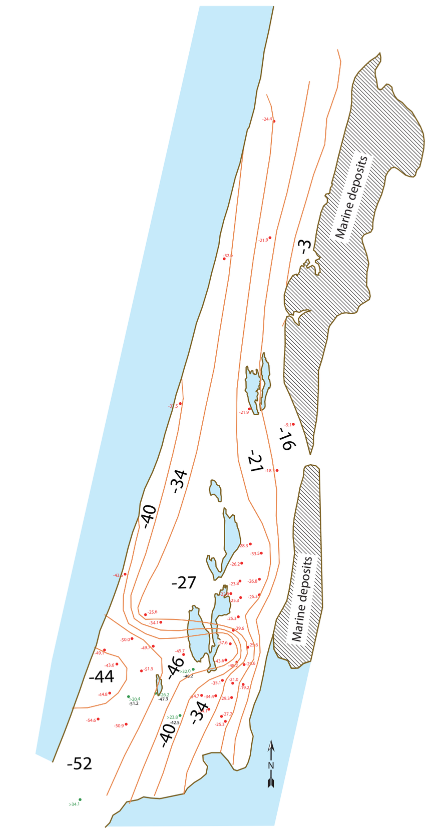


Figure 4. This is my first attempt to utilize the borehole bottom information in Figure 2 to set up inhomogeneities for a GFLOW, analytic-element simulation. The inhomogeneity boundaries follow the Robison (1973) contours, with the bottom elevations being the average of the two bounding contours. The small depressions beneath and north of Horsfall-Spirit lake were not modeled, since they do not add anything to the flow field. The bottom elevations shown are in meters below mean sea level. Convergence in this model was sometimes difficult to achieve, and calibration to the water table lakes did not work without unreasonably tweaking parameters outside the dune field.

The difficulty may have arisen because of the long, narrow nature of the inhomogeneities, and from the crossing of the boundaries of the inhomogeneities with the linesink strings of the lakes. Model run times were high, on the order of an half hour.

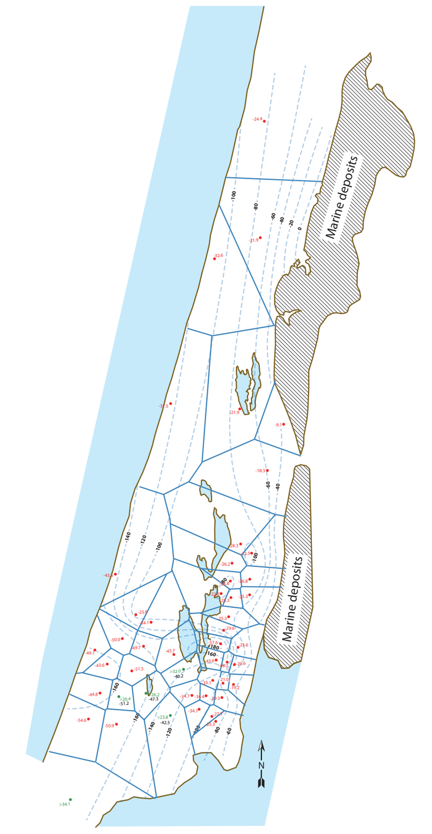


Figure 5. Robison's contour lines (Figure 3) are his interpretaion of the bottom elevation vis a vis the borehole data, and not necessarily correct. My next approach was to use Thiessen polygons; one per data point, and to give the polygon bottom value the bottom depth of its associated hole. For the drill holes that did not reach the bottom, an interpolation scheme was utilized, comparing the subject polygon to those immediately adjacent. Run times were very high; on the order of three or more hours, depending on the solver settings. Although some positive results were achieved, small changes in factors such as pumping or to recharge could cause the model to crash.

Another approach was needed.

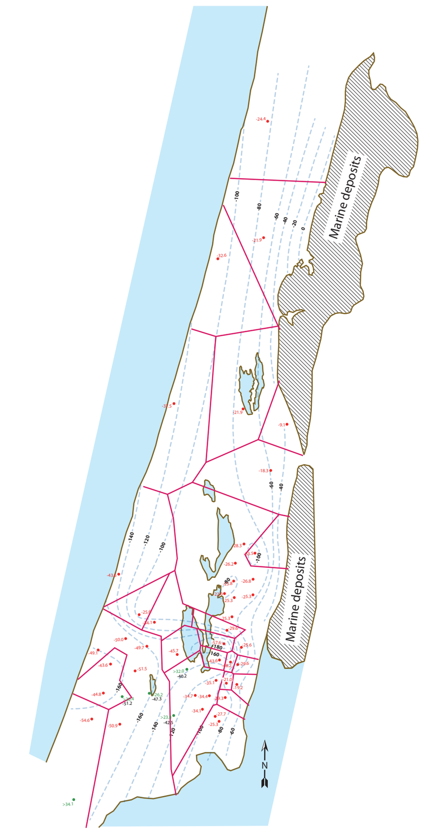


Figure 6. My first attempt at simplification was to group polygons who's bottom elevations were within 10% of each other. An improvement in run times was achieved, but they were still on the order of an hour or more; and sensitivity to crashing was not much improved.

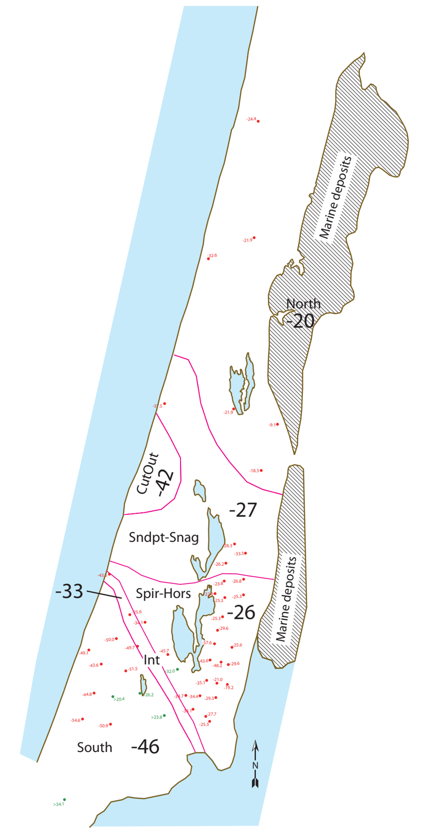
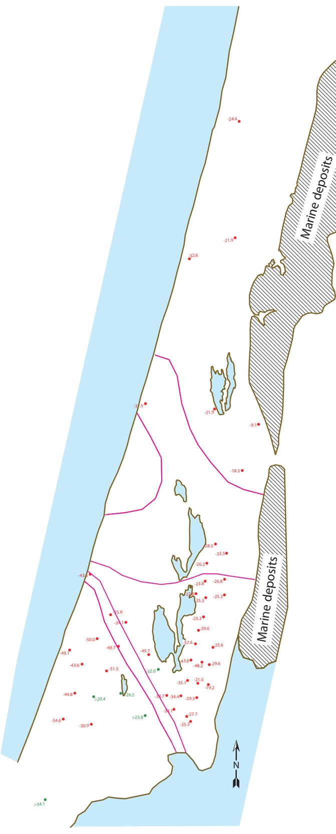


Figure 7. Finally I decided on a more statistical approach. I used six zones and gave each zone a value close to the average of the borehole data in that zone. The name of each zone is shown above. The zone labled "CutOut" was constructed without borehole data. Its presence in the model is solely interpretive. It is a construct that allowed the water levels of Spirit and Snag lakes to be approximately the same in the pre-pumping, steady-state simulation.

The model, as in prior scenarios, was calibrated to the water levels of the water table lakes at their normal high stand. I attempted to maintain the bottom elevation for each of the polygons within one half standard deviation of the mean of the borehole values within it. I did not use the borehole data that did not reach the base of the aquifer.

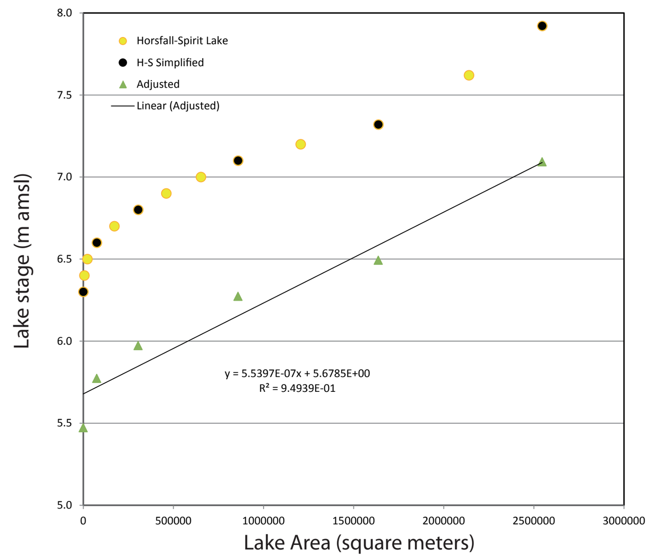
This model proved to be the most stable, responded well to the addition of pumping, and had run times on the order of three to six minutes, depending on solver settings.

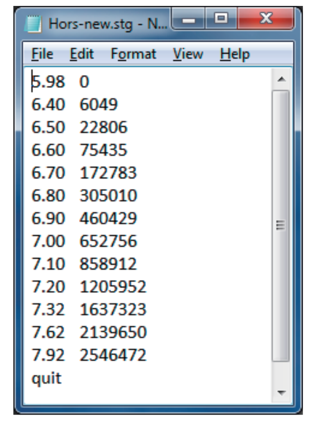


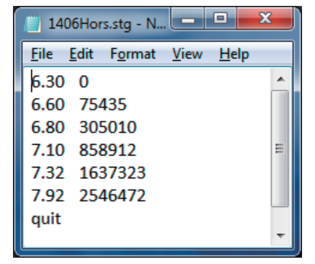
South	Intermd	Spir-Hors	Sndpt-Snag	North
50.0	43.6	26.8	37.5	18.3
51.5	25.6	23.8	28.3	9.1
49.1	49.7	22.6	26.2	21.9
43.6		25.3	33.5	32.6
44.8		25.3		21.9
54.6		25.3		24.4
34.1		29.6		
50.9		57.6		
		25.6		
		43.6		
		48.2		
		29.6		
		21.0		
		19.2		
		35.1		
		29.3		
		27.7		
		25.3		
		34.1		
		34.4		
		34.7		
		32.0		
		45.7		
		34.1		

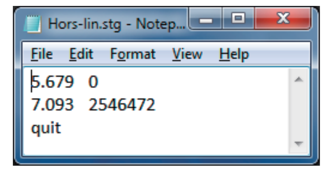
	X mean	s	n	X_m-s	X_m+s	X_m-0.5s	X_m+0.5s	X_m-2s	X_m+2s
South	47.33	6.42	8	40.91	53.75	44.12	50.54	34.49	60.17
Spir-Hors	31.98	9.42	25	22.56	41.4	27.27	36.69	13.14	50.82
Sndpt-Snag	31.38	5.11	4	26.27	36.49	28.825	33.935	21.16	41.6
North	21.37	7.69	6	13.68	29.06	17.525	25.215	5.99	36.75

The statistical approach shown in Figure 7 used the borehole bottom data shown at left. The standard deviation values are shown above. I was able to maintain model base values for each polygon within one half deviation of the mean.



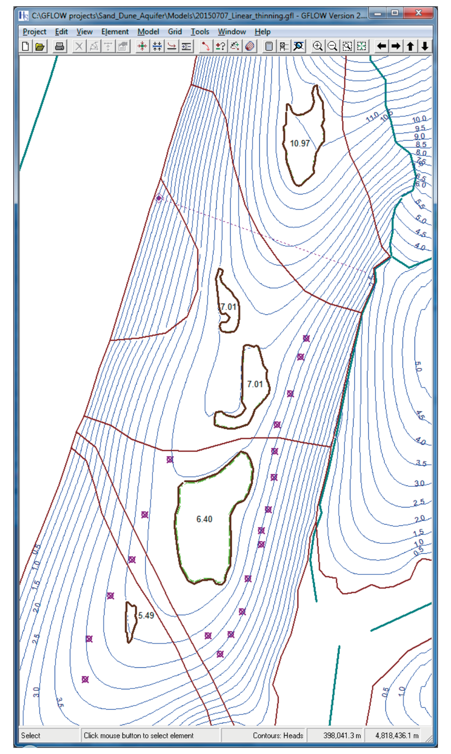






Also problematical was the detailed bathymetry supplied to me by Rich Miller, a research assistant at Portland State University in Portland, Oregon. Rich obtained the data personally in the field in 2014. I first attempted to use the complete, detailed bathymetry, as shown in the Hors-new.stg file above, which has a value for lake area at intervals of one tenth of a meter. These break points are represented on the graph above by the yellow dots. The model would give sometimes illogical results, including decrease in head with increased recharge. I determined that the stage versus area files may have been more complex than necessary, so I used the simpler 1406Hors.stg file above, shown as the black dots in the graph at left. It became plain with detailed analysis of the results, that problems were occurring at the break points. In the end I chose a linear approach, shown in the stage file Hors-lin.stg at right. The problem was solved and the model ran with very few crashes.

Evident in the graph at left is also a vertical displacement. This was a result of Rich informing me that the vertical datum was probably incorrect. The vertical land data was in NGVD27, while the data he gathered was in NAVD88. The difference is about a meter. This correction also stabilized the model.



Bluebill lake	5.49	5.522
Spirit-Horsfall lake	6.4	6.378
Sandpoint lake	7.01	6.959
Snag lake	7.01	7.068
Beal lake	10.97	11.016

These are the pre-pumping results for the final, most stable model. The target water levels, in meters above mean sea level, are highlighted above. The model water levels for these lakes is to the right. All are within a tenth of a meter. The model, at left, shows the one-half meter contours. The model exercise was successful, and now pumping effects may be evaluated.

Conclusion

Working with real data can be problematical. Modeling without data is generally far easier. The more data available, the more difficult the decisions that must be made as to what to do with it. Throwing out data because is looks like it doesn't match is unsatisfactory. Utilizing every data point independently is also unsatisfactory in the same way that using Thiessen polygons for precipitation in Florida is; just how representative of the entire polygone is that data point? Sometimes it works, sometimes it doesn't. In this case, grouping the points, averaging them, and using values within one-half standard deviation of the mean proved the most productive.